Sensorimotor Synchronization and Individual Differences in Intelligence: A Chronometric Perspective on Music Evolution

CHRISTOPHER JENKINS

Below are speaking notes as used in a dissertation defense presentation. For full details of the theoretical basis, methods, and results of this study, see:
jenkinsc.com/docs/d.pdf.

My research has looked at the process of synchronizing movement with an external beat ("sensorimotor synchronization" or SMS) and how measuring performance error in SMS might be a window into how musical rhythm could have evolved as a display of intelligence.

The musical beat is part of human nature

All cultures perform music with:
- An equal-interval ("isochronous") beat
- Beat groupings
- Percussion instruments
- Dance
- Strong emotional and cultural significance

All cultures have music that is structured around an equal-interval, or "isochronous," beat. And with a few rare exceptions, almost all musical styles within each culture have this as well. These are also universally grouped into a hierarchy of repetitions above the level of the beat, which are represented in Western music by bars and measures. All cultures use percussion instruments, dance along with musical performances, feel strong emotions in response to music, and assign music an important place in cultural rituals and social life.

The musical beat is part of human nature

In traditional societies:
◦ Group-based, inclusive
◦ Timing changes associated with emotional intensity

Reliably develops in early infancy:
◦ Recognition of changes to isochronous sequences
◦ Increased attention to strong vs. weak rhythmic beats

Ringo Starr, caveman/percussionist

Where music has been studied in societies that live in traditional, kin-based groups, solo musical performance is rare. Music is more often performed in large group gatherings, and it's broadly inclusive, without the clear division between the performers and the audience that you see in mainstream Western culture.

One very common feature of music, including these kinds of group performances, is for in increase in tempo to occur as emotional intensity increases throughout a performance.

Humans reliably develop ability to interpret the beat in music, are motivated to synchronize movements with it. Infants can recognize isochronous sequences and develop expectancies around them, so that disturbances are noticed. This is particularly the case for beats that adults hear as "strong," in the way that makes people want to clap on particular beats.

Human universals like the musical beat suggest the need for an explanation in the context of human cognitive evolution.

Why do humans like to create beat-based rhythmic performances?

By-product of auditory attention biases; self-stimulation technology (Steven Pinker)

Communication to coordinate group-level motivational states (Stephen Brown)

Group-selection for amplification via synchronization (Bjorn Merker et al.)

Demonstrating group qualities to other groups (Hagen & Bryant)

Mating display (Darwin), Sexually-selected fitness indicator (Miller)

There has been no shortage of speculation about this, but relatively little in the way of empirical research programs derived from evolutionary hypotheses.

Stephen Pinker is well known among musicologists for calling music "auditory cheesecake" -- in other words, a self-stimulation technology. Stephen Brown and other theorists in ethnomusicology have focused on group-level benefits that music might have, such as the idea that it allows motivational states to be coordinated across a group.

Hypotheses specifically about rhythmic entrainment in music include:
--that it benefits groups by amplifying synchronized sounds and maximizing the range at which migrating women could notice the existence of a group
--a paper by Hagen and Bryant that suggested that groups that can be conspicuously synchronous might be perceived as more formidable.

But here, I'm focusing on the idea that musical virtuosity could be a reliable indicator of individual traits, which, as Geoffrey Miller has argued, could have evolved through sexual selection.

How could beat-based rhythmic synchronization signal traits?

Music in general (Miller):

◦ Physical endurance, strength, health, motor coordination
◦ Self-confidence, conscientiousness, free time / sexual availability
◦ Creative intelligence
◦ Neurological differences in complex sequencing ability

Geoffrey Miller's chapter on musical fitness signaling suggested a range of mechanisms by which musical performance might reliably indicate fitness. An impressive musical performance might require physical endurance, extended practice time, confidence, and creativity. Rhythmic accuracy in particular might demonstrate the functionality of particular neurological differences that are correlated with more general differences in general intelligence and genetic fitness.

How could beat-based rhythmic synchronization signal traits?

Time-sensitive indicators of psychometric intelligence ("Mental Chronometry" research):

Four-choice reaction time:
r around .5 (Jensen, Deary)

Interval discrimination performance:
r = .43 (Helmbold, et al., 2006)

Self-paced equal interval tapping ("ISIP"): short-term variability:
r = -.44 (Madison, et al., 2009)

Error in synchronization with an isochronous beat ("SMS"):
r = -.12 to .29 (Lorås, et al., 2013)

Research on how timed tasks relate to psychometric intelligence go back a long way in psychology. The reaction-time/intelligence relationship (studied by Arthur Jensen and Ian Deary) is by far the best established of these. But other tasks start to make a connection to musical rhythmic ability seem more plausible, such as discrimination between time intervals of different lengths and variability in motor actions performed in repeating intervals. Researchers have begun comparing these to individual differences in intelligence.

Temporal information processing

Relationships among temporal task performance (Helmbold, Troche, & Rammsayer, 2007):
◦ Detecting differences in time intervals (same/different)
◦ Judging precedence in near-simultaneous events (e.g., light and sound)
◦ Detecting changes in rhythmic stimuli (yes/no response)
◦ Choice reaction time tasks

SEM analysis: a general temporal-performance factor accounts for the relationship between reaction time tasks and psychometric intelligence

A few researchers have investigated whether the association between reaction time and intelligence may be due to a more general effect. A 2007 study found that performance across a series of time-sensitive tasks. A general "temporal performance" factor accounted for the relationship between reaction time tasks and psychometric intelligence. So timing resolution could be a fundamental determining factor in reaction time and time-sensitive task performance.

In group-based musical performance

Time-sensitive tasks:
- Inferring a target beat among overlapping stimuli
- Quickly responding to one's own peripheral / skeletomuscular inaccuracies
- Quickly responding to unintentional and intentional changes in the overall beat produced by other group members

Could accurate accommodation of errors and changes, resulting in precise synchronization of actions with other musical performers, be a strong, human-perceptible cue of general intelligence?

What are some time-sensitive reactions required for rhythmic performance?
--Responding to changes in a stimulus beat, is never perfectly isochronous in real performance settings

Experimental research on basic beat-keeping processes

Isochronous serial interval production (ISIP)
◦ Self-paced

Sensorimotor synchronization (SMS)
◦ Paced by equal-interval pulse

There were two kinds of rhythmic performance this research focused on measuring.

Isochronous serial interval production (ISIP)

Performers attempt to keep a steady, equal-interval beat at a given tempo

Self-correcting processes

Analyze variability in "inter-tap intervals" (ITIs)

ISIP tasks start with entrainment to an external stimulus. (…) Variability in interval durations – the time between performed taps – across the task is analyzed as a means of modeling the cognitive mechanisms behind movement timing. Tapping rate is maintained by a self-correcting timekeeping process without continuous feedback. In the social structure of ancestral music performance, this kind of performance was probably seen by fewer people, and performed less often, making it a poorer candidate as a signaling behavior than synchronized performance.

[The process is self-correcting: performers compare time intervals with recent intervals in working memory. Interonset interval (IOI) varies from about 200 ms to 3000 ms -- tapping faster than this is not mechanically possible for most people, and tapping more slowly seems to invoke a different process, such that some regularities in the findings within this range tend to break down for slower rates.]

ISIP drift
(IOI = 750 ms)
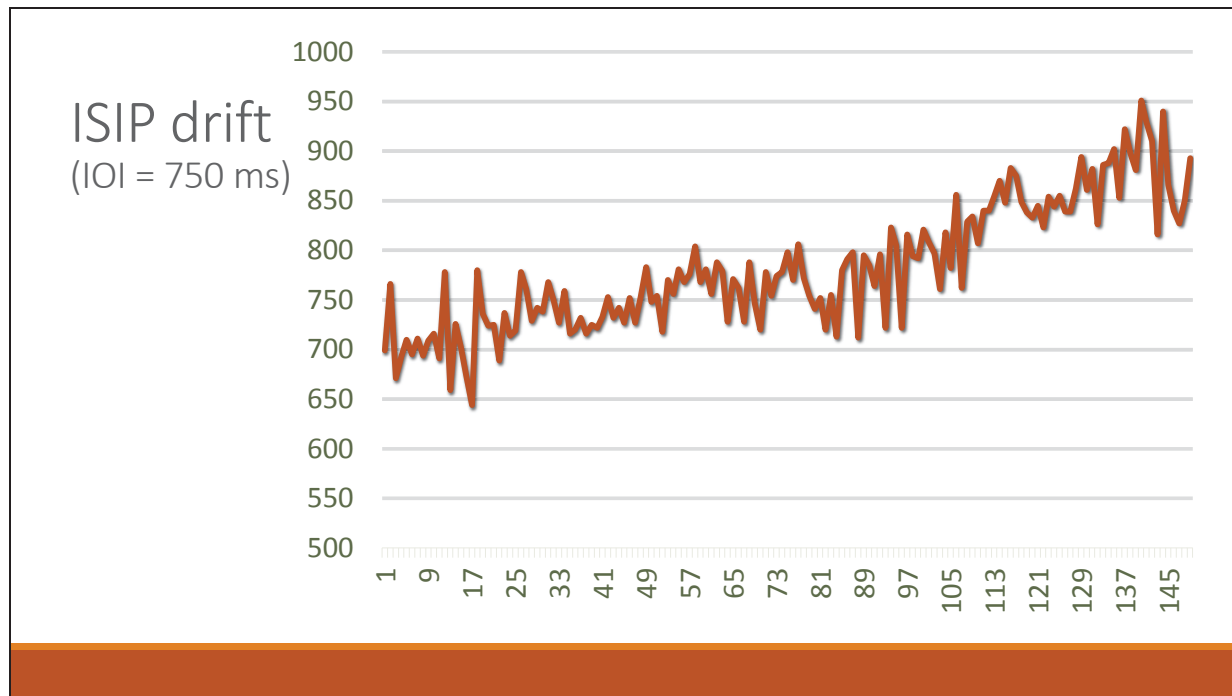
Long-range correlation among the interval durations.

In this study, I adjusted for this drift effect by analyzing the deviations of each interval from a moving mean of the previous four intervals.

Individual differences in ISIP

Previous research: Raven's Progressive Matrices test (brief IQ proxy)

Local variability & RPM score:

  highest correlation when stimulus IOI is about 800 ms (r = about −0.4)

Drift variability & RPM score: the highest correlation is found when stimuli occur at about 500 ms (r = about −0.4)
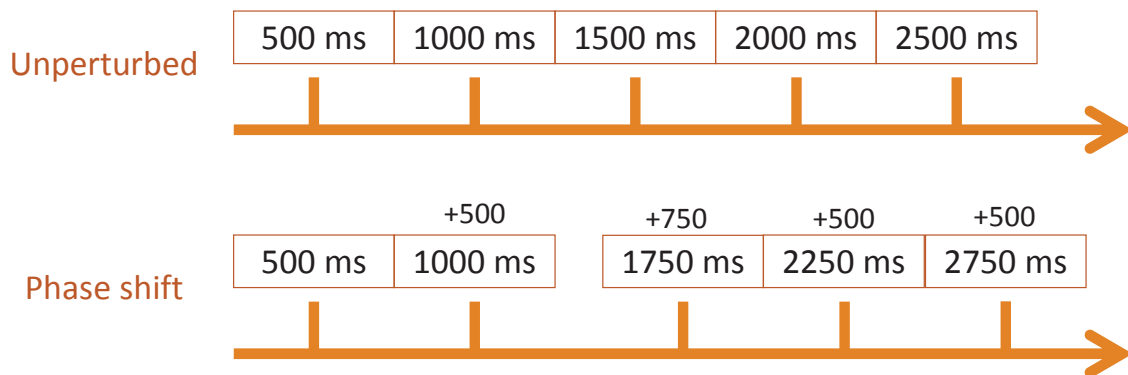
Sensorimotor Synchronization (SMS)

Performers entrain to a steady, equal-interval ("isochronous") set of auditory stimuli

Processes of correction for phase and period changes

Auditory dominance

Performers entrain to a steady, equal-interval ("isochronous") set of auditory stimuli. Interonset interval (IOI) varies from about 200 ms to 3000 ms. (Tapping faster than this is not mechanically possible for most people, and tapping more slowly seems to invoke a different process, such that some regularities in the findings within this range tend to break down for slower rates.) Possible evidence of design relevant to music– synchronized performance preferentially attunes to *auditory* over *visual* cues.
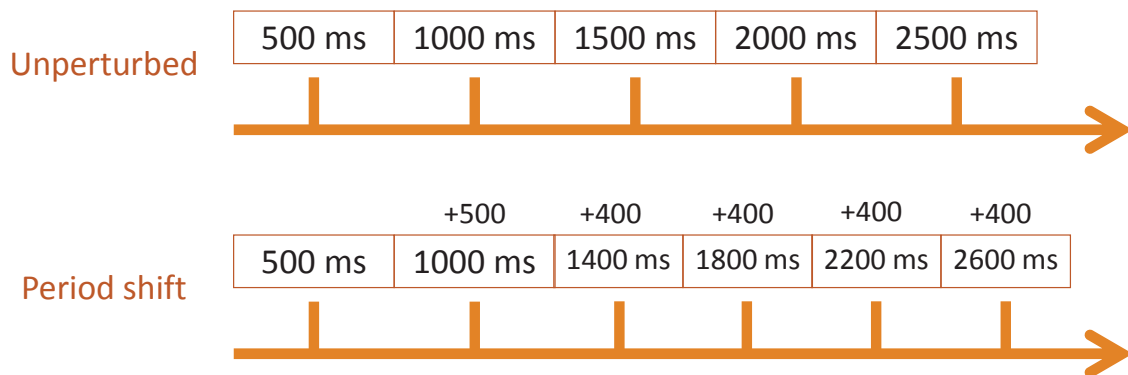
Phase shift in pacing stimuli

| Unperturbed | 500 ms | 1000 ms | 1500 ms | 2000 ms | 2500 ms |

| | | +500 | | +750 | +500 | +500 |
| Phase shift | 500 ms | 1000 ms | | 1750 ms | 2250 ms | 2750 ms |

A phase shift describes a set of intervals where the series moves forward or backward in time but the interval between stimuli remains the same after the shift. In this case, interonset interval (IOI) is 500 ms.

Participants are generally aware when they are accommodating a phase shift, and the process tends to be interfered with when competing cognitive demands are present, so it's considered to be primarily under executive control.

Period change in pacing stimuli

Unperturbed	500 ms	1000 ms	1500 ms	2000 ms	2500 ms

		+500	+400	+400	+400	+400
Period shift	500 ms	1000 ms	1400 ms	1800 ms	2200 ms	2600 ms

A change in the period of stimuli is just another way of saying that the interonset interval increases or decreases. So rather than shifting, this could be thought of as in increase or decrease in tempo.

This process can occur without awareness, and tends not to be interfered with by competing cognitive demands, so it's considered to be more unconscious and less dependent on executive control than phase shifts.

Individual differences in SMS error

Among musicians:

Working memory (WAIS-III digit span and letter-number sequencing) predicts accuracy in synchronization with patterned rhythms (Bailey and Penhune, 2010)

In a college sample (nonmusicians and musical amateurs):

Cognitive ability predicts performance in brief SMS tasks (IOIs between 500 and 950 ms); *r* about 0.3 (Lorås et al, 2013)

A study of musicians found that equal-interval tapping along with musical rhythms was related to differences in working memory.

Only simple SMS tasks, *without* phase or period shifts, have been compared to general intelligence in previous research.

Single-stimulus, isochronous SMS/ISIP: reliability and validation (2008)

Can SMS and ISIP be reliably measured in a UNM student sample with my software/hardware?

Do quantifiable differences in SMS/ISIP error relate to human-perceptible differences in the quality of improvised creative performances on drum pads?

Low levels of musical experience in sample; removed "drummers" from analysis

My first study, which I'll just touch on very briefly, addressed two main questions.

(…)

(this first iteration used custom software running on a Windows PC to time the stimulus delivery and to record the participant response timing. That may have introduced a lot of noise, so a better setup was used for the newer study, which I'll talk about in a minute.)

Single-stimulus, isochronous SMS/ISIP: reliability and validation (2008)

Task battery:
◦ Simple SMS tasks at IOIs of 250, 500, 750 ms; alternating-hands; simple patterns (2:1, 3:1)
◦ ISIP tasks at IOIs of 250, 500, 750 ms
◦ Creative performance tasks: with & without metronome-like stimuli

5 in-lab raters: accuracy, creativity of performance

Simple (unchanging) SMS tasks where the stimuli had interonset intervals of (…)

-- 2 : 1 pattern, (that is, "tap tap pause, tap tap pause")

In a sample of college students, where most in this sample had very little musical experience.

SMS task types included patterned and alternating-hands tasks, not considered in present study.

Relationships between Quantitative Outcomes
and Creative Performance Ratings (Pearson r)

	Self-paced creative improvisation		Synchronized creative improvisation	
	Accuracy	Creativity	Accuracy	Creativity
SMS error (10-task composite)	−.14	.00	−.34*	−.43**
ISIP Local error (8-task composite)	−.31*	−.30*	−.39**	−.40**

*p < .05, **p < .01

As expected, it appears to be easier to judge whether a performer is capable of rhythmic accuracy, when observing a creative performance that is set to the equivalent of a metronome. (As opposed to one that's entirely self-paced).

On the other hand, the component of rhythmic ability that allows one to perform well on the self-paced ISIP tasks does appear to be perceptible in a self-paced creative improvisation.

What aspects of group-based beat synchronization weren't measured?

Change over time to the IOI ("period change")

Adjustments to timing that don't affect the IOI ("phase shift")

The complexity of a background beat that must be inferred from the performances of multiple imperfectly-synchronized performers

That provided a preliminary validation of the idea that we could take quantitative measurements of individual differences that might have some real consequences for social perception of musical ability.

This first study didn't measure elements that might be important ways that synchronization invokes timing ability – and therefore intelligence– when performed against an error-prone social group of other performers.

Research goals

Stimulus timing variations and predictors of error

◦ IQ, personality factors, musical experience or skill

◦ Replicating finding: IQ as a predictor of ISIP error (Madison 2008)

◦ Replicating finding: IQ as predictor of (unchanging) isochronous SMS error (Loras et al, 2013)

In order to address this, my second study varied stimulus timing and measured traits that might predict synchronization error: IQ, personality, and musical experience/skill.

Research goals

Stimulus timing variations and predictors of error

◦ Response to change in SMS entrainment sequences: predictable period changes, unpredictable phase shifts

◦ Do tasks that require quick adaptation to changes relate more strongly to IQ?

◦ Is the IQ/error relationship independent of musical experience or skill?

The new variations used here varied timing in the stimuli, so that quick adaptation responses would be necessary to maintain accurate synchrony: (a) steady, predictable changes in the period, and (b) sudden, unpredictable changes to the phase.

Sample qualifiers

Right-handed / ambidextrous

English fluency

Self-rated musical instrument skill attainment:
Not "professional-level"

Right-hand-dominant or ambidextrous (removed two participants who didn't meet these criteria despite a stated requirement in recruitment materials).

English fluency (removed six participants who didn't meet this criterion)
Advertisement: "You are a fluent English speaker and have been primarily English-speaking since childhood"
"Professional-level" self-rated musical instrument skill attainment (removed two) - keeps these results comparable with other studies not sampling professional musicians; and with only two participants reporting this value, musicians wouldn't be adequately represented in any case.

N = 107 before removals
N = 97 after removals

Measures

Between-subjects measures
(as continuous predictors of SMS & ISIP error):

General intelligence: WASI-II FSIQ-2

Big Five personality factors: BFI-44

Musical performance skill self-rating
(Performance experience, practice time, achieved skill level: instruments, singing, dance)

I'll describe the sets of between-subjects measures in turn before getting to the rhythmic task manipulations that were used.

THE Wechsler Abbreviated Scale of Intelligence "Full-scale IQ 2" measure, which is a brief approximation of the full-scale IQ score from the full WAIS-4 battery.

Correlation between this brief measure and the full WAIS-4 IQ scale is about 0.84.

This consisted of a vocabulary test…

Copyrighted / protected test material
removed from this slide.

a vocabulary subtest, where participants were asked to describe what a word means and were scored on a two-point scale for each,

Copyrighted / protected test material
removed from this slide.

And a matrix reasoning subtest, where participants completed visual patterns.

Participants in this sample had a distribution of IQ outcomes that was close to the general population, with a mean IQ score of about 104.

(103.6 and a standard deviation of 10.)

Measures

Between-subjects measures
(as continuous predictors of SMS & ISIP error):

General intelligence: WASI-II FSIQ-2

Big Five personality factors: BFI-44

Musical performance skill self-rating
(Performance experience, practice time, achieved skill
level: instruments, singing, dance)

The big five personality factors were measured with a free scale called the Big Five Inventory 44.

1	2	3	4	5
Disagree Strongly	Disagree a little	Neither agree nor disagree	Agree a little	Agree strongly

I am someone who…

1. _____ Is talkative

2. _____ Tends to find fault with others

3. _____ Does a thorough job

4. _____ Is depressed, blue

5. _____ Is original, comes up with new ideas

6. _____ Is reserved

7. _____ Is helpful and unselfish with others

8. _____ Can be somewhat careless

9. _____ Is relaxed, handles stress well

23. _____ Tends to be lazy

24. _____ Is emotionally stable, not easily upset

25. _____ Is inventive

26. _____ Has an assertive personality

27. _____ Can be cold and aloof

28. _____ Perseveres until the task is finished

29. _____ Can be moody

30. _____ Values artistic, aesthetic experiences

31. _____ Is sometimes shy, inhibited

This has a fairly typical format, with a five-point agreement scale for self-descriptions that were summed to scale totals.

Measures

Between-subjects measures
(as continuous predictors of SMS & ISIP error):

General intelligence: WASI-II FSIQ-2

Big Five personality factors: BFI-44

Musical performance skill self-rating
(Performance experience, practice time, achieved skill
 level: instruments, singing, dance)

…and I administered a set of questionnaire items to get information about whether participants have ever spent time practicing an instrument, singing, or dancing, and if so, how much time they have spent playing in the past year and how skilled they have become.

The only measure of performance skill that was well-distributed across participants was a single item asking them to rate their musical instrument skill level, so the analyses focused on that item.

(Paying special attention to this because it ended up being a very important item.)

Self-rated Instrument Skill Level

3. Have you ever taken music lessons, or spent time learning to play a musical instrument? (Please include any informal practice, and any lessons you might have received in your life, including early schooling.)

(Yes)   (No)

IF YES:

    a) What instrument(s) have you learned to play?

    (Please list below. If more than one, please start with the instrument that you have the most experience with.)

Coded as zero unless contradicted in part (b)…

b) Which of these best describes the highest level of skill you have achieved in musical performance on any instrument?

(1) (Beginner)

(2) (Able to play very simple material)

(3) (Able to learn and play complex material after extensive practice)

(4) (Able to play complex material comfortably, with creative expressiveness)

(5) (Professional-level technical skill)

Two participants in the sample selected "5," professional-level technical skill, and I decided to remove them from all analyses.

(A lot of research on musicians groups participants into musical professionals and nonprofessionals, and I thought it would be useful to be able to describe the study population as limited to musical amateurs and nonmusicians. In retrospect it would also have been a good idea to ask about whether people would actually self-describe as professional musicians.)

Rhythmic performance task set

Twelve SMS tasks combined levels of:

(3) Presence of **phase and period changes** across the task

So those were the major between-subjects, continuous measures. The study manipulated a number of elements of synchronization tasks within-subjects as well.

There were three SMS task variations used to measure the effects of phase and period changes versus unperturbed, isochronous stimuli.

Rhythmic performance task set

Twelve SMS tasks combined levels of:

(3) Presence of **phase and period changes** across the task

(2) **Complexity/ambiguity** of each beat's target **stimuli**

There were two variations used to measure the effects of two factors…

Rhythmic performance task set

Twelve SMS tasks combined levels of:

(3) Presence of **phase and period changes** across the task

(2) **Complexity/ambiguity** of each beat's target **stimuli**

(2) **Interonset interval** of stimuli (500 ms, 800 ms)

…and these were each tested with stimulus intervals of 500 and 800 milliseconds.

Rhythmic performance task set

Twelve SMS tasks combined levels of:

(3) Presence of **phase and period changes** across the task

(2) **Complexity/ambiguity** of each beat's target **stimuli**

(2) **Interonset interval** of stimuli (500 ms, 800 ms)

Two ISIP tasks (500 ms & 800 ms IOI)

Simple self-paced tapping tasks were run at these IOI levels as well. All participants received all of the levels of these within-subjects factors in partially randomized orders. We'll see exactly what each of these task differences means in a minute, but first I'll talk a bit about the apparatus used to run the tasks and the structure of the data that they produce.

Apparatus:

Roland PD-8 Drum Trigger

Participants completed the tasks by tapping with a drumstick held in the right hand on a Roland PD-8 drum trigger. This is just a rubber pad with a piezoelectric vibration sensor. It's connected with an audio cable to…

Roland TD-6 Percussion Sound Module

…the "drum module,", which reads that vibration sensor every 0.5 milliseconds, and registers a tap whenever it spikes above a given threshold. So this is the device that's directly connected to the drum pad and headphones used by participants.

Roland TD-6 Percussion Sound Module

All of the stimuli that were used are part of the drum module's built-in set of percussion sounds.

The simplest part of that is that it gives feedback from taps, just like if it were being used for regular drumming practice. So the sound of tapping on the right-hand drum pad was like a tom:

[Audio playback]
All of the tasks used that same feedback.

Roland TD-6 Percussion Sound Module

(to headphones)

(from trigger pads)

MIDI input/output

It has output and input ports for sending and receiving MIDI commands. MIDI is a protocol used for electronic music that allows one device tell another device to immediately turn a note on or off. The module sends out a MIDI command whenever a drum is tapped. (So the drum module doesn't measure performance timing or meter out when stimuli should be played – it sends signals to the next piece of equipment to be timed.)

Arduino Mega 2560 R3

The timing of stimulus output and performance measurement used a dedicated circuit on an Arduino platform, which is basically just a microcontroller chip, the Atmel ATmega2560, with some components built around it for input and output and supplying power. The microcontroller runs code written in C and allows sub-millisecond timing for metering out stimulus commands and recording the timing of participant input.

Stimuli out: Arduino → Module → headphones

To drum module (MIDI commands)

"Start task #15"...

Run programmed tasks

Meter out stimuli

(Note stimulus output time for comparison with input)
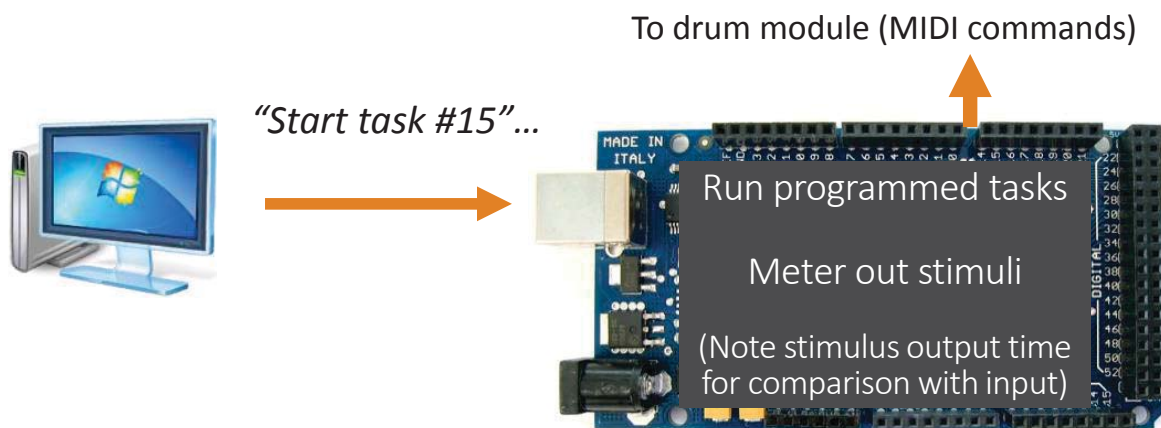
The Arduino was programmed to run the tasks autonomously, except that it waited to receive a command to start each task. A program running on a Windows PC was used to keep track of the task orders assigned to each participant, and to send the command to start the next task as needed.

Stimuli out: Arduino → Module → headphones

Stimulus MIDI commands
metered out
(e.g., at 500-ms intervals)

So for the simpler synchronization tasks, the Arduino sent commands at the specified intervals to play metronome-like ticking sounds out to participants' headphones, and those served as the entrainment stimuli.

[audio]

Data in: Module → Arduino → CSV

Tap detected: Send MIDI command

Save timestamp in memory
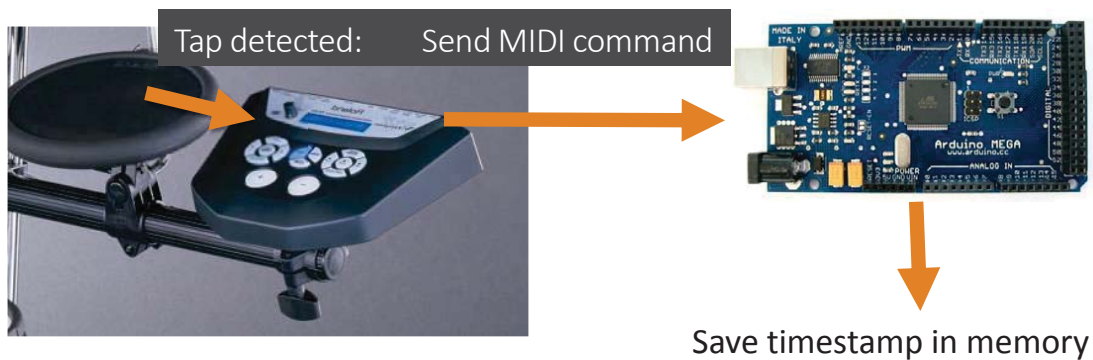
The same Arduino device was used to measure participants' performance timing. When it registers a tap, in addition to playing that tom sound to the participant, it tells the Arduino that a tap has occurred.

Data in: Module → Arduino → CSV
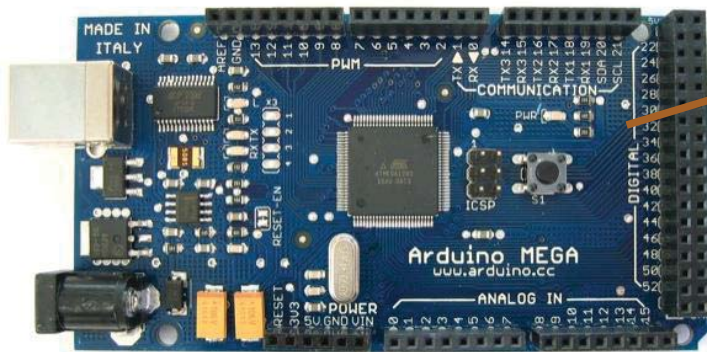
High-resolution timestamps added as input is received

Structure as CSV lines (task name, channel, instrument, velocity, timestamp)

Send CSV text to PC

The microcontroller adds a timestamp and saves it in memory, and later sends out a set of recorded data points out to the connected PC.

Raw recorded data (single-stim. SMS task)

taskName	i	channel	pitch	velocity	microseconds
T1_SMS_5	0	5 *[stimulus]*	42 *[tick]*	127	588064100
T1_SMS_5	1	5	42	127	588563816
T1_SMS_5	2	1 *[performance]*	48 *[tom]*	35	588793524
T1_SMS_5	3	5	42	127	589063512
T1_SMS_5	4	1	48	43	589137392
T1_SMS_5	5	1	48	42	589558604
T1_SMS_5	6	5	42	127	589564248
T1_SMS_5	7	1	48	45	590050744
T1_SMS_5	8	5	42	127	590063944
…	…	…	…	…	…

Some performance taps will precede a target stimulus, and others will follow it…

SMS target/response comparisons

channel	milliseconds	Beat	Absolute Deviation	Percent Deviation
(stimulus)	0	0		
(stimulus)	499.716	1		
(performance)	729.424	1	+229.71 ms	**+45.9%**
(stimulus)	999.412	2		
(performance)	1073.292	2	+73.88 ms	**+14.8%**
(performance)	1494.504	3	-5.64 ms	**-1.1%**
(stimulus)	1500.148	3		
(performance)	1986.644	4	-13.20 ms	**-2.6%**
(stimulus)	1999.844	4		
…	…	…	…	…

So we can categorize the taps into beats, and compare to the appropriate stimulus. This produces positive or negative asynchronies. On average, across studies, participants tend to tap a little before each stimulus, which is a "negative asynchrony."

SMS target/response comparisons

Percent Deviation

+45.9%

+14.8%

-1.1%

-2.6%

…

→ Standard deviation:

SDasy (as percent of IOI)

Finally, after some filtering procedures, the variability in asynchronies is calculated, and this serves as the outcome measure. I'll describe these in terms of a percentage of stimulus intervals, since the amount of error tends to scale up linearly with the interval duration. (A natural log transformation was applied to this prior to modeling)

ISIP data

Next, we'll see what data from the unpaced interval production tasks looks like.

Recorded data (ISIP task)

taskName	i	channel	pitch	velocity	milliseconds
ISIP_5	82	1	48	57	300503.3
ISIP_5	83	1	48	49	301015.4
ISIP_5	84	1	48	52	301543.1
ISIP_5	85	1	48	52	302048.9
ISIP_5	86	1	48	46	302579.8
ISIP_5	87	1	48	54	303105.6
ISIP_5	88	1	48	51	303607.6
ISIP_5	89	1	48	56	304120.3
ISIP_5	90	1	48	51	304639.0
ISIP_5	91	1	48	56	305137.4
ISIP_5	92	1	48	54	305673.9
…	…	…	…	…	…

Here, there are no stimuli, so we just look at the time intervals between
each performed tap.

milliseconds	interval (ms)
300503.3	
301015.4	512.1
301543.1	527.7
302048.9	505.8
302579.8	530.9
303105.6	525.8
303607.6	502.0
304120.3	512.7
304639.0	518.7
305137.4	498.4
305673.9	536.5
...	...

But if we calculated a standard deviation of these, it wouldn't tell us much about the inaccuracies that an observer might perceive moment-to-moment in a person's performance. Most of that would be due to the drift effect.

milliseconds	interval (ms)
300503.3	
301015.4	512.1
301543.1	527.7
302048.9	505.8
302579.8	530.9
303105.6	525.8
303607.6	502.0
304120.3	512.7
304639.0	518.7
305137.4	498.4
305673.9	536.5
…	…
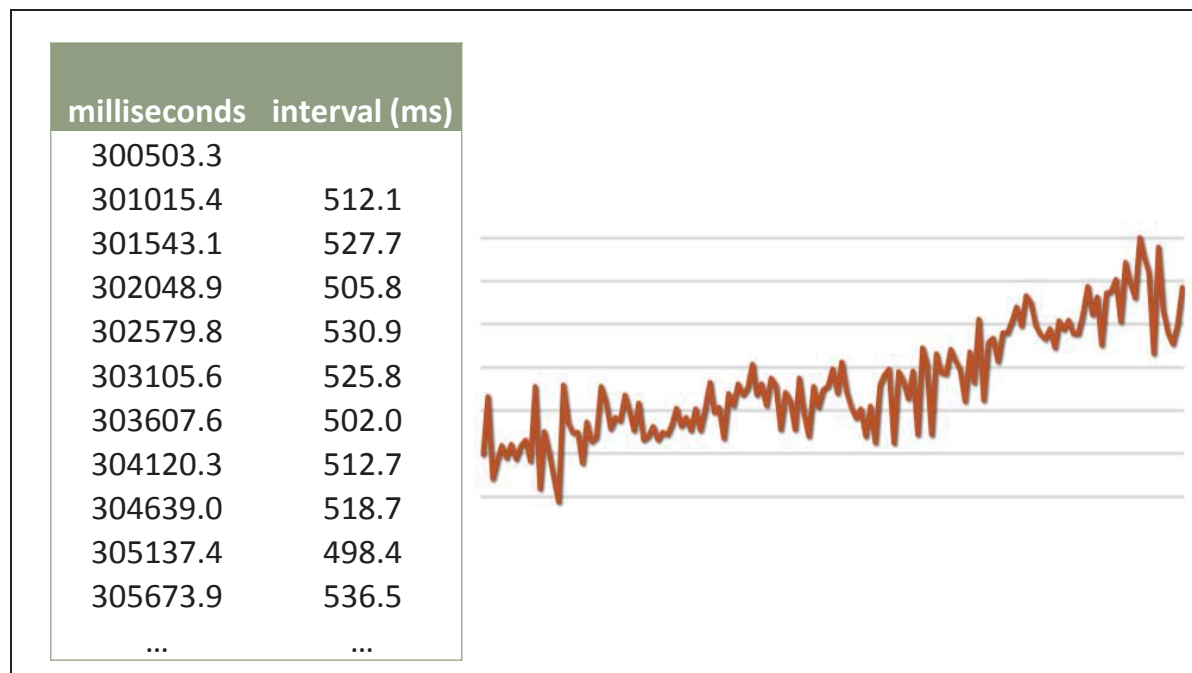
So to remove the longer-term trends in the interval durations…

interval (ms)	moving 4-mean	absolute local deviation
512.1		
527.7		
505.8		
530.9		
525.8	519.1	+6.7
502.0	522.5	-20.5
512.7	516.1	-3.4
518.7	517.8	+0.9
498.4	514.8	-16.4
536.5	508.0	28.5
…	…	…

Calculate variability, represent as a percentage of participant's mean interval duration: "**ISIP Local error**"

I looked at each interval's deviation from the mean of the preceding four intervals. This was referred to as ISIP Local error. (A natural log transformation was applied to each of the two local error measures prior to modeling.)

Within-subjects factors (3 x 2) for SMS task set:

    Timing changes

    Stimulus target ambiguity

Separate analysis: ISIP local error

These two major task types had a number of variations that were tested within subjects. The set of synchronization tasks looked at effects of two factors. (…)

First, we'll look at these timing change variations…

Within-subjects factors (3 x 2) for SMS task set:

    Timing changes

    Stimulus target ambiguity

Separate analysis: ISIP local error

These two major task types had a number of variations that were tested within subjects. The set of synchronization tasks looked at effects of two factors: (…)

First, we'll look at these timing change variations…

WS factor: Timing change type

Isochronous tasks
 ◦(each at 500 & 800 ms IOI)

Phase-shifting tasks
 ◦(each at 500 & 800 ms IOI)

Linear period-changing tasks
 ◦(from 500 to 800 ms; from 800 to 500 ms)

- No changes, which is a basic isochronous SMS task. (…)

Timing changes	Interonset Interval	1st: Single-stimulus SMS	2nd: Grouped-stimulus SMS
Isochronous	500ms		
	800ms		
Phase shifts	500ms		
	800ms		
Linear change	500 ms -> 800 ms		
	800 ms -> 500 ms		

Here's what the within-subjects design looks like for SMS tasks. We'll start with some examples from the single-stimulus tasks, which were administered first.

Isochronous task (single-stimulus)

Timing changes	Interonset Interval	1st: Single-stimulus SMS	2nd: Grouped-stimulus SMS	3rd: ISIP
Isochronous	500ms	🔊	🔊 (demo)	🔊
	800ms	🔊	🔊	
Phase shifts	500ms			
	800ms	🔊		
Linear change	500 ms -> 800 ms			
	800 ms -> 500 ms		🔊 🔊	

Single-stimulus tasks used a metronome-like ticking sound. These were similar to tasks that have been run in many studies before, including the studies that looked at relationships with intelligence.

Grouped-stimulus tasks were meant to introduce an element of group-based human musical performance. The variability in timing exaggerated from what would normally be seen in a social setting– it was meant to introduce enough difficulty to increase error variability but not enough that participants would fail to complete the task.

ISIP task:
Entrainment period (~40 intervals)
Measurement period (continue for about 1m10s)

Isochronous (ex. 800 ms, single-stimulus)
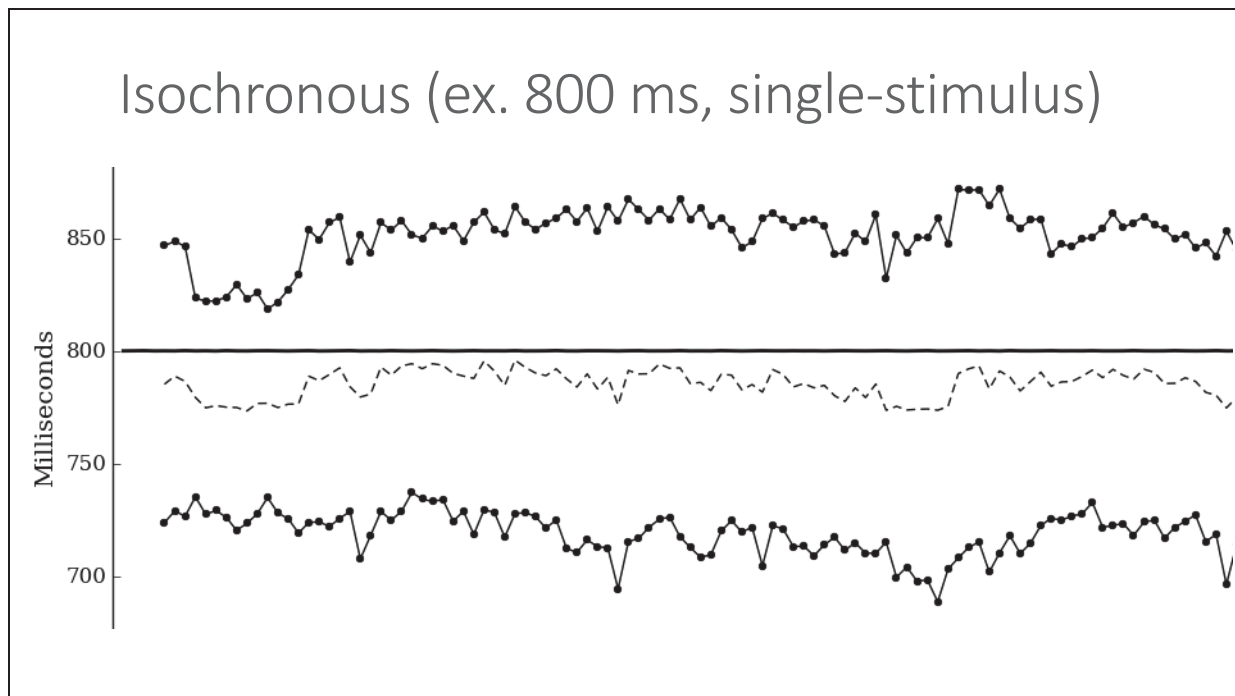
So for this task, the interval between each stimulus was constant at 800 milliseconds across the whole task: that's the solid black line.

Incidentally, the dotted line and the lines with circle markers show the mean and standard deviation of participants' actual deviations from the targets for each beat across the task.
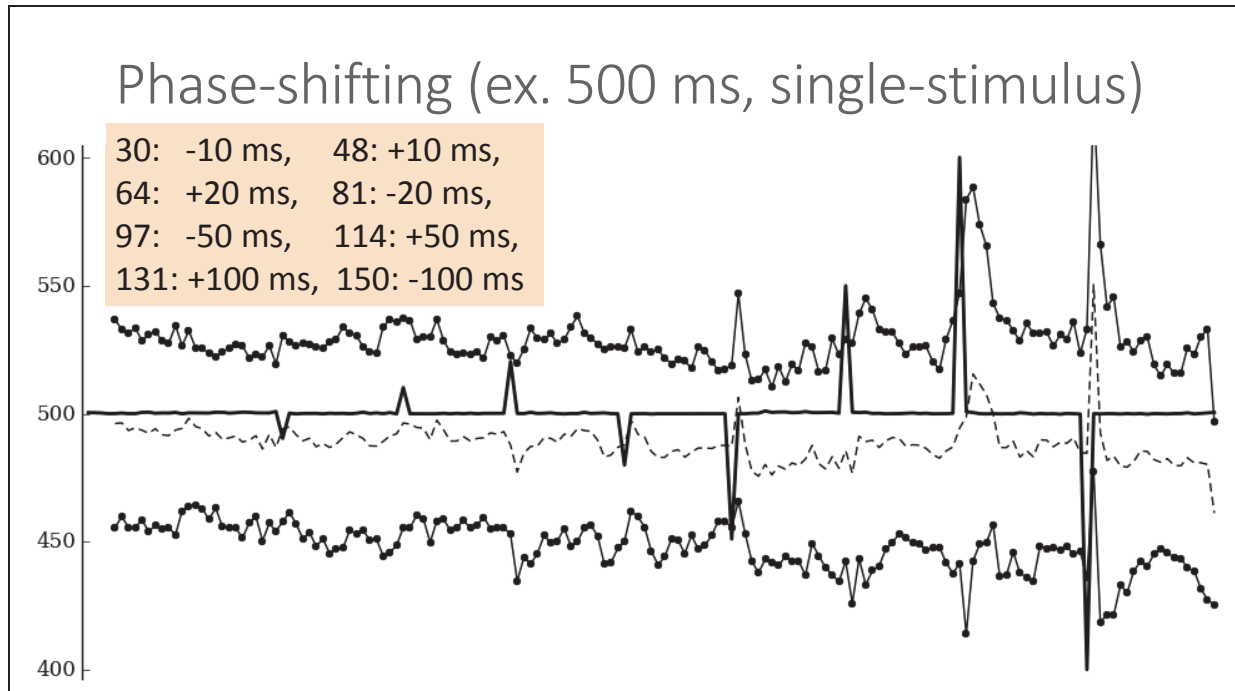
Phase-shifting task (single-stimulus)

Timing changes	Interonset Interval	1st: Single-stimulus SMS	2nd: Grouped-stimulus SMS	3rd: ISIP
Isochronous	500ms	🔊	🔊 (demo)	🔊
	800ms	🔊	🔊	
Phase shifts	500ms			
	800ms	🔊		
Linear change	500 ms -> 800 ms			
	800 ms -> 500 ms		🔊 🔊	

After a moment we'll hear an example of a negative phase shift, where the participant will need to adjust his phase to be a bit earlier, but then resume tapping at the same tempo.

Phase-shifting (ex. 500 ms, single-stimulus)
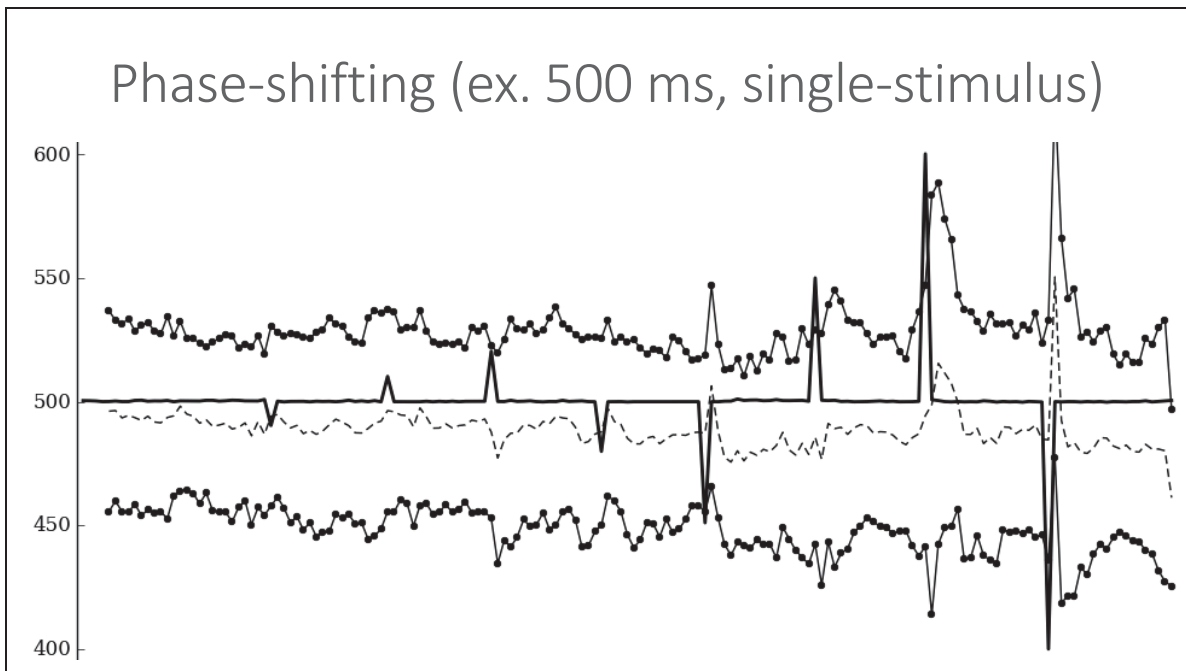
This shows the set of shifts that occurred across the 500-millisecond version of the phase-shifting task.

The first four seemed to have very little effect, but you see an increase in individual differences in variability for the four in the second half. (along with the overall movement that results from the adjustment)

Phase-shifting (ex. 500 ms, single-stimulus)

30: -10 ms,   48: +10 ms,
64: +20 ms,   81: -20 ms,
97: -50 ms,   114: +50 ms,
131: +100 ms, 150: -100 ms

They're not quite at regular intervals, so participants couldn't precisely guess when one would occur.

These settings were constant across all of the phase-shift tasks for all participants, except that larger shifts were used for the 800-millisecond version of the task, such that the proportional shifts were roughly equal.

Linear-change task (single-stimulus)

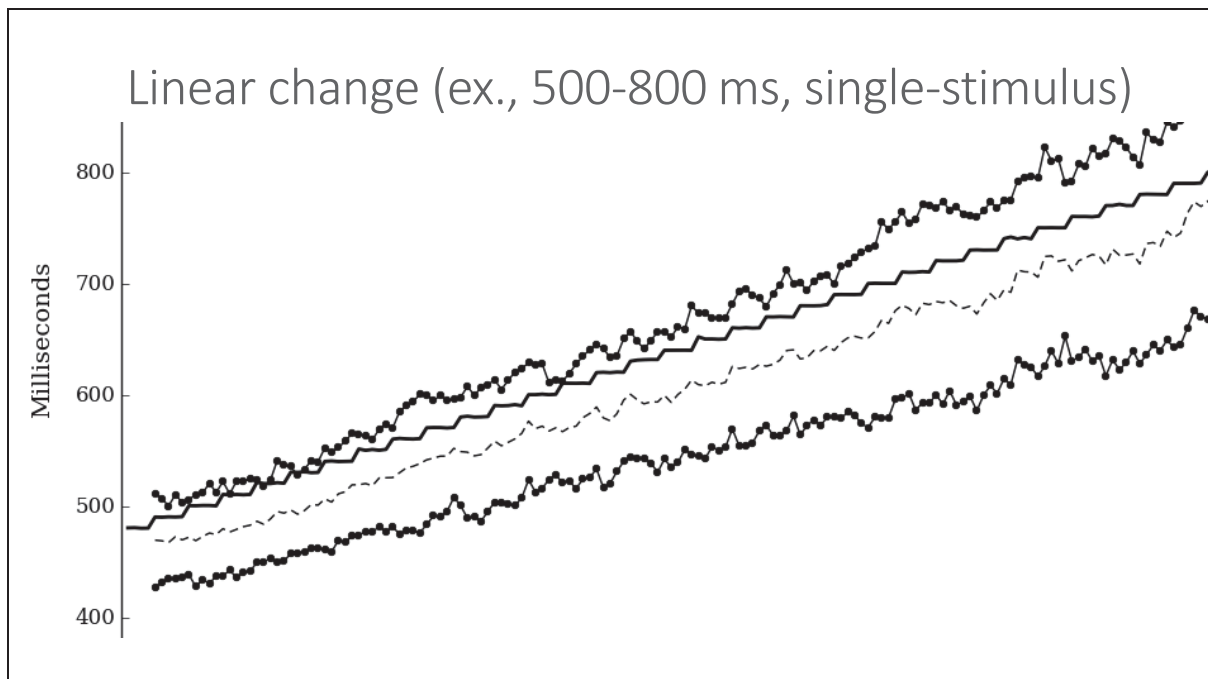| Timing changes	Interonset Interval	1st: Single-stimulus SMS	2nd: Grouped-stimulus SMS	3rd: ISIP
Isochronous	500ms	🔊	🔊 (demo)	🔊
	800ms	🔊	🔊	
Phase shifts	500ms			
	800ms	🔊		
Linear change	500 ms -> 800 ms			
	800 ms -> 500 ms		🔊 🔊	

This would take a long time to demonstrate in an audio sample, because the changes are imperceptibly small at any given point.

Linear change (ex., 500-800 ms, single-stimulus)

For this task, the IOI increased by ten milliseconds every five intervals.

Stimulus target ambiguity: "grouped-stimulus" tasks

Three instruments "all trying to play to the same beat"

Random deviations added to each instrument
(uniform distribution from -10% to +10% of IOI)

Within-subjects factor (2: single-stimulus, grouped-stimulus)

Single-stimulus tasks used a metronome-like ticking sound. These were similar to tasks that have been run in many studies before, including the studies that looked at relationships with intelligence.

For half of the synchronization tasks in this study, participants heard stimuli that sounded like three instruments. The centers of these beats are set to the same timing as each equivalent single-instrument task, but the center of the generating distribution is unheard.

Cowbell, Conga, Muted Triangle

Stimulus target ambiguity: "grouped-stimulus" tasks

Three instruments "all trying to play to the same beat"

Random deviations added to each instrument
(uniform distribution from -10% to +10% of IOI)

Within-subjects factor (2: single-stimulus, grouped-stimulus)

Single-stimulus tasks used a metronome-like ticking sound. These were similar to tasks that have been run in many studies before, including the studies that looked at relationships with intelligence.

For half of the synchronization tasks in this study, participants heard stimuli that sounded like three instruments. The centers of these beats are set to the same timing as each equivalent single-instrument task, but the center of the generating distribution is unheard.

Cowbell, Conga, Muted Triangle

Isochronous task: Grouped-stimulus

Timing changes	Interonset Interval	1st: Single-stimulus SMS	2nd: Grouped-stimulus SMS	3rd: ISIP
Isochronous	500ms	🔊	🔊 (demo)	🔊
	800ms	🔊	🔊	
Phase shifts	500ms			
	800ms	🔊		
Linear change	500 ms -> 800 ms			
	800 ms -> 500 ms		🔊 🔊	

Grouped-stimulus tasks, again, were meant to introduce an element of unpredictable error as seen in human group performance. [Audio]

Linear change task: Grouped-stimulus

Now I'll play an example of how one of the linear change versions of the grouped-stimulus task sounds near the start and end of the task.

Timing changes	Interonset Interval	1st: Single-stimulus SMS	2nd: Grouped-stimulus SMS	3rd: ISIP
Isochronous	500ms	🔊	🔊 (demo)	🔊
Isochronous	800ms	🔊	🔊	
Phase shifts	500ms			
Phase shifts	800ms	🔊		
Linear change	500 ms -> 800 ms			
Linear change	800 ms -> 500 ms		🔊 🔊	

ISIP task example

Finally, I'll demonstrate how the entrainment and self-paced tapping in an isochronous serial interval production task sound, which is relatively simple.

Timing changes	Interonset Interval	1st: Single-stimulus SMS	2nd: Grouped-stimulus SMS	3rd: ISIP
Isochronous	500ms	🔊	🔊 (demo)	🔊
	800ms	🔊	🔊	
Phase shifts	500ms			
	800ms	🔊		
Linear change	500 ms -> 800 ms			
	800 ms -> 500 ms		🔊 🔊	

ISIP task:

Entrainment period (~40 intervals)

Measurement period (continue for about 1m10s)

General procedure

1. Random order: (A) rhythmic task battery, (B) cognitive testing

2. Questionnaire: general demographic questionnaire

3. Questionnaire: musical behavior, practice, skill

4. Questionnaire: personality (BFI-44)

Looking at the overall procedure of the study, the order of cognitive testing versus that whole rhythmic task battery was also randomized for each participant.
After those two sets of tasks, participants completed questionnaires (…)

Rhythmic task administration order

1. Practice tasks (SMS)

2. Single-stimulus SMS tasks
   - Random order: task timing variation (isochronous., phase shifts, linear change; randomized once for both single and grouped stimulus sets)
   - Random order: IOI variations (sequential within each timing variation; randomized once for all tasks)

3. Grouped-stimulus SMS tasks (same as above)

4. ISIP tasks (same IOI order)

5. Improvised creative performance tasks (not analyzed)

And the overall picture of how tasks are presented in the rhythmic battery is:

[These orders were randomized completely independently for each subject, which means that there is unequal representation of the various orders and combinations of orders, but the SPSS GLM procedures I used make appropriate adjustments to avoid non-independence problems.]

Results

Relationships among non-timing variables

Univariate description of SMS outcomes by task type

Reliability across IOI levels

Bivariate relationships with task type combinations

Multiple regression (ISIP, seven predictors)

Two-way within-subjects ANCOVA (SMS factors, seven predictors)

So next we'll look at the results.. We'll go through the relationships among non-timing variables, (…)

	IQ	E	A	C	N	O	IS
WASI-II IQ	1						
BFI-44 E	-.033	1					
BFI-44 A	-.153	.202*	1				
BFI-44 C	-.369**	.197	.228*	1			
BFI-44 N	.085	-.423***	-.306*	-.405***	1		
BFI-44 O	.389***	-.009	-.017	-.074	-.008	1	
Inst. Skill	.353**	-.037	.057	-.246*	.022	.309*	1

*$p < .05$, **$p < .01$, ***$p < .0001$

To start with, we can take a look at the zero-order correlations among the seven main between-subjects predictors.

The main thing to notice here is that the outcome of the intelligence test was correlated with participants' self-ratings of their skill attainment with a musical instrument. It also had a smaller positive correlation with openness, which isn't too surprising, and a smaller *negative* correlation with conscientiousness.

Instrument skill is a categorical scale from zero to four (after the fives, with professional-level skill, were excluded from analyses). So we can break this down into categories…

WASI-II FSIQ-2

Self-rated instrument skill level
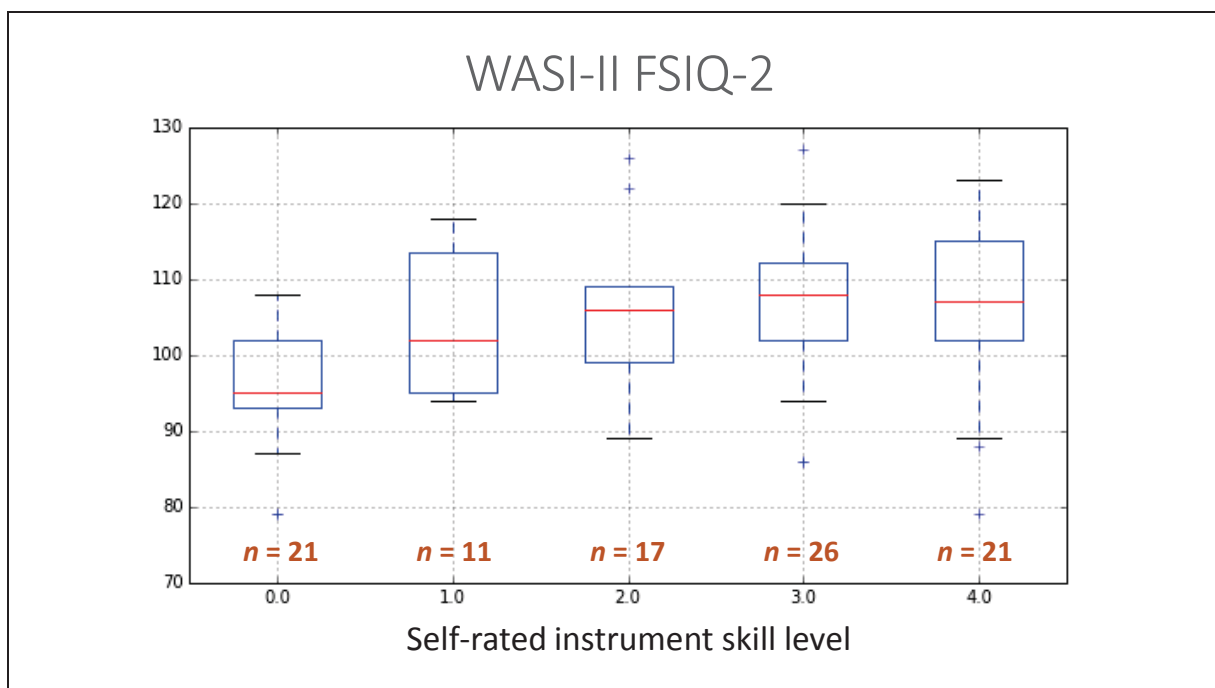
"Zero" corresponded to never spending any time in a serious attempt to learn an instrument. One and two could be considered near-beginner levels; three and four referred to an ability to learn and play complex material. The median intelligence score clearly appears to increase between skill levels one and three. (It might not be quite linear throughout the range, but I've been treating it as linear in these analyses.)

[Notes: the center of each box is the median; the box edges are the interquartile range (1st to 3rd); whiskers extend out 1.5 interquartile ranges; multivariate outliers NOT removed for this visualization.]

SMS task outcomes

We can start looking at the synchronization task outcomes descriptively.

SD asy (% IOI)

	Single-stimulus IOI (ms)					
	500			800		
Timing change	M	SD	N	M	SD	N
Isochronous	4.82	1.65	96	4.42	1.64	93
Phase-shift	6.97	2.60	97	7.37	4.78	97

	Single-stimulus IOI range (ms)					
	500 to 800			800 to 500		
Timing change	M	SD	N	M	SD	N
Linear-change	5.47	1.75	91	5.13	1.56	93

The typical effect of IOI on error was seen here, where error increases linearly with IOI. Since these values are reported as percentage of IOI, the values stay roughly constant between pairs of matching 500 ms and 800 ms tasks, as expected.

For the most part, this held for the altered forms of SMS task involving timing changes and grouped stimuli.

SDasy (% IOI)

	Grouped-stimulus					
	IOI (ms)					
	500			800		
Timing change	M	SD	N	M	SD	N
Isochronous	5.83	1.20	96	6.08	1.48	94
Phase-shift	8.30	3.11	97	8.30	4.46	97

	Grouped-stimulus					
	IOI range (ms)					
	500 to 800			800 to 500		
Timing change	M	SD	N	M	SD	N
Linear-change	6.48	1.60	88	8.09	1.57	94

For the most part, this held for the altered forms of SMS task involving timing changes and grouped stimuli, although the grouped-stimulus, accelerating linear change task may have increased the error rate more than the matching *decelerating* task.

[[This is probably because the distribution of the grouped stimuli was fixed at the start of the task, rather than changing along with the IOI. So as the stimuli accelerated, the scattering of the different instrument sounds became greater relative to IOI.]]

Composite measures of SMS task type combinations

Composite across IOI levels (two-task mean): 6 outcomes based on 12 tasks

Dependent variable (SMS):
$z(\log(\text{mean of \% } SD_{asy} \text{ for two IOI levels}))$

Dependent variable (ISIP):

$\log(\text{mean of \% Local error for two IOI levels})$

Next, for all analyses below, a composite measure was formed from the two tasks within each of the six task variation combinations. This was calculated as the mean of the 500 ms and 800 ms version of each task. A natural log transformation was applied to this due to the characteristic long-tailed distribution of timing data. Finally, because main effects of task variations by themselves were not relevant to the study hypotheses, I used z scores for all of these timing outcomes to simplify interpretation.

That last step (z transformation) wasn't important for interpreting the simpler regression analysis on ISIP outcomes, so those just used the unstandardized log of the two-task mean.

Task outcome correlations between IOI pairs (500 ms / 800 ms)

	Single stimulus	Grouped stimulus	ISIP Local
Isochronous	0.739	0.525	0.605
Phase shift	0.78	0.626	-
Linear change	0.819	0.475	-

We can think of the correlations (here) between pairs of tasks at the two IOI levels as measures of reliability of each of these two-task composites.

So we can see already that the grouped stimulus variations, which aren't as reliable as their single-stimulus counterparts, might not be capturing the additional variability in performer traits that we wanted it to capture. And this was generally the case across the analyses the followed. On the other hand, the measurements of the different timing change tasks seemed to be very stable among single-stimulus tasks.

Correlations: ISIP local vs. SMS tasks

	ISIP Local
Isochronous, Single-stimulus	.723**
Isochronous, Grouped-stimulus	.658**
Phase-shifting, Single-stimulus	.505**
Phase-shifting, Grouped-stimulus	.402**
Linear-change, Single-stimulus	.668**
Linear-change, Grouped-stimulus	.405**

**. Correlation is significant at the 0.01 level (2-tailed).

We can also look a bit more broadly at the two basic types of tasks—the individual differences tapped by SMS and ISIP tasks were strongly related for tasks that were most similar to simple entrainment to a metronome, and weaker for some of the more complex combinations.

IQ & instrument skill / task outcomes

	WASI-II IQ	Instrument Skill
WASI-II IQ	1	
Self-rated Musical Instrument Skill	.353**	1
SMS: Isochronous, single-stimulus	-0.203*	-.436***
SMS: Isochronous, grouped-stimulus	-0.124	-.365**
SMS: Phase-shifting, single-stimulus	-.269**	-.402***
SMS: Phase-shifting, grouped-stimulus	-.204*	-.293**
SMS: Linear-change, single-stimulus	-0.186	-.494***
SMS: Linear-change, grouped-stimulus	-0.051	-.307**
ISIP Local error	-0.152	-.362**

*$p < .05$, **$p < .01$, ***$p < .0001$

Before getting to the full linear models, we can see how some of the individual task outcomes relate to predictor variables.

First, there were weak but significant relationships between IQ and each of the two phase-shifting tasks. One of the two had a P value of .049, though, and these results are uncorrected for multiple comparisons.

Second, there's no indication that the grouped stimulus presentation that was devised for the study succeeded in increasing the cognitive requirements of the task such that outcomes would reflect general intelligence.

IQ & instrument skill / task outcomes

	WASI-II IQ	Instrument Skill
WASI-II IQ	1	
Self-rated Musical Instrument Skill	$.353^{**}$	1
SMS: Isochronous, single-stimulus	-0.203^{*}	$-.436^{***}$
SMS: Isochronous, grouped-stimulus	-0.124	$-.365^{**}$
SMS: Phase-shifting, single-stimulus	$-.269^{**}$	$-.402^{***}$
SMS: Phase-shifting, grouped-stimulus	$-.204^{*}$	$-.293^{**}$
SMS: Linear-change, single-stimulus	-0.186	$-.494^{***}$
SMS: Linear-change, grouped-stimulus	-0.051	$-.307^{**}$
ISIP Local error	-0.152	$-.362^{**}$

$*p < .05, **p < .01, ***p < .0001$

Next, we can see that the self-ratings of musical instrument skill consistently show small to moderate negative relationships with synchronization and interval production error. This seems intuitively reasonable, but it was unexpected given previous research, and I'll say a bit more about this later on.

Given the relationship between IQ and instrument skill, we might expect that these task outcomes might not be *uniquely* predicted by IQ at all. That did turn out to be case for the most part.

ISIP local error composite

	Predictor	β	t	p
Model 1 $(R^2 = .134)$	(Constant)		3.218	.002
	Instrument Skill	-.366	-3.807	.000
Model 2 $(R^2 = .169)$	(Constant)		2.833	.006
	Instrument Skill	-.301	-3.011	.003
	BFI-44 Openness	-.200	-2.002	.048

That's beside the point for the analysis of ISIP Local error, which didn't have a bivariate relationship with IQ here. I used the stepwise regression procedure in SPSS to select from the seven continuous between-subjects predictors, as predictors of the ISIP local error composite. As we can see, instrument skill predicted ISIP Local error, and a weaker effect of openness was found. (Negative betas indicate that greater skill and openness reduced error variability in performance.)

Predictors entered: WASI-II FSIQ-2, BFI-44 (5 factors), Self-rated Instrument Skill Level)

SMS model

Composite of SD_{asy} across IOI levels (two-task mean)

D.V. = $z(\log(\text{composite}))$

Timing changes (random order)	1st: Single-stimulus SMS	2nd: Grouped-stimulus SMS
Isochronous		
Phase shifts		
Linear change		

The SMS task outcomes, again, were based on the standard deviation of asymmetries between targets and taps, for each of the six combination of timing changes and stimulus types. I used the GLM repeated-measures procedure in SPSS to model this.

SMS model

Two within-subjects factors (2 x 3)

3 between-subjects factors: Dummy codes for order effects

7 between-subjects covariates: IQ, Instrument Skill, Personality

Timing changes (random order)	1st: Single-stimulus SMS	2nd: Grouped-stimulus SMS
Isochronous		
Phase shifts		
Linear change		

I used the GLM repeated-measures procedure in SPSS to model this. The predictors included

--two within-subjects factors for task timing variations,

--three between-subjects dummy codes to represent the randomized orders,

--seven continuous between-subjects variables, which were IQ, instrument skill level, and the five personality factors.

SMS model: planned contrasts

Key questions: how do the BS covariates interact with the WS task-type factors as predictors of SD_{asy}?

Timing changes (random order)	1st: Single-stimulus SMS	2nd: Grouped-stimulus SMS
Isochronous		
Phase shifts		
Linear change		

The important outcomes of the analysis are the interactions of between-subjects variables and within-subjects factors. These will tell us how the elements of timing shifts and complex stimuli might determine the extent to which synchronization error predicts individual differences in intelligence, instrument skill level, and the five personality traits.

So we'll need to contrast the levels of timing change variation, if any interactions are present. Here, I used comparisons of each of the two task types that alter the timing (Phase, Linear) with the simpler, isochronous task type.

Between-subjects predictors of $z(\log(SD_{asy})$ across WS factors

Source	Type III SS	df	Mean Square	F	p	Partial eta squared
WASI-II FSIQ-2	0.537	1	0.537	1.735	0.192	.021
Self-Rated Instrument Skill	4.626	1	4.626	14.961	0.0002	.159
BFI-44 Conscientiousness	1.324	1	1.324	4.281	0.042	.051
BFI-44 Openness	2.250	1	2.250	7.278	0.009	.084
(order / personality factors)	…	(6)	…	…	…	…
Error	24.426	79	0.309			

First, we can have a look at the overall between-subjects effects. So here we're looking across both of the within-subjects factors at the average effect of these predictors on SMS error. There was no significant unique effect of IQ—

--But we see that instrument skill level, openness, and conscientiousness predicted SMS error (and we can't see it here, but conscientiousness predicted *greater* error.) The other personality factors and task-order variables also did not significantly affect the mean error across the SMS task set.

Interactions with stimulus type variations (single, grouped)

Source	Type III SS	df	MS	F	**Sig.**	Partial η^2
Stimulus type	.835	1	.835	1.990	**.162**	.025
Stimulus type x FSIQ	.271	1	.271	.645	**.424**	.008
Stimulus type x Instrument Level	2.131	1	2.131	5.079	**.027**	.060
Stimulus type x Agreeableness	1.414	1	1.414	3.370	**.070**	.041
(Interactions with personality, presentation order, n.s.)	(...)	(6)	(...)	(...)	**(...)**	(...)
Error	33.143	79	.420			

Next, we can look at the interaction effects. These all use the Huynh-Feldt* correction for violation of sphericity.

The stimulus type variations interacted with instrument skill level. We don't need a contrast to interpret that, but we could look at regression slopes or correlations for instrument level to see how they differed between the two stimulus types.

*["winn-felt"]

Interpretation of interaction effects

	Instrument Level	WASI-II FSIQ-2	BFI-44 Agreeableness	BFI-44 Neuroticism
MM stimulus: single	-.498**	-.183	-.091	-.101
MM stimulus: grouped	-.389**	-.122	.070	-.170
MM timing: iso	-.456**	-.141	-.022	-.095
MM timing: phase	-.407**	-.225*	-.043	-.168
MM timing: linear	-.451**	-.107	.010	-.137

$* \ p < .05$

$** \ p < .01$

Here we're looking at marginal means for the within-subjects factors. For the two stimulus type marginal means, that's averaging across the three levels of timing change. (We'll come back to this table for another interaction effect as well.)

Interpretation of interaction effects (stimulus x inst. lv.)

	Instrument Level	WASI-II FSIQ-2	BFI-44 Agreeableness	BFI-44 Neuroticism
MM stimulus: single	-.498**	-.183	-.091	-.101
MM stimulus: grouped	-.389**	-.122	.070	-.170
MM timing: iso	-.456**	-.141	-.022	-.095
MM timing: phase	-.407**	-.225*	-.043	-.168
MM timing: linear	-.451**	-.107	.010	-.137

* $p < .05$

** $p < .01$

Source	Type III SS	df	MS	F	Sig.	Partial η^2
Stimulus type x Instrument Level	2.131	1	2.131	5.079	.027	.060

So in the orange region we can see that the interaction (reproduced below) indicates that self-rated instrument skill predicted single-stimulus accuracy more strongly than grouped-stimulus accuracy. (Marginal means for the three levels of timing change– so that's averaging across the two levels of stimulus type.)

Interactions with timing change variations (isochronous, phase shift, linear change)

Source	Type III SS	df	MS	F	Sig.	Partial η^2
Timing change type	.679	2	.339	1.284	.280	.016
Timing change type x Order: phase-shift before isochronous	1.884	2	.942	3.564	.031	.043
Timing change type x FSIQ	1.896	2	.948	3.586	.030	.043
Timing change type x Instrument Level	2.135	2	1.067	4.039	.019	.049
(Interactions with personality, presentation order, n.s.)	(…)	(14)	(…)	(…)	(…)	(…)
Error	41.762	158	.264			

Looking at interaction effects with the three-level factor of timing change type now:

First, we can see an interaction with one of the randomized presentation order dummy codes. I won't go into detail here about the interaction between presentation order and the timing change variations, but essentially what it found was that participants tended to perform better on some of the tasks if they were presented earlier, so participant fatigue seems to have been more important than any learning curve that might pull in the other direction.

Interactions with timing change variations (isochronous, phase shift, linear change)

Source	Type III SS	df	MS	F	Sig.	Partial η^2
Timing change type	.679	2	.339	1.284	.280	.016
Timing change type x Order: phase-shift before isochronous	1.884	2	.942	3.564	.031	.043
Timing change type x FSIQ	1.896	2	.948	3.586	.030	.043
Timing change type x Instrument Level	2.135	2	1.067	4.039	.019	.049
(Interactions with personality, presentation order, n.s.)	(...)	(14)	(...)	(...)	(...)	(...)
Error	41.762	158	.264			

What's more interesting is that both IQ and instrument skill level both interacted with the levels of timing change type. We'll need to look at the contrasts between isochronous tasks and the other tasks to see where this is coming from.

Contrasts: timing change type interactions

		Type III SS	df	F	sig.	Partial η^2
Timing change type x FSIQ	Phase vs. Isochron.	.702	1	2.131	.148	.026
	Linear vs. Isochron.	.278	1	1.075	.303	.013
Timing change type x Instrument Level	Phase vs. Isochron.	1.822	1	5.529	**.021**	.065
	Linear vs. Isochron.	.036	1	.141	.709	.002
Error	Phase vs. Isochron.	26.034	79	.330		
	Linear vs. Isochron.	20.418	79	.258		

The planned contrasts did not capture the interaction effect with IQ. We'll take a look at those relationships anyway in a minute.

For instrument level, there was a significant difference between the phase-shifting and isochronous task types in how strongly their outcomes related to instrument skill level.

Interpretation of interaction effects (instrument level)

	Instrument Level	WASI-II FSIQ-2	BFI-44 Agreeableness	BFI-44 Neuroticism
MM stimulus: single	-.498**	-.183	-.091	-.101
MM stimulus: grouped	-.389**	-.122	.070	-.170
MM timing: iso	-.456**	-.141	-.022	-.095
MM timing: phase	-.407**	-.225*	-.043	-.168
MM timing: linear	-.451**	-.107	.010	-.137

* $p < .05$

** $p < .01$

		Type III SS	df	F	sig.	Partial η^2
Timing change type x Instrument Level	Phase vs. Isochron.	1.822	1	5.529	.021	.065
	Linear vs. Isochron.	.036	1	.141	.709	.002

The isochronous task had a stronger relationship with instrument skill level.

Looking at the IQ relationships, it appears that an interaction contrast that compared phase-shift outcomes to the mean of the other two might have shown an effect, but that won't be terribly meaningful theoretically. (It would be interesting to see how this would have turned out with a larger sample size, given that the difference between isochronous and phase shift tasks is in the predicted direction.)

$z(\log(SD_{asy}))$ marginal mean: Isochronous tasks

Self-rated instrument skill level
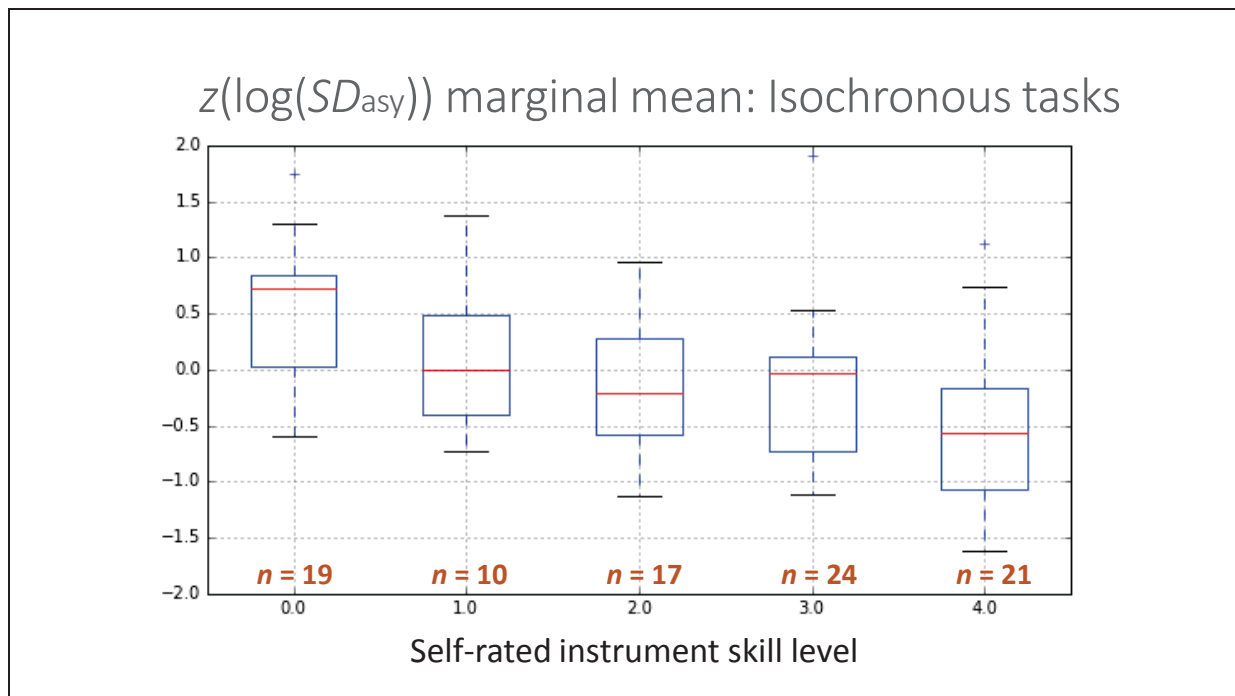
We can look in more detail about what's going on with instrument skill level by looking again at the individual level ratings. Again, the isochronous task outcomes (shown here) related to instrument skill significantly more strongly than the phase-shifting tasks did.

z(log(*SD*asy) marginal mean: Phase-shift tasks
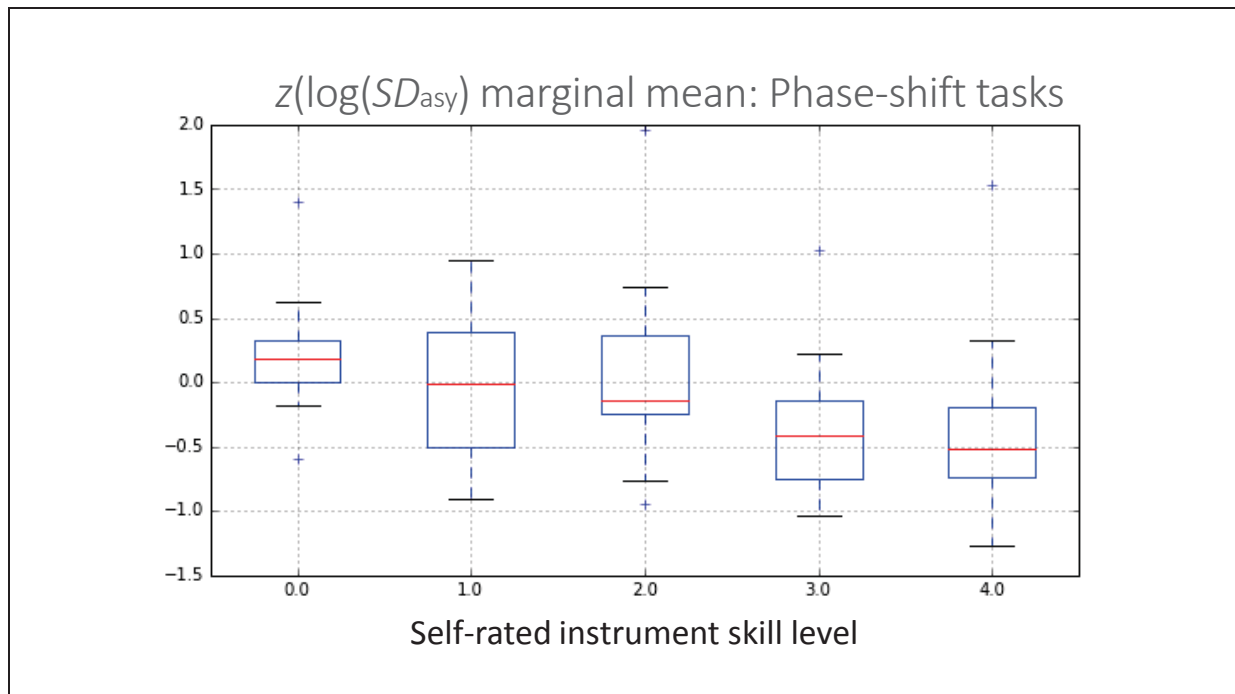
Phase shift tasks and self-rated instrument skill. (This relationship was significantly weaker than Instrument Skill / Isochronous task)

There's very little change in these profiles except at the zero and three skill levels.

z(log(*SD*asy) marginal mean: Linear-change tasks

Self-rated instrument skill level

n = 19   n = 10   n = 17   n = 24   n = 21
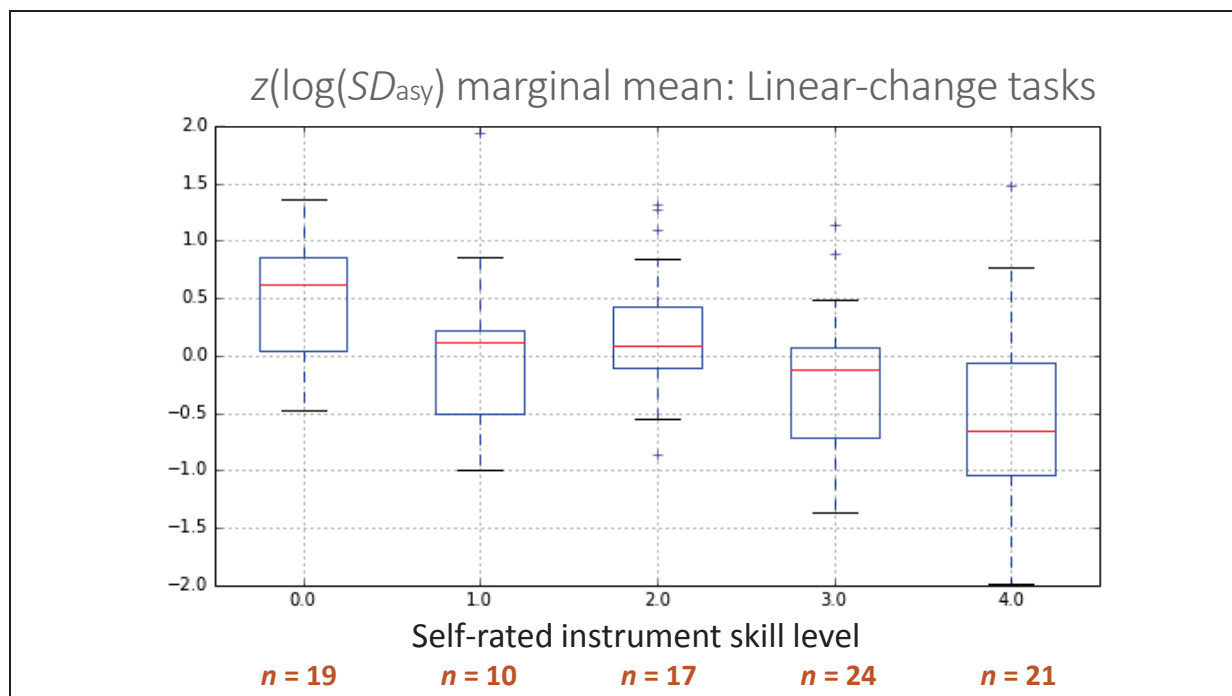
Just as another point of comparison, this is the profile for the linear change tasks, which weren't found to be significantly different from the isochronous tasks in how they relate to instrument skill level.

3-way interaction (WS cells)

Source	Type III SS	df	MS	F	Sig.	Partial η^2
(Stimulus x Timing)	.378	2	.189	1.169	.313	.015
(Stimulus x Timing) x Order: Phase before Isochronous	1.143	2	.572	3.534	.032	.043
(Stimulus x Timing) x Neuroticism	1.618	2	.809	5.001	.008	.060
…	…	(16)	…	…	…	…
Error	25.556	158	.162			

Finally, we see a three-way interaction with neuroticism– we could do a simple effects test here, but it's unclear what kind of interpretable result could be found. Just taking a quick look at correlations with the cell means….

Contrasts, 3-way interaction (WS cells)

		Type III SS	df	F	sig.	Partial η^2
(Stimulus x Timing) x Neuroticism	Phase vs. Isochron.	.119	1	.194	**.661**	.002
	Linear vs. Isochron.	4.040	1	5.912	**.017**	.070
Error	Phase vs. Isochron.	48.624	79	.615		
	Linear vs. Isochron.	53.983	79	.683		

Phase-shifting, grouped-stimuli		
	Pearson Correlation	-.250
	Sig. (2-tailed)	.017
	N	91

(but no relationships between individual isochronous or phase-shifting task cells and neuroticism)

…the only significant relationship there was that participants higher in neuroticism tended to have *lower* error on one of the most complex task types, the grouped-stimulus, phase-shifting combination.
So it's not clear what that might tell us, although it's actually a stronger effect than any of the others we've been looking at.

Summary/conclusions

Synchronization with a phase-shifting sequence is weakly related to general intelligence among nonmusicians

Self-rated skill achievement in amateur musicianship predicted accuracy across all synchronization and self-paced interval tapping tasks examined

Controlling for self-rated skill achievement, no unique variability in synchronization error was associated with general intelligence

Relationship between general intelligence and ISIP Local error not replicated in this sample

So, to summarize what was found here:

(…)

A parsimonious explanation for this is that the causal effect on intelligence is entirely through its effects on musical instrument skill attainment.

Amateur musical skill and SMS in previous research

Repp (2010): No difference in SD_{asy} between amateur musicians and participants with no musical training at all

Hove, et al. (2010): No effects of music training among college students in synchronization accuracy

Repp & Su (2013), reviewing: "Thus, it seems that only a high level of rhythmic experience reduces the variability of tapping in SMS."

This consistent prediction of rhythmic task outcomes from self-rated musical instrument skill was unexpected, given previous research.

Limitations to interpretation

Single-item musical achievement measure; IQ/achievement relationship could be artifactual

Causes of differences in non-specialists' musical skill attainment not explored (e.g., efficiency of practice; interest in creative expression; interest in musical performance specifically)

Western/educated population not representative of traditional inclusiveness of musical performance; greater self-selection of musical practice behavior

Differences in musical performance ability (e.g., rhythm): hypotheses

Differences reflect general intelligence (as well as domain-specific cognitive skills and practice)
◦ (Ruthsatz, Detterman, Griscom, et al. 2008)

Differences due primarily to levels of practice
◦ (Ericsson, Nandagopal, & Roring, 2005)

To put this in context, there's been a vigorous debate about the relative impact of talent and focused practice on specialized skills like musical virtuosity.

The amount of practice required to reach musical expertise may be much lower among high-intelligence performers, as a paper by Ruthsatz found in musicians, so rhythmic precision might still indicate intelligence if the members of a group being compared on skill are assumed to have similar practice opportunity.

So studying participants' history of rhythmic and musical practice more carefully, in addition to the outcome of skill attainment, could make it possible to determine why the overlap between intelligence and musical skill as predictors of synchronization error was found.

Next steps

Follow-up study: gather participant ratings of creative improvised performances from current sample
  ◦ (Intelligence, personality, instrument skill, and rated creativity and accuracy)

Rhythmic endurance: playing near physical exhaustion would be ecologically valid and might result in different kinds of relationships with synchronization accuracy

Reducing fatigue and boredom and increasing musical motivation

Unique intelligence effects on ratings could suggest that the quantitative measures used in the current study missed important elements of rhythmic performance ability.

The fatigue effect demonstrated by interactions with presentation order could be a problem, because participants may not have been indicating their optimal abilities if they were bored of the tasks and reducing effort.

Promoting perception of the process as musical could avoid this problem and increase the ecological validity of the tasks.

Next steps

Adaptive testing of responses to phase shifts (Repp, Keller, Jacoby, 2012)

Compare with choice reaction time tasks; rhythmic change detection tasks with discrete responses

Some useful measures to compare with those in this set would be a method of quantifying the phase correction response more precisely, through adaptive testing, that was used by Bruno Repp.

Finally, it would clearly be useful to proceed by comparing outcomes like those in this study to actual reaction time tasks, since the hypothesis predicts common sources of variability.