Journal of Experimental Psychology:
Human Perception and Performance
2006, Vol. 32, No. 3, 780–786

0096-1523/06/$12.00 DOI: 10.1037/0096-1523.32.3.780

OBSERVATION

Evidence for Tempo-Specific Timing in Music Using a Web-Based Experimental Setup

Henkjan Honing
Universiteit van Amsterdam

Perceptual invariance has been studied and found in several domains of cognition, including those of speech, motor behavior, and object motion. It has also been the topic of several studies in music perception. However, the existing perceptual studies present rather inconclusive evidence with regard to the perceptual invariance of expressive timing under tempo transformation in music performance. The current study used a novel experimental methodology that took advantage of new technologies, such as an online Internet setup, high-quality audio, and state-of-the-art tempo-transformation techniques. The results show that listeners could detect which was the original performance when asked to compare 2 recordings, 1 of which was tempo-transformed to make both similar in overall tempo. This result is taken as support for the tempo-specific timing hypothesis—which predicts that a tempo-transformed performance will sound less natural than an original performance—and as counterevidence for the relationally invariant timing hypothesis, which predicts that a tempo-transformed performance will sound equally natural.

Keywords: music cognition, Internet-based experiment, expressive timing, tempo, rate

Invariance and variability have been important topics in the cognitive sciences for several decades now. Perceptual invariance is concerned with whether certain objects or event properties remain perceptually constant under transformation (Shepard & Levitin, 2002). Perceptual invariance has been studied and found in several domains of cognition, including speech (Perkell & Klatt, 1986), motor behavior (Heuer, 1991), and object motion (Shepard, 2001). It has also been the topic of several studies in music perception (Handel, 1992; Hulse, Takeuchi, & Braaten, 1992; Repp, 1995). A well-known and relatively uncontroversial example is melody (Dowling & Harwood, 1986).[1] When a melody is transposed to a different register, it not only maintains its frequency ratios in performance, but it is also perceived as the same melody (i.e., melody remains perceptually invariant under transposition). With respect to other aspects of music, such as rhythm, there is less agreement in the literature. Whereas one might expect rhythm to scale proportionally with tempo in production and to be perceptually invariant under tempo transformation, several studies have shown that this is not always the case (Handel, 1992; Monahan & Hirsh, 1990). Rhythms are timed differently at different tempi (Repp, Windsor, & Desain, 2002), and listeners often do not recognize proportionally scaled rhythms as being identical (Desain & Honing, 2003; Handel, 1993).

Another aspect of music whose perceptual invariance under tempo transformation has been studied is expressive timing, that is, the minute deviations from regularity that contribute to the quality of a musical performance (Clarke, 1999; Palmer, 1997). The existing perceptual studies present rather inconclusive evidence. Repp (1994) asked listeners to distinguish tempo-transformed from original performances that were recorded and played back on an electronic MIDI keyboard instrument that used a ±15% tempo range. The responses turned out to be barely above chance level. However, Repp (1995) found a small but significant effect of tempo in a subjective rating task with the same material (using a +44/−23% tempo range). Reed (2003) found no effects of tempo in an identification task (using a ±20% tempo range) but some effects in a rating and ranking task. Recently, Honing (in press-b) found a significant effect of tempo in an identification task using stimuli from a variety of musical genres (using a ±20% tempo range). However, the last two studies did not control for the effect of tempo preference or for the effect of artifacts caused by the tempo-transformation method, and this could have biased the results.

Portions of this work were presented at the 10th Rhythm Perception and Production Workshop, University of Ghent, Ghent, Belgium (July 2–6, 2005), and at the International Computer Music Conference, Universitat Pompeu Fabra, Barcelona, Spain (September 5–9, 2005). Special thanks to Jordi Bonada (Music Technology Group/Universitat Pompeu Fabra) for providing his time-scale modification algorithm and for preparing the tempo-transformed examples, to Bruno Repp (Haskins Laboratories, New Haven, Connecticut) for various suggestions related to the experimental design, to Marijke Engels (Universiteit van Amsterdam, Amsterdam, the Netherlands) for statistical advice, and to all beta-testers at the Universiteit van Amsterdam and Northwestern University for their time and their suggestions as to how to improve the Internet version of the experiment.

Correspondence concerning this article should be addressed to Henkjan Honing, Music Cognition Group, ILLC/Universiteit van Amsterdam, Nieuwe Doelenstraat 16, 1012 CP Amsterdam, the Netherlands. E-mail: honing@uva.nl

[1] Note that when timbre is considered as well, the invariance of melodic patterns under transposition is less clear, as is suggested by, for example, the tritone paradox (Deutsch, 1987).

These inconclusive results might have been caused by several factors. One factor could be the particular structural properties of the musical material that was used in the experiments. Honing (2005) argued that the rhythmic structure might have an effect on the use of tempo fluctuations (*tempo rubato*), with rhythmically varied music being less susceptible to relational invariance because of differentiated durations, present in the composition used in Desain and Honing's (1994) study (i.e., theme and first variation on "Nel cor più non mi sento," WoO 70, by L. van Beethoven) but not in Repp's (1994, 1995) studies (i.e., "Träumerei," Opus 15, No. 7, by R. Schumann).

Another factor that could have influenced the results was the kind of stimuli used. Repp (1995) presented MIDI performances at different tempi played back on an electronic keyboard. These performances included several "regularizations" applied to, for example, onset asynchronies and articulation. These regularizations could well have interfered with the perceived quality of the performances and, arguably, might have made it more difficult to make judgments on the naturalness of the performances. In that sense, audio recordings, as used in Reed's (2003) and Honing's (in press-b) studies, can be considered more ecologically valid stimuli. However, as said, these studies did not control for the effect of possible artifacts of the tempo-transformation method used.

Hence, the study reported here applied a different experimental design than that used in the perceptual studies mentioned above. To minimize the influence of tempo preference, I used two different recordings of the same composition. These were made similar in tempo by tempo-transforming each of them to the overall tempo of the other, which resulted in two pairs that were presented to different groups of listeners. These groups were asked (through the use of a comparison task) to indicate which of the two performances with the same overall tempo was an original recording (Experiment 1) by focusing on the expressive timing of the performances. Furthermore, to control for the effect of artifacts, in a second experiment, I asked a control group to focus on the sound quality of the recordings and to indicate whether they contained an artifact that could be attributed to the signal processing method (Experiment 2).[2]

Two hypotheses are considered, namely the relationally invariant timing hypothesis and the tempo-specific timing hypothesis. In the experimental design used, the first hypothesis is in fact the null hypothesis. It predicts no significant difference in responses between the original and the tempo-transformed excerpts: Because both excerpts will sound equally natural, the respondents will consider both versions musically plausible performances and, consequently, simply guess which one is the original recording.

However, the tempo-specific timing hypothesis is supported if a significant proportion of the respondents are able to identify the original. This hypothesis is based on the idea that expressive timing in music performance (defined as the local deviations from isochrony, as well as more global changes in tempo) is intrinsically related to global tempo. When expressive timing is simply scaled to another tempo (i.e., slowed down or sped up proportionally), the performance might sound awkward or unnatural and, hence, easier to identify as a tempo-transformed version. Furthermore, one could argue that if performers adapt their timing to the global tempo in a nonproportional way (as has been shown for at least some music repertoires; Desain & Honing, 1994; Friberg & Sundström, 2002), it might well be that listeners are sensitive to this as well: A performance that has been tempo-transformed might sound awkward because the expressive timing is not adapted in the way a musician would normally do it.

There are several advantages to this experimental design. First, expert performances, which are widely available as commercial recordings, can be used in an experimental setting. This has advantages over researchers using performances recorded in a laboratory setting with MIDI instruments, which might well influence the musician's performance.

Second, the task is relatively simple, yet challenging: Listeners are asked to compare two different interpretations of the same composition in the same tempo. This is a task similar to that which a panel of experts might be asked to do in a piano competition or master class; listeners interested in music performance tend to find this attractive.[3]

Third, such a comparison task resolves the issue of tempo preference (which was problematic in some of the earlier studies) because the performances that are compared are presented at roughly the same tempo.

Fourth, because the comparison task is generally considered attractive and challenging, it is relatively easy to recruit large numbers of participants. Large numbers are necessary not only to be able to analyze categorical frequency data (contrary to the earlier studies, which relied on analyses of variance and just 10 respondents) but also to allow for testing hypotheses on the perception of timing and tempo among potentially very different types of participants (e.g., jazz vs. classical musicians, amateurs vs. experts, Western vs. non-Western listeners) simply because of the widespread availability of Internet and high-quality audio playback facilities.[4]

Experiment 1

In the first experiment, the participants were asked to listen to seven pairs of audio fragments from commercially available recordings by well-known pianists (see Table 1). These were presented in an Internet-listening experiment. One stimulus of the pair was an original recording, the other a manipulated, tempo-transformed recording. The latter was originally performed at a different tempo but had been time stretched (or time compressed) to become close in tempo to the other performance of the pair. The task was to judge which of the two performances was an original recording while focusing on the use of expressive timing.[5]

---

[2] Because piano music was used in this study, a design that uses, for example, a MIDI grand piano (i.e., modern pianola) is an attractive alternative. This would avoid the problem of artifacts in the manipulation of audio data. However, the current setup was preferred (over MIDI synthesized or MIDI grand piano rendering) to take advantage of the wide variety of audio recordings currently available.

[3] This is based on some of the comments given in the questionnaire at the end of the experiment.

[4] Related experiments (Honing, 2004; in press-b) have attracted hundreds of interested respondents.

[5] No qualitative information was collected. However, in a related pilot study that used an identification task and a different set of stimuli (Honing, 2004), participants were asked for an explicit motivation for their judgments. Some examples of these motivations are as follows: "X had a more natural feeling," "X has more energy and vibrancy," "X seems too slow and sluggish," "X sounds too fast, uncomfortable pacing for the music," "X just sounds better," "X invites dancing," or "X sounds like it is tripping over itself." These qualitative responses confirm that the participants were able to focus on the musical quality of the sound examples used.

Table 1
Recordings Used in Experiments 1 and 2

Code	Pianist	Composition	Record label	Recording date	Tempo (beats per minute)	Duration
01 Ga	Glenn Gould	J. S. Bach, Goldberg Variations (1981), BWV 988, Variation 1	Sony, SMK 64126, 1999	1981	83	01’ 11
02 Gb	Glenn Gould	J. S. Bach, Goldberg Variations (1955), BWV 988, Variation 1	Sony, SK 52594, 1992	1955	136	00’ 45
03 GG	Glenn Gould	J. S. Bach, English Suite No. 4, BWV 809, Allemande	Sony, SK 87766, 2001	1974–1976	87	00’ 33
04 SR	Sviatoslav Richter	J. S. Bach, English Suite No. 4, BWV 809, Allemande	Delos, GH 5601, 2004	1991	70	01’ 27
05 GG	Glenn Gould	J. S. Bach, WTC II, BWV 890, Prelude 21	Sony, SX4K 60150, 1997	1971	152	00’ 50
06 RT	Rosalyn Tureck	J. S. Bach, WTC II, BWV 890, Prelude 21	BBC, BBCL 4116-2, 2002	1976	93	01’ 23
07 AR	Arthur Rubinstein	L. van. Beethoven, Piano Sonata No. 14, Op. 17, No. 2. Allegretto	RCA, 09026-63056-2, 1999	1976	56	01’ 03
08 VA	Vladimir Ashkenazy	L. van. Beethoven, Piano Sonata No. 14, Op. 17, No. 2. Allegretto	Decca, 452 982-2, 1997	<1997	75	00’ 51
09 CA	Claudio Arrau	F. Chopin, Grande Valse Brillante, Op. 18	Philips, 468 391-2, 2001	1979	70	01’ 01
10 VA	Vladimir Ashkenazy	F. Chopin, Grande Valse Brillante, Op. 18	Decca, 417 798-2, 1990	1983–1985	88	00’ 52
11 VH	Vladimir Horowitz	R. Schumann, Kinderszenen, Träumerei	DGG, 474 370-2, 1991	1985–1989	87	00’ 39
12 CA	Claudio Arrau	R. Schumann, Kinderszenen, Träumerei	Philips, 468 391-2, 2001	1974	70	01’ 01
13 GG	Glenn Gould	J. S. Bach, WTC II, BWV 880, Fugue 11	Sony, SX4K 60150, 1997	1969	135	00’ 50
14 RT	Rosalyn Tureck	J. S. Bach, WTC II, BWV 880, Fugue 11	BBC, BBCL 4116-2, 2002	1976	102	01’ 05

Note. BWV = Bach-Werke-Verzeichnis; WTC = Das Wohl Temperierte Clavier; BBC = British Broadcasting Corporation; RCA = Radio Corporation of America; Op. = opus; DGG = Deutsche Grammophon Gesellschaft. SMK, SK, GH, SX4K, and BBCL are parts of record label identification.

Method

Participants. The participants ($N = 162$) responded to an invitation that was sent to a variety of professional mailing lists and to music students at the University of Amsterdam and Northwestern University.[6] Three gift certificates were raffled among those who responded. Of all participants, 46% reported to be an “expert” (musician), 40% reported to be “experienced” (listen frequently to music), and 14% reported to be “average” (listen casually to music).

Equipment. I collected the responses in an online Internet version of the experiment using standard Web browser technologies (i.e., HTML, CGI, and Java scripts). The stimuli were excerpts of commercially available recordings (see Table 1). These excerpts were converted to the MPEG4 file format to guarantee optimal sound quality on different computer platforms and at different data transmission rates.[7] I generated the experimental setup and stimuli using POCO (a computer program; Honing, 1990).

Materials and stimulus presentation. The experiment used 14 original and 14 tempo-transformed recordings. The two stimulus pairs derived from each performance pair (A/B) were presented to two different groups of listeners. Group 1 ($n = 81$) was presented with seven A/B′ pairs (prime indicating a transformed recording), whereas Group 2 ($n = 81$) was presented with seven A′/B pairs. This was done to prevent the respondents from remembering characteristics of the stimuli in one pair and using them to make a response to the other pair.

I made the tempo-transformed versions using state-of-the-art time–scale modification software (Bonada, 2000).[8] For each recording, the tempo of the first four bars was measured with a metronome and checked perceptually by synchronizing it with the music. The resulting tempo estimate was used to calculate the tempo-scaling factor to make the stimulus pairs similar in tempo (see Table 1 for the tempo estimates in beats per minute). All sound excerpts were taken from the beginning of a recording (see Table 1 for the exact durations). The presentation of the stimuli was randomized within and between pairs for each participant, as was assignment of participants to either Group 1 or Group 2.

Procedure. Participants were asked to visit a nonpublic Web page of the online experiment. First, they were asked to test their computer and audio system with a short sound excerpt and to adjust the volume to a comfortable level. Next, they were referred to a Web page containing the actual experiment (see Figure 1). Here, the following instructions were given:

> You will be presented with seven pairs of audio fragments: one being an original recording (by one pianist), the other a manipulated, tempo-transformed recording (by another pianist). The tempo-transformed recording originally had a different tempo, but it has been time-stretched (or time-compressed) to become close in tempo to the other performance of the pair. Your task is to decide which is which. 1) Listen to a pair of audio fragments once and in their entirety. 2) Focus on the use of expressive timing by the performer, such as note asynchrony, tempo rubato and articulation. (Please ignore any phenomena related to audio-recording quality, like noise, ticks, and/or miking technique. The sound quality of the recordings is *not* relevant here.) 3) Then answer the two questions listed next to the excerpts, namely: Which is the real, original recording (i.e., the most natural performance), the top or the bottom excerpt? And are you sure? 4) Please do this for all seven pairs of audio fragments.

At the end of the experiment, the respondents were asked to fill in a short multiple-choice questionnaire to obtain information on, for example, their musical experience. The experiment took, on average, 16 min to complete.

Analysis. The response forms were automatically sent by e-mail to the author and converted into a tabulated file for further analysis with POCO (Honing, 1990). The responses to the “Which is the original?” question

[6] The participants were in fact invited during a previous study (Honing, in press-b) when they indicated, after doing the experiment, that they would like to be invited to take part in related experiments. Most of them participated in this study as well. This, in principle, allows for comparisons between experiments, as is planned for a future study.

[7] See http://www.apple.com/mpeg4/ for technical details.

[8] See http://www.hum.uva.nl/mmm/exp2/ for the stimuli used.

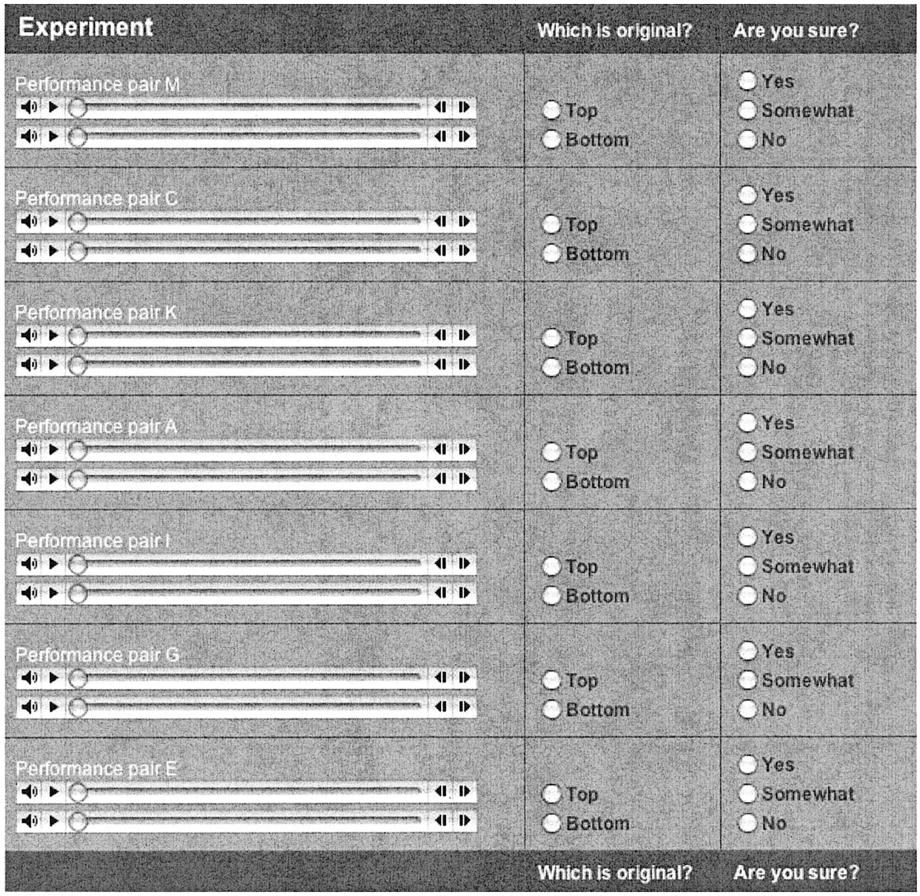

Figure 1. A fragment of the Internet user interface showing the presentation of seven pairs of sound excerpts.

were converted to percentage correct for each stimulus pair, whereas the responses to the "Are you sure?" question were converted to numerical confidence ratings (1 = *Yes*, 0.5 = *Somewhat*, and 0 = *No*). JMP (Version 5.0; SAS Institute, 2003) was used for the statistical analyses.

Results

The results of the comparison task are shown in Figure 2. The participants correctly identified the original recording 70.1% of the time ($SD = 14.6\%$). For 12 of the 14 stimulus pairs, the percentage of correct responses was significantly above chance (see Figure 2). The mean confidence ratings for individual stimulus pairs correlated positively with percentage of correct responses ($r = .37$).

To test the overall significance of the results, I computed the mean percentage of correct responses for each quadruplet of stimuli (i.e., two stimulus pairs of the same composition). I tested the resulting seven values (one for each composition) against chance (i.e., 50% correct) using a *t* test. The difference was significant, $t(6) = 4.91$, $p < .01$. Furthermore, individual participants' percentages of correct responses were found to be significantly above chance level, $t(161) = 17.43$, $p < .01$. In summary, these tests confirmed that the results are indeed significantly different than the null hypothesis.

Experiment 2

To make sure that possible artifacts of the signal processing method (Bonada, 2000) did not bias the responses, I performed a control experiment using the same stimuli as in Experiment 1. These were judged individually for artifacts by a control group that consisted mainly of audio experts.

Method

Participants. The participants ($n = 43$) responded to an invitation that was sent to the *auditory* mailing list. Three gift certificates were raffled among those who responded. Of all participants, 56% reported to be an "audio expert," 26% reported to be "experienced" (listen frequently to music), and 18% reported to be "average" (listen casually to music).

Equipment. Same as for Experiment 1.

Materials and stimulus presentation. The same 28 stimuli used in Experiment 1 were presented individually and in random order to each participant.

Procedure. Participants were asked to visit a nonpublic Web page of the online experiment. First, they were asked to test their computer and audio system with a short sound excerpt and to adjust the volume to a comfortable level. Next, they were referred to a Web page containing the actual experiment. Here, the following instructions were given:

> This listening experiment investigates whether an advanced time-stretching method used in a related experiment causes any audible artifacts. Please do the following: 1) Listen to each excerpt once and in its entirety, using headphones. 2) Focus on possible timbral artifacts (unnatural transients, phasiness, loss of attack sharpness, etc.) in the audio recording. (Please ignore any performance related phenomena. The musical quality, e.g., the timing or tempo used, is *not* relevant here - this is the topic of a parallel study.) 3) Then answer the two

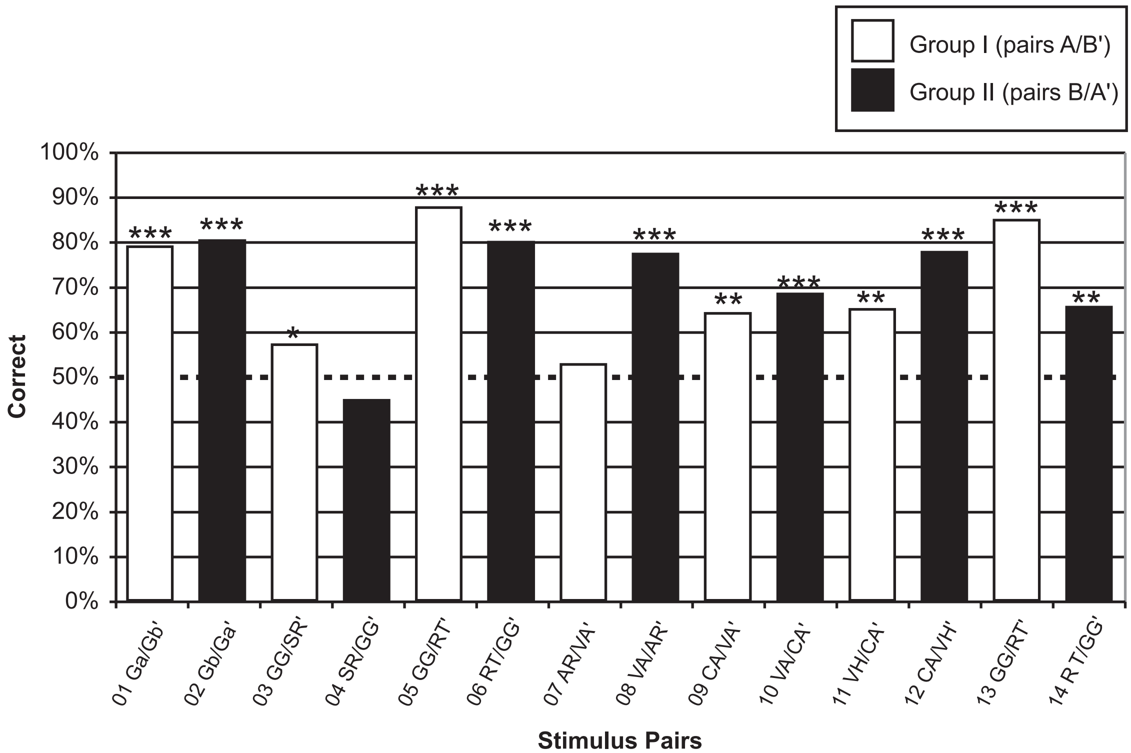

Figure 2. Results of Experiment 1 (Group 1, $N = 81$; Group 2, $N = 81$) showing the percentage of correct responses. Significance levels are indicated with asterisks (binomial test; $^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$).

> questions listed next to the excerpts, namely: Has the recording been manipulated in some way (or is it an original recording), and, are you sure? 4) Please do this for all 28 sound excerpts.

Furthermore, for each sound excerpt, the recording date was mentioned (see Table 1). At the end of the experiment, the participants were asked to fill in a short multiple-choice questionnaire to obtain information on, for example, their listening experience. The experiment took, on average, 32 min to complete.

Analysis. Same as for Experiment 1.

Results

The results of the identification task are shown in Figure 3. The participants correctly identified the original recording 51.2% of the time ($SD = 16.9\%$). To test the significance of these results, for each quadruplet of stimuli, I computed the mean percentage of correct responses and tested against chance (50% correct) using a t test. The difference was nonsignificant, $t(6) = 0.51$, $p < .32$. Furthermore, individual participants' percentages of correct responses were found to be nonsignificant as well, $t(42) = 0.45$, $p < .66$. In summary, these tests confirmed that the participants did not do better than chance over the whole set of 28 stimuli.

There were, however, some individual exceptions. Stimuli 02 Ga′, 05 RT′, and 13 RT′ (all three are tempo-transformed excerpts) attracted a significantly higher number of correct responses than would be expected by chance. Apparently, these did contain artifacts; consequently, the responses to the pairs containing these stimuli in Experiment 1 (i.e., 02 Gb/Ga′, 05 GG/RT′, and 13 GG/RT′) could have been biased. The confidence ratings were positively correlated ($r = .42$) with percentage of correct responses, indicating that when an artifact was spotted, the respondents tended to be confident.

To make sure that these exceptions were not responsible for the overall results, two additional tests were performed on the scores (percentage correct) of Experiment 1 and Experiment 2. This was to ensure that the results of Experiment 1 were indeed significantly different than those of Experiment 2. The first test was a paired-sample t test across compositions, and the second was a t test with unequal sample sizes across participants. Both turned out to be significant, $t(6) = 2.87$, $p < .05$, and $t(161) = 12.96$, $p < .01$, respectively. Thus, in conclusion, we can be certain that the results of Experiment 1 are indeed different from those of Experiment 2.

Discussion

The two experiments reported here were concerned with the question of whether listeners can distinguish an original audio recording by one pianist from a tempo-transformed recording of the same composition by another pianist. Experiment 1 used a comparison task in which listeners were instructed to focus on the expressive timing of the performance (while ignoring the sound quality) and to indicate which was the original and which was the tempo-transformed recording. Experiment 2 presented stimuli singly, and listeners were instructed to focus on the sound quality (while ignoring the musical aspects) and to indicate whether they heard an artifact that could be attributed to the signal processing method.

The results of Experiment 1 are highly significant. Apparently, listeners can often recognize what is an original recording by focusing on the expressive timing of a performance. For Experi-

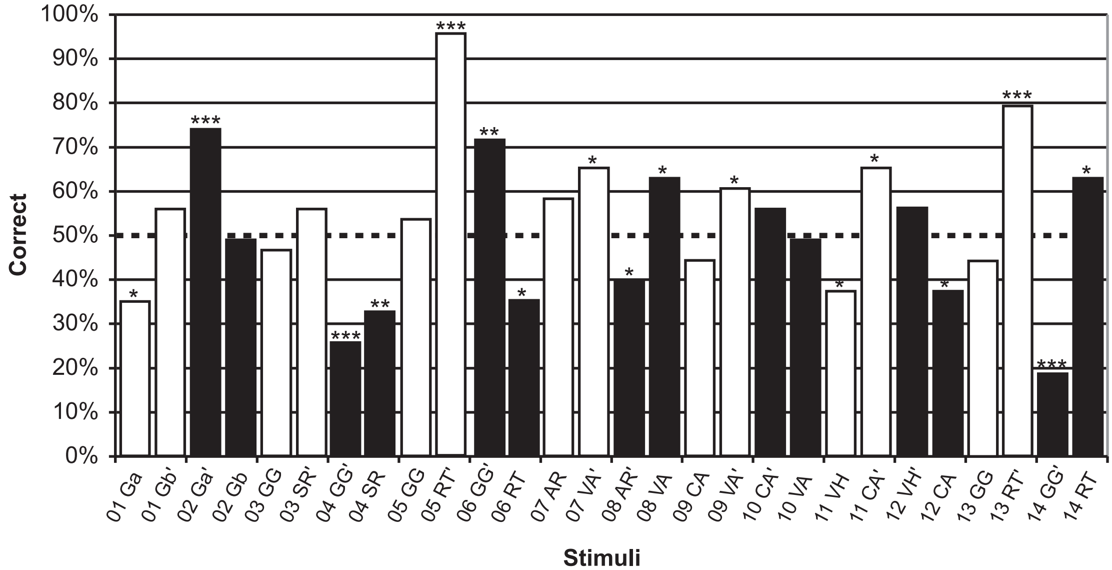

Figure 3. Results of Experiment 2 ($N = 43$) showing the percentage of correct responses. Significance levels are indicated with asterisks (binomial test; *$p < .05$; **$p < .01$; ***$p < .001$). (The black/white coding of the bars is similar to that used in Figure 2, to allow for comparisons.)

ment 2, the responses to only a few stimuli differ significantly from chance, whereas the overall results are not significantly different from chance. Therefore, artifacts of the tempo transformation method did not seriously bias the responses, and the results of Experiment 1 can be taken as support for the tempo-specific timing hypothesis, which suggests that the relationship between timing variations and global tempo can function as a cue for the identification of a real performance. The results are counterevidence for the relationally invariant timing hypothesis, which predicts no preference for the original over the tempo-transformed version: Both versions are predicted to sound equally natural.

Nevertheless, the music performance literature provides some support for the relationally invariant timing hypothesis. Next to a possible effect of the musical material used (as suggested by Desain & Honing, 1994; Repp, 1994), the different results might also be explained by methodological differences (e.g., with fragments in MIDI vs. audio format, with rating vs. identification or comparison tasks).

The current study did not control for a possible effect of familiarity. If listeners were familiar with a particular recording and thought that they recognized the performer, they could have based their judgment (against instructions) on tempo instead of on the expressive timing used. The two famous recordings of the *Goldberg Variations* by Glenn Gould (stimulus pairs 01 and 02) could well be susceptible to such an effect. It has been shown that expert pianists can, at least to some extent, recognize their own performances (Repp & Knoblich, 2004). However, the extent to which listeners are capable of remembering and/or recognizing the timing details of performances by others is less clear. With respect to memory for tempo (Levitin & Cook, 1996), it was shown that the phenomenon of absolute tempo is apparent in pop and rock music but less clear in music from the classical repertoire. The difference between these genres might be attributable to the generally larger variety of tempi used for one composition in classical music compared with pop or rock music, the first having less of an effect on an iconic memory for tempo. Hence, one could argue that the current experimental design is less susceptible to such an effect.

The results presented in this article are important for models of rhythm perception and production in music. Had relational invariance been observed, this would have been an indicator of the existence of a generalized motor program with a variable rate parameter (Heuer, 1991). Several models of expressive timing in music performance indeed suggest this (cf. Honing, 2005); they predict timing to be relationally invariant with global tempo (or rate). However, timing and tempo in music perception and music production are clearly far more intimately coupled: One cannot be changed without affecting the other (Honing, 2001, in press-a).

References

Bonada, J. (2000). Automatic technique in frequency domain for near-lossless time-scale modification of audio. In *Proceedings of International Computer Music Conference* (pp. 396–399). San Francisco: Computer Music Association.

Clarke, E. F. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), *Psychology of music* (2nd ed., pp. 473–500). New York: Academic Press.

Desain, P., & Honing, H. (1994). Does expressive timing in music performance scale proportionally with tempo? *Psychological Research, 56,* 285–292.

Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception, 32,* 341–365.

Deutsch, D. (1987). The tritone paradox: Effects of spectral variables. *Perception & Psychophysics, 41,* 563–575.

Dowling, W. J., & Harwood, D. (1986). *Music cognition.* New York: Academic Press.

Friberg, A., & Sundström, A. (2002). Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern. *Music Perception, 19,* 333–349.

Handel, S. (1992). The differentiation of rhythmic structure. *Perception & Psychophysics, 52,* 497–507.

Handel, S. (1993). The effect of tempo and tone duration on rhythmic discrimination. *Perception & Psychophysics, 54,* 370–382.

Heuer, H. (1991). Invariant relative timing in motor-program theory. In J. Fagard & P. H. Wolff (Eds.), *The development of timing control and temporal organisation in coordinated action* (pp. 37–68). Amsterdam: Elsevier.

Honing, H. (1990). POCO: An environment for analysing, modifying, and generating expression in music. In *Proceedings of the 1990 International Computer Music Conference* (pp. 364–368). San Francisco: Computer Music Association.

Honing, H. (2001). From time to time: The representation of timing and tempo. *Computer Music Journal, 35*(3), 50–61.

Honing, H. (2004). *Is timing tempo-specific? An online Internet experiment on perceptual invariance of timing in music.* Retrieved from the Institute for Logic, Language, and Computation Web site: http://www.illc.uva.nl/Publications/ResearchReports/PP-2004–34.text.pdf.

Honing, H. (2005). Is there a perception-based alternative to kinematic models of tempo rubato? *Music Perception, 23,* 79–85.

Honing, H. (in press-a). Computational modeling of music cognition: A case study on model selection. *Music Perception.*

Honing, H. (in press-b). Is expressive timing relationally invariant under tempo transformation? *Psychology of Music.*

Hulse, S., Takeuchi, A. H., & Braaten, R. F. (1992). Perceptual invariances in the comparative psychology of music. *Music Perception, 10,* 151–184.

Levitin, D., & Cook, P. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics, 56,* 414–423.

Monahan, C. B., & Hirsh, I. J. (1990). Studies in auditory timing: II. Rhythm patterns. *Perception & Psychophysics, 47,* 227–242.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48,* 115–138.

Perkell, J. S., & Klatt, D. H. (1986). *Invariance and variability in speech processes.* Hillsdale, NJ: Erlbaum.

Reed, R. (2003). Tempo change and interpretation preference. In *Proceedings of the European Society for the Cognitive Sciences of Music* (pp. 558–561). Hannover, Germany: University of Hannover.

Repp, B. H. (1994). Relational invariance of expressive microstructure across global tempo changes in music performance: An exploratory study. *Psychological Research, 56,* 269–284.

Repp, B. H. (1995). Quantitative effects of global tempo on expressive timing in music performance: Some perceptual evidence. *Music Perception, 13,* 39–57.

Repp, B. H., & Knoblich, G. (2004). Perceiving action identity: How pianists recognize their own performances. *Psychological Science, 15,* 604–609.

Repp, B. H., Windsor, W. L., & Desain, P. (2002). Effects of tempo on the timing of simple musical rhythms. *Music Perception, 19,* 565–593.

SAS Institute. (2003). JMP (Version 5.0) [Computer software]. Cary, NC: Author.

Shepard, R. (2001). Perceptual–cognitive universals as reflections of the world. *Behavioral and Brain Sciences, 24,* 581–601.

Shepard, R., & Levitin, D. (2002). Cognitive psychology and music. In D. Levitin (Ed.), *Foundations of cognitive psychology: Core readings* (pp. 503–514). Cambridge, MA: MIT Press.

Received July 26, 2005
Revision received November 14, 2005
Accepted November 15, 2005 ■