KEEPING AN EYE ON GUITAR SKILL: VISUAL REPRESENTATIONS OF GUITAR CHORDS

MATTHEW J. C. CRUMP
Brooklyn College of CUNY

GORDON D. LOGAN, & JERRY KIMBROUGH
Vanderbilt University

WE INVESTIGATE A ROLE FOR VISION IN SKILLED GUITAR playing, focusing on visual contributions to the representation of basic first-position root chords (C, A, G, E, D). Experiment 1 involved naming or playing guitar chords displayed in different visual formats (letter, photograph, chart) and orientations. Experiment 2 employed a Stroop-like design, involving identification of the visual or auditory dimension for congruent or incongruent pairs of chord photographs and sounds. Our results demonstrate that visual representations of guitar chords are orientation sensitive and associated with their corresponding actions and sounds. We discuss the implications of our findings for understanding the multimodal nature of musical skill, and consider how the format of visual information can impact acquisition of musical skill.

Received August 11, 2010, accepted October 4, 2011.

Key words: guitar, Stroop, orientation, vision, association

MUSICAL PERFORMANCE IS A BROAD CLASS OF SKILLS involving sight-reading, playing from memory, playing by ear, and improvisation (McPherson, 1994). All of these skills rely on several component processes mediating perceptual, cognitive, and motor aspects of musical structure. The present experiments are aimed specifically at the role of vision in representations of musical structure, a topic that has not received extensive treatment in the literature. Prior research has focused primarily on skill sets common among formally trained musicians, particularly pianists (Palmer, 1997). For example, the role of vision has been investigated most extensively in terms of sight-reading skills (Sloboda, 1976; Wolf, 1976; for a review see Lehmann & McArthur, 2002), and to a lesser extent in terms of how visual feedback guides action during sight-reading (Banton, 1995; Lehmann & Ericsson, 1996; Ronkainen & Kussi, 2009). We are interested in the role that vision may play in developing representations of musical structure derived from watching oneself and others play an instrument. We investigated these issues in skilled guitarists, a group that we assumed would rely strongly on visual input during performance, and during acquisition of their skill.

Our interest in visual aspects of guitar skill stems from two intuitions derived from our own experience as guitar players. First, we assume the guitar itself encourages reliance on visual information. The guitar allows the same note to be played in different positions on different strings. For example, the Gibson Les Paul Junior guitar we used in our experiments has a range of 46 semitones spread across six strings and 22 frets (132 fretboard positions). Because of this ambiguity, guitar chords and scales are represented in print using chord charts (dots on a grid depicting a portion of the fretboard) or tablature (numbers indicating fret position placed on a grid specifying individual strings), specifying the to-be fretted positions. We assume that visual attention to these spatial patterns, both in print form and by visual inspection of the guitar during performance, allows guitarists to specify note and finger positions to choose the next note or chord quickly and efficiently. Indeed, guitar players often watch their hands when they play.

Second, we assume that formal and informal music training emphasize different aspects of visual knowledge. In formal training with instruments in the Western classical music tradition, musicians learn to sight-read musical notation, which could de-emphasize visual information from the instrument and the effectors used to play it. By contrast, as with learning of instruments in most non-Western cultures, many guitarists are trained informally; they learn to play by ear, many never learn to read music, and few rely on music notation while they play. Playing by ear frees attention from a musical score and encourages attention to visual aspects of the guitar (the fretboard and the strings) and the hands (finger placements in scales and chords). We assume that musical skill development in this context will encourage representation of musical structures in terms of their visual expressions on the guitar. Thus, in addition to representing musical structures in terms of auditory or kinesthetic events, we propose that guitarists represent notes, chords,

Music Perception VOLUME 30, ISSUE 1, PP. 37–47. ISSN 0730-7829, ELECTRONIC ISSN 1533-8312. © 2012 BY THE REGENTS OF THE UNIVERSITY OF CALIFORNIA ALL RIGHTS RESERVED. PLEASE DIRECT ALL REQUESTS FOR PERMISSION TO PHOTOCOPY OR REPRODUCE ARTICLE CONTENT THROUGH THE UNIVERSITY OF CALIFORNIA PRESS'S RIGHTS AND PERMISSIONS WEBSITE, HTTP://WWW.UCPRESSJOURNALS.COM/REPRINTINFO.ASP.DOI: 10.1525/MP.2012.30.1.37

scales, and musical phrases visually in formats obtained from watching themselves and others play the guitar.

We report two experiments that investigated the role of visual representations in skilled guitar playing, focusing on the identification and production of basic guitar chords. We had three major aims. First, we investigated whether guitarists develop visual representations of chords that depend on their guitar playing experience. Second, we investigated whether visual representations of chords are associated with actions to produce the chords. Third, we investigated whether visual representations of chords are associated with their corresponding sounds. We discuss the implications of our findings for understanding the multimodal nature of musical skill, and consider how the format of visual information can impact acquisition of musical skill.

Experiment 1

Experiment 1 investigated whether guitarists develop visual representations of guitar chords that depend on their experience with the guitar. A second aim of Experiment 1 was to determine whether these visual representations are associated with the actions required to produce the chords.

A hallmark of visual expertise in general is the finding that visual stimuli are easier to identify in familiar than unfamiliar orientations (Yin, 1969). Rotation effects are ubiquitous across practiced visual skills from identifying faces to cars (Diamond & Carey, 1986; Rossion & Gauthier, 2002), and offer a diagnostic tool for measuring the existence of experience-dependent visual representations. To address our first aim, we investigated whether guitar chord identification would be influenced by rotations applied to visual depictions of basic guitar chords (C, A, G, E, D, major chords played with some open string positions). Guitarists should identify chords faster in orientations they have more experience with than in orientations they have less experience with. Evidence of rotation effects in guitar chord identification would provide a first demonstration that guitarists acquire and utilize experience-dependent visual representations to recognize and produce chords.

We measured response time (RT) to vocally name the root of chords that were presented in one of three different visual forms: letter names denoting the root pitches of the chords, graphical chord charts, and photographs of a hand fingering chords on a fretboard (see Figure 1). Across trials, chord displays were presented in one of four rotations: 0°, 90°, 180°, or 270° clockwise. Letters are commonly viewed upright, so we expected faster RTs for 0° than the 90°, 180°, and 270° rotations. Chord charts are depicted with the guitar nut (a grooved strip holding the strings where the headstock is joined with the top of the fretboard) on top (0°), and sometimes with the nut on the left (270°), so we expected faster RTs for 0° and 270° than the 90° and 180° rotations. Guitarists would be familiar with viewing finger

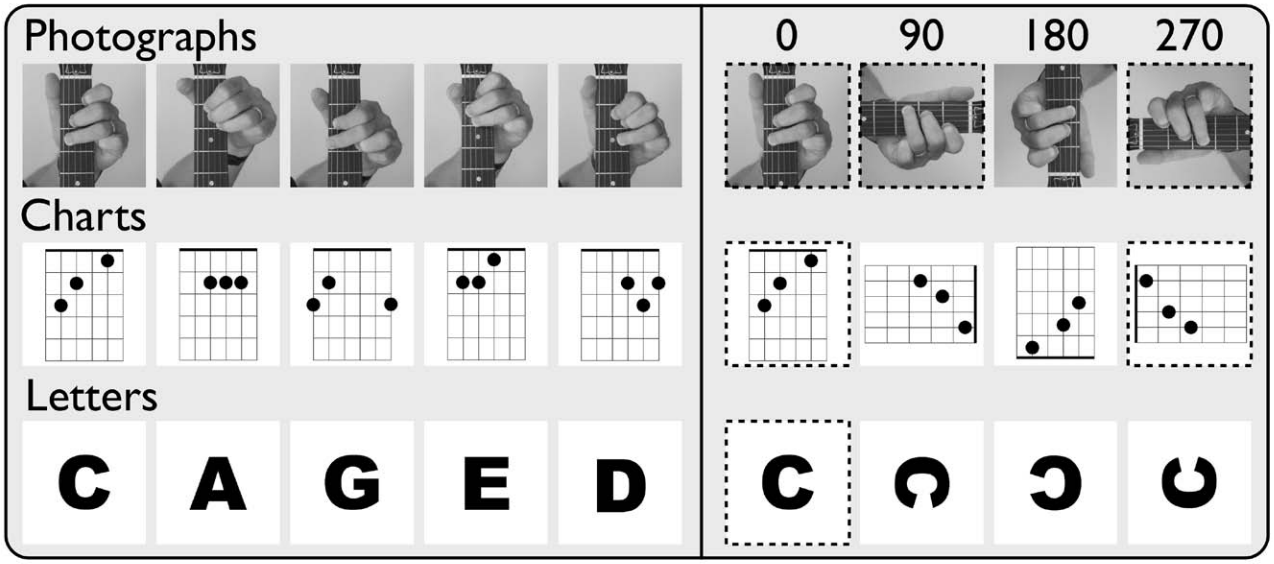

FIGURE 1. The left side shows visual formats (photograph, chart, letter) for each guitar chord (C, A, G, E, D) employed in Experiment 1. The photograph format was also employed for Stroop items in Experiment 2. The right side shows examples of the rotation manipulation (0°, 90°, 180°, 270°) for each display type. Displays framed by a dotted line represent rotations familiar to a guitarist, and unframed displays represent unfamiliar rotations.

placement on the neck at the 0°, 270° (most similar to watching yourself play guitar), and 90° (watching someone else play guitar) rotations, so identification RTs for photographs should be faster for 0°, 90°, and 270° than the 180° rotation. These predicted rotation effects would demonstrate that learning guitar produces experience-dependent visual representations for guitar chords.

Next, we aimed to investigate associations between visual representations and the motor actions required to play each chord on the guitar. We assessed visual-motor associations in a second phase instructing guitarists to identify chords by strumming them on a guitar. Comparing vocal RTs and guitar RTs offers insight into visual-motor associations. If there are no direct associations between visual representations and motor actions, then guitar RTs should measure the time required to name and then play a chord. Thus, guitar RTs should be longer than vocal RTs, which only involve naming time, for all target types (letters, charts, and photographs). However, if there are direct associations between visual representations and motor actions, then guitar RTs need not be longer than vocal RTs. Guitar RTs for letters may be longer than vocal RTs because subjects may name the letter and then play the named chord, but guitar RTs for charts and photographs may not be longer than vocal RTs because they may not require the intermediate step of naming the chord. However, we expected to find the same pattern of rotation effects as in the naming task.

Method

SUBJECTS

Thirty experienced guitarists were recruited from the Nashville community. They had played for an average of 20.9 years (range: 5-53), had started playing at 13.8 years of age (range: 11-20), and had 3.4 years of formal training (range: 0.5-11). Twenty had training in music theory, 24 were able to read music, and 28 were able to read tabulature. When asked to play for a few minutes to demonstrate their skill, no subject used sheet music. Some played music from memory but most improvised. All subjects were paid 20 dollars per 1 hour of participation. They had normal or corrected-to normal vision and spoke English as a first language. All were right-handed and fingered the guitar with their left hand.

APPARATUS & STIMULI

The experiment was conducted on an Intel 2.4 GHz MacBook Pro attached to a 15" SVGA monitor. The experiment was controlled by MATLAB using Psychophysics toolbox (Brainard, 1997; Pelli, 1997). Vocal responses were collected using a SHURE SM58 microphone connected via an ALESIS USB2 mixer fed into MATLAB. Guitar responses were collected from a Gibson Les Paul Junior electric guitar (graciously donated by Gibson Musical Instruments).

There were three chord display types (letter, chart, photograph). Each represented the open string (i.e., some strings not depressed), root positions for one of five major chords named by the letters: C, A, G, E, D. The letter displays were 6.5 cm x 6.5 cm displayed in Helvetica font. The charts depicted were 16.5 cm in height x 16.5 cm in width. The images depicted black lines indicating strings and frets (5 frets), with black dots indicating chord fingering. The letter and chord chart stimuli were presented on a white background square, 16.5 cm in height x 16 cm in width. The chord photograph displays were obtained by photographing the second author fingering each of the chords on the fretboard of a Gibson Les Paul Junior electric guitar. The digitized photographs were standardized to 16.5 cm x 16 cm square, and each photograph was cropped from the guitar nut down to just above the 5th fret on the fretboard. The width of the guitar neck in each photograph (as displayed on the computer screen) was 4.5 cm.

DESIGN AND PROCEDURE

There were two identification tasks involving vocal (name aloud the chord) or guitar responses (play the chord). Task order was counterbalanced across subjects. Chord display types were presented intermixed and in random order, and with equal frequency in each of the four rotations (0°, 90°, 180°, 270° clockwise). Zero degrees refers to upright letters and to chord charts and photographs with the nut of the guitar in the top position (see Figure 1). Each task involved 60 letter, chord chart, and photograph trials, for a total of 180 trials per task, and 360 trials for the experiment.

Trials began with a fixation cross (500 ms), followed immediately by the chord display. Depending on the task, subjects named the chord aloud or strummed the chord on the guitar. Chord displays remained onscreen until a response was registered, whereupon it was replaced by a blank screen. No feedback was presented. The next trial began 500 ms after the response. RTs for each trial were measured using the Psychophysics Toolbox voice-key script. The voice-key was triggered by the microphone or the guitar input by sounds that exceeded a loudness threshold. The threshold was set high enough so as not respond to ambient background noise, but low enough to be highly sensitive to all voice and guitar responses.

Results

For each subject, correct RTs in each condition were submitted to an outlier analysis (Van Selst & Jolicoeur, 1994), which trimmed an average of 3% of the observations in each condition. Vocal responses were analyzed for accuracy by hand. Guitar responses were analyzed for accuracy using MATLAB's fast Fourier transform function, which fed into a chord-classifying algorithm.[1] Mean RTs and error rates for each condition are displayed in Figure 2. Unless otherwise noted, an alpha criterion of .05 was adopted for all statistical tests.

The RT data were submitted to a 2 (response: vocal vs. guitar) x 3 (display type: letter, chart, photograph) x 4 (rotation: 0°, 90°, 180°, 270°) repeated measures analysis of variance (ANOVA). There was a main effect of response, $F(1, 29) = 5.74$, $MSE = 118766.00$, $\eta^2_p = .17$. Vocal RTs (708 ms) were shorter than guitar RTs (769 ms). There was a main effect of display type, $F(2, 58) = 102.23$, $MSE = 26542.90$, $\eta^2_p = .78$. RTs were much shorter for letters (621 ms) than chord charts (765 ms), $F(1, 29)=95.90$, $MSE = 26010.80$, $\eta^2_p = .77$, which were shorter than photographs (829 ms), $F(1, 29)=45.97$, $MSE = 10453.40$, $\eta^2_p = .61$. This finding suggests that RT was influenced by the visual complexity of the images. Finally, there was a main effect of rotation $F(3, 87) = 13.92$, $MSE = 7122.51$, $\eta^2_p = .32$. RTs in the 0° (715 ms) condition were shorter than RTs in the 90° (731 ms), 180° (771 ms), and 270° (737 ms) conditions, $F(1, 29) = 55.36$, $MSE =2301.40$, $\eta^2_p = .67$. RTs for 90° and 270° were shorter than 180°, $F(1, 29) = 9.73$, $MSE = 17135.20$, $\eta^2_p = .25$, but were not significantly different from one another, $F < 1$. These main effects were further qualified by second order interactions, which allowed us to investigate the presence of guitar-specific visual representations, and their associations to motor actions.

Importantly, the display type by rotation interaction was significant, $F(6, 174) = 7.40$, $MSE = 4948.12$, $\eta^2_p = .20$. We predicted specific patterns of rotation effects for each display type based on guitarists' pre-experimental familiarity with display types in particular rotations. The normal viewing orientation for letters is upright, and indeed RTs were significantly shorter for the 0° (608 ms) than 90° (629 ms), 180° (622 ms), and 270° (625 ms) rotations,

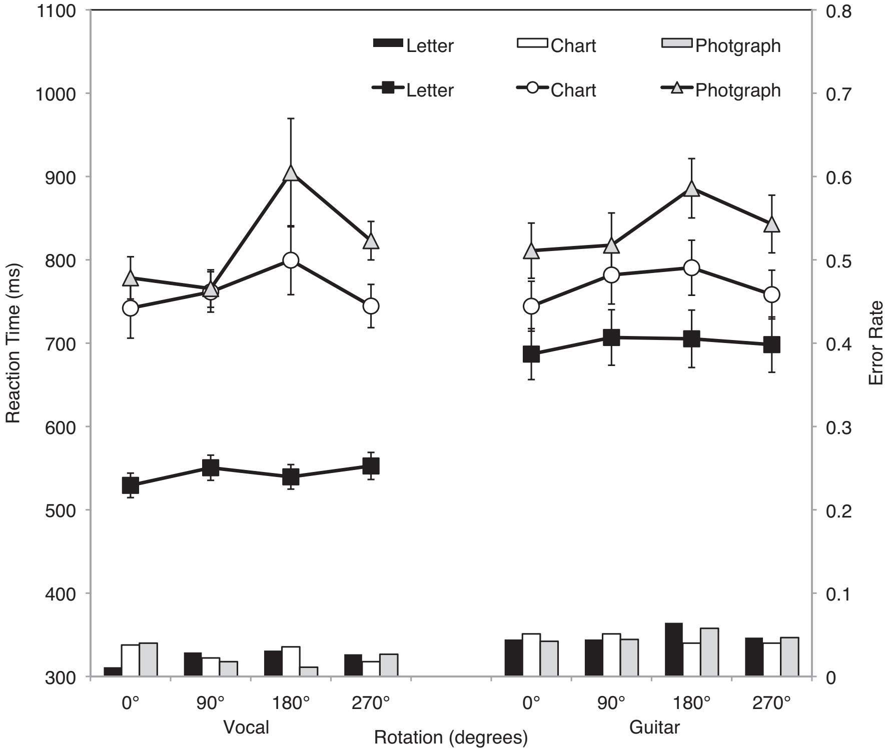

FIGURE 2. Mean reaction times (ms, with standard errors) and error rates from 30 guitarists for each condition in Experiment 1. RTs for each response (vocal vs. guitar), stimulus (photograph, chart, letter), and orientation (0°, 90°, 180°, 270°) are presented as lines, with milliseconds on the left y-axis. Error rates are presented as bars, with error rate on the right y-axis.

---

[1] The chord-classifying algorithm used Matlab's Fast Fourier Transform routine to find the power spectrum of each recording. For each recording, the frequency values for the top 10 amplitudes in the power spectrum were calculated. Each frequency value was assigned its corresponding note name and for each recording whichever note name occurred most frequently among the 10 frequencies was chosen as the chorded response. Spoils occurred when the classifying process did not yield a clear winner (i.e., when there were at least two note names chosen as the chorded response). The average spoil rate was 11%, and these trials were omitted from the RT and error analysis.

$F(1, 29) = 13.72$, $MSE = 991.89$, $\eta^2_p = .32$. For chord charts we assumed that guitarists would be familiar with 0° and 270° rotations (standard chord-chart orientations). RTs for 0° (743 ms) and 270° (751 ms) were not significantly different, $F < 1$, and were shorter than 90° (771 ms) and 180° (795 ms), $F(1, 29) = 37.55$, $MSE = 2075.93$, $\eta^2_p = .57$. For chord photographs we assumed that guitarists would be familiar with 0° and 270° rotations (most similar to first person viewpoints), and the 90° rotation (watching another guitarist finger chords). RTs for 0° (795 ms) and 90° (792 ms) were not significantly different, $F < 1$, and were shorter than RTs for 180° (895 ms) and 270° (833 ms), $F(1, 29) = 18.16$, $MSE = 16653.70$, $\eta^2_p = .39$. Finally, RTs for 270° were shorter than RTs for 180°, $F(1, 29) = 6.01$, $MSE = 19366.00$, $\eta^2_p = .17$, suggesting that guitarists had learned visual representations of guitar chords played by themselves and by other guitarists. The findings demonstrate that guitar players develop orientation sensitive visual representations for different visual depictions of chords.

Our second aim was to assess evidence for visual-motor associations comparing vocal and guitar RTs across display types. Importantly, there was a significant response by display type interaction, $F(2, 58) = 36.60$, $MSE = 11118.80$, $\eta^2_p = .56$. Planned contrasts indicated that letter RTs were shorter for vocal (543 ms) than guitar (699 ms) responses, $F(1, 29) = 25.34$, $MSE = 57828.30$, $\eta^2_p = .47$; this is a trivial effect, as people are well-practiced at naming letters. More important, chord chart RTs were not significantly different for vocal (762 ms) and guitar (769 ms) responses, $F < 1$; and photograph RTs were not significantly different for vocal (818 ms) and guitar (839 ms) responses, $F < 1$. This demonstrates that an equivalent amount of time was needed to name or play chords presented in chart or photograph formats, and suggests that these visual representations are directly associated with motor actions required to produce the chords.

A corresponding repeated measures ANOVA was conducted on the error rates for each condition. The main effect of task was marginally significant, $F(1, 29) = 3.95$, $MSE = 0.0225$, $p < .06$, $\eta^2_p = .12$. Mean error rates were higher in the guitar (.05) than vocal (.03) response conditions. The remaining main effects and higher order interactions were not significant.

Discussion

Experiment 1 establishes that guitarists have experience-dependent, orientation-specific visual representations of chords, and that these visual representations are associated with actions to produce each chord. The evidence for orientation specific visual representations of chords was particularly compelling, in that the pattern of rotation effects depended on pre-experimental familiarity with each display type. Specifically, RTs for the familiar viewing angles for chord charts (0° and 270°) were faster than RTs for unfamiliar angles, and RTs for the familiar viewing angles for chord photographs (0°, 90°, and 270°) were faster than RTs for unfamiliar angles. These rotation effects are expected given knowledge of the experience of guitarists. For letter and chord chart displays, we know that guitarists probably view letters in the upright position, and we can guess that chord charts are viewed in the 0° and 270° orientations by examining published chord charts, so we would expect faster RTs for these rotations than unfamiliar rotations. Most important were the rotation effects for chord photograph displays, which demonstrate that guitarists have experience viewing their own hands and other guitarists hands. We would not have predicted these effects by assuming that our guitarists were like classically trained musicians, who are trained to look at a musical score.

The second aim was to determine whether visual representations of chords are associated with motor actions. The response by display type interaction indicated that influences of perceptual and response complexity on RT were not additive. Letter RTs were shorter for vocal than guitar responses, but chart and photograph RTs were equivalent for vocal and guitar responses. Thus, RTs for vocal and guitar RTs were underadditive for chart and photograph displays, indicating these display types are independently associated with actions to name and play each chord. In support of this interpretation, recent studies investigating imitation learning of hand actions for guitar chords have shown that guitar chord photographs (similar to the photographs used in our study) activate the mirror neuron system that supports coding of finger positions for potential future action (Buccino et al., 2004), and that these neural substrates are utilized by guitarists when viewing and constructing a chord (Vogt et al., 2007). We would expect our photographs to activate similar brain regions, and it is interesting to speculate that chord charts – which convey finger position in a more abstract format – and letters – which produced the fastest guitar RTs, but conveyed no finger position information – also directly activate areas coding specification of actions. This suggestion fits with previous demonstrations in pianists that learning to read musical notation – which abstractly specifies finger-to-note mappings on the piano – activates motor-related cortical areas, both for explicit and implicit notation reading tasks (Stewart et al., 2003).

Experiment 2

The visual experience of watching oneself play a guitar chord is concurrent with the motor actions that execute the chord and the sonic consequences of the motor actions. The aim of Experiment 2 was to investigate whether visual representations of guitar chords are also associated with their auditory equivalents.

We employed a Stroop-like task (Stroop, 1935) to measure associations between photographs and sound of guitar chords. In a standard Stroop task subjects identify the ink-color of a color word. Reaction times are typically shorter for congruent (ink-color matches word) than incongruent (ink-color mismatches word) items, and this difference is termed the Stroop effect. Stroop effects are commonly used to investigate selective attention abilities, but can also index associations between target (e.g., color) and distractor (e.g., word) dimensions (MacLeod & Dunbar, 1988; Melara & Algom, 2003; Musen & Squire, 1993). Indeed, Stroop-like tasks have been investigated in pianists to tap learned associations between music notation and responses (Stewart, Walsh, & Frith, 2004; Zakay & Glicksohn, 1985), and between sounds and responses (Drost, Rieger, Brass, Gunter, & Prinz, 2005a, 2005b; Drost, Rieger, & Prinz, 2007).

We created congruent and incongruent chord photograph/sound pairs using the photographs from Experiment 1 and recordings of each chord being strummed. This created five congruent visual-audio items (same visual and auditory chord), and 20 incongruent visual-audio items (different visual and auditory chords).

We conducted separate visual and auditory chord identification versions of the task. The visual and auditory stimuli were presented simultaneously on each trial. The visual task involved naming the visual and ignoring the auditory version of the chord. The auditory task involved naming the auditory and ignoring the visual version of the chord. We tested both guitarists and nonguitarist musicians.

For guitarists, we expected that years of practice would produce strong associations between visual depictions of the chord fingerings on a fretboard and the corresponding sounds of each chord being strummed. If the Stroop task is sensitive to learned visual-audio associations for guitar chords, then we expected faster and more accurate identification for congruent than incongruent items. Moreover, if associations are bidirectional, we expected interference both from visual distractors on auditory identification and from auditory distractors on visual identification. For the nonguitarist musicians, we predicted no interference in either of the visual or auditory identification tasks, as these musicians would not have learned associations between visual and auditory forms of guitar chords.

Nonguitarists were given a short training session familiarizing them with the appropriate S-R mappings. We expected that the auditory task would be especially demanding for both groups as it requires an absolute judgment; nevertheless, we expected that participants would be able to perform the task. Accuracy is good when absolute judgments of pitch are given to small set-sizes of unidimensional stimuli (Pollack, 1952), and accuracy is even better for multidimensional auditory stimuli (Pollack & Ficks, 1954). The auditory chords contained sounds from all six strings, and offered multiple dimensions of information for the absolute judgment. Finally, we presented the correct response as visual feedback following each trial for all conditions and groups to facilitate learning of the S-R mappings.

Method

SUBJECTS

The 30 guitarists from Experiment 1 were tested. Nonguitarists were 30 experienced musicians from the Nashville community who reported no experience playing guitar. The nonguitarists had played for an average of 12.4 years (range: 4-22), had started playing at 8.2 years of age (range: 3-17), and had 8.3 years of formal training (range: 1-14). Twenty-eight reported the ability to read music notation. All subjects were paid 20 dollars per 1 hour of participation. All subjects had normal or corrected-to normal vision and spoke English as a first language.

APPARATUS & STIMULI

The experiment was conducted on the same apparatus used in Experiment 1. Auditory stimuli were presented over Sennheiser headphones.

The visual stimuli were the photographs for five open string, root position guitar chords (C, A, G, E, and D major chords) in 0° orientation as employed in Experiment 1 (see Figure 1).

The auditory stimuli were recordings of each of the five guitar chords, obtained by recording the third author (a professional guitarist) strumming each chord (a single, downward strum from the low to high E string) on the Gibson Les Paul Junior depicted in the photographs. Recordings were cropped to 2000 ms and normalized for volume.

DESIGN AND PROCEDURE

The factorial combination of the five visual and audio chords into visual-audio pairs yielded five congruent pairs and 20 incongruent pairs. Each task involved 50% congruent pairs and 50% incongruent pairs. Each congruent pair was presented 16 times and each incongruent pair was presented 4 times, for a total of 160 trials. Trials were presented intermixed in a randomized order with no constraints.

Guitarists received 160 trials in the visual identification task and 160 trials in the auditory identification task, and did not receive pre-training. Nonguitarists received the same 160 trials in the visual and auditory identification tasks, but did receive pre-training. Pre-training involved 40 trials of identifying the five visual chords presented without audio, and 40 trials of identifying the five audio chords presented without the photographs, with visual feedback indicating the correct response. For both practice and experimental trials, the correct response was always given as visual feedback after each trial. The order of visual/audio pre-training was counterbalanced, as was the order of visual and auditory identification tasks.

Subjects were seated in front of the microphone approximately 57 cm from the computer monitor. Trials began with a central fixation cross (500 ms), immediately followed by a target item. Subjects were instructed to name the target as quickly and accurately as possible. Targets were removed following the onset of the vocal response, and feedback indicating the correct response was displayed on screen. For both tasks, subjects were instructed to keep their eyes on the screen.

Results

For each subject correct RTs in each condition were submitted to an outlier analysis (Van Selst & Jolicoeur, 1994), which trimmed an average of 3% of the observations in each condition. Vocal responses were analyzed by hand for accuracy. Mean RTs and error rates for guitarists and nonguitarists in each condition are displayed in Table 1.

The primary question of interest was whether visual and auditory identification judgments made by guitarists would produce Stroop-like effects. For guitarists, we submitted mean RTs to a 2 (Task: visual vs. audio identification) x 2 (Congruency: congruent vs. incongruent) repeated measures ANOVA. The main effect of task was significant, $F(1, 29) = 65.35$, $MSE = 80766.50$, $\eta^2_p = .69$. Visual RTs (907 ms) were faster than auditory RTs (1327 ms). The main effect of congruency was significant, $F(1, 29) = 35.78$, $MSE = 4684.88$, $\eta^2_p = .55$, as was the task x congruency interaction, $F(1, 29) = 11.99$, $MSE = 4677.40$, $\eta^2_p = .29$. Separate one-way repeated measures ANOVAs for each task demonstrated Stroop-like effects for both visual and auditory tasks. Visual RTs were shorter for congruent (892 ms) than incongruent (923 ms) items, $F(1, 29) = 28.43$, $MSE = 523.92$, $\eta^2_p = .50$. Similarly, auditory RTs were shorter for congruent (1267 ms) than incongruent (1386 ms) items, $F(1, 29) = 23.62$, $MSE = 8838.35$, $\eta^2_p = .45$. The task x congruency interaction indicates larger Stroop-like effects for auditory than visual tasks.

The second question of interest was whether Stroop-like interference would be observed for the nonguitarists. For nonguitarists, mean RTs were submitted to the same repeated-measures ANOVA. The main effect of task was significant, $F(1, 29) = 40.44$, $MSE = 134477.00$, $\eta^2_p = .58$. Visual RTs (1427 ms) were shorter than auditory RTs (1853 ms). However, the main effect of congruency, $F < 1$, and the task x congruency interaction were not significant, $F < 1$. A separate mixed design ANOVA was conducted including guitarists vs. nonguitarists as a between-subjects factor. The critical group by congruency interaction was significant, $F(1, 58) = 10.69$, $MSE = 6760.37$, $\eta^2_p = .16$, indicating that congruency effects for guitarists were indeed larger than the nonsignificant congruency effects for nonguitarists.

Error rates for both guitarists and controls were submitted to the same repeated measures ANOVAs. For guitarists,

TABLE 1. **Mean Visual and Auditory Target Response Latencies (in Milliseconds, with Standard Errors in Parentheses, and Error Rates) for 30 Guitarists and 30 Controls in Experiment 2.**

	Target		Congruent (C)	Incongruent (I)	Stroop (I-C)
Guitarists	Visual	RT	892 (58)	923 (58)	31(6)
		ER	0.02	0.02	
	Audio	RT	1268 (64)	1386 (67)	118 (24)
		ER	0.08	0.16	
Control	Visual	RT	1429 (56)	1426 (54)	-3 (18)
		ER	0.05	0.06	
	Audio	RT	1847 (60)	1860 (62)	13 (29)
		ER	0.31	0.35	

the pattern of error rates resembled the pattern of RTs. The main effect of task was significant, $F(1, 29) = 60.91$, $MSE = 0.005$, $\eta^2_p = .68$; auditory error rates (.12) were higher than visual error rates (.02). The main effect of congruency was significant, $F(1, 29) = 41.83$, $MSE = 0.002$, $\eta^2_p = .60$; error rates for incongruent items (.09) were higher than for congruent items (.05). The task by congruency interaction was significant, $F(1, 29) = 38.46$, $MSE = 0.0013$, $\eta^2_p = .57$. For the auditory task, error rates were higher for incongruent (.16) than for congruent (.08) items, $F(1, 29) = 44.63$, $MSE = 0.0023$, $\eta^2_p = .61$; however for the visual task, error rates for incongruent (.02) and congruent (.02) items were not significantly different, $F < 1$.

For the nonguitarists, the main effect of task was significant, $F(1, 29) = 118.42$, $MSE = 0.019$, $\eta^2_p = .80$; auditory (.33) error rates were much higher than visual (.06) error rates. The main effect of congruency was significant, $F(1, 29) = 6.88$, $MSE = 0.0025$, $\eta^2_p = .19$; error rates for incongruent items (.21) were higher than for congruent items (.18). Last, the task x congruency interaction was not significant, $F(1, 29) = 1.72$, $MSE = 0.0025$, $p < .20$, $\eta^2_p = .06$.

Discussion

Experiment 2 demonstrates for the first time a Stroop-like effect in the domain of skilled guitar playing. For skilled guitarists, but not the for the nonguitarists, visual identification speed was influenced by auditory chord distractors, and auditory identification speed was influenced by visual chord distractors. Furthermore, the Stroop-like effect in guitarists was asymmetrical, with larger interference from visual onto the auditory dimension than from auditory onto the visual dimension. Asymmetries in Stroop interference are common in other variants of the task (Melara & Algom, 2003) and suggest underlying differences in the salience of target and distractor dimensions. For guitar chords, it appears that visual information is dominant over auditory information. This is consistent with findings that both groups were faster and more accurate when identifying the visual version of chords.

The presence of Stroop-like interference for visual and auditory targets suggests that visual representations of chords establish associations with chord sounds and vice versa. However, the present data do not specify whether the Stroop-like effects are driven by direct stimulus-stimulus associations or stimulus-response associations (Kornblum, Hasbroucq, & Osman, 1990). A stimulus-stimulus account would assume that distracting visual or auditory information directly interferes with perceptual processing of the target. On this view, the Stroop-like effect may be similar to other multisensory interference effects like the McGurk effect, where perception of an auditory phoneme /ba/ is perceived as /da/ when a viewer is simultaneously presented with a movie depicting a speaker pronouncing the phoneme /ga/ (McGurk & MacDonald, 1976). On the other hand, a stimulus-response account would not assume that distracting visual information directly interfered with auditory perception; instead visual information may have retrieved irrelevant response information that interfered with auditory identification at a response level. Although the present data do not rule out a stimulus-response account, we note that Experiment 1 demonstrated direct visual-action associations for chords, which at least establishes the possibility that the Stroop-like effects in Experiment 2 were driven by direct visual-auditory associations.

It is interesting that accuracy was higher for visual than audio chords in both guitarists and nonguitarists. We assume that visual chords are easier to identify than auditory chords because visual information affords fixed points of reference that allow people to identify a stimulus in an absolute fashion; whereas, without absolute pitch abilities, people are poor at identifying the pitch of isolated sounds, and instead use relative pitch to make their judgments. The visual advantage underscores the importance of visual representations in developing guitar skill: It is easier to classify the appearance of a chord than the sound of a chord even in novices, so guitarists may come to rely on visual representations as they develop skill, particularly when they learn from watching other guitarists play.

Given that auditory pitch identification is difficult without absolute pitch ability, it is perhaps surprising that our guitarists and nonguitarists (both of whom we assumed did not have absolute pitch) were able to accurately identify the auditory chords at above chance levels. Guitarists were more accurate than nonguitarists at auditory identification, and we assume this difference reflects the guitarists' experience with hearing and playing these particular chords. Our nonguitarists were also musicians, and their music training may partly explain their above chance performance. However, we do not assume that absolute pitch was necessary for accurate auditory identification. First, it is well known that people in general can identify absolute pitch for a small set-size of alternatives, as was the case in our experiment (Pollack, 1952); and, accuracy is better for absolute judgments of multidimensional auditory stimuli (Pollack & Ficks, 1954), and our guitar chords were multidimensional. As well, the correct answer was presented as visual feedback after each trial, and we assumed that subjects would learn diagnostic features of the chords as they progressed through the experiment. For these reasons we expected above chance performance in the auditory identification task. The specific auditory cues used by subjects to make their judgments remain unclear. It is possible that they were sensitive to average differences in pitch from all of the strings, and it is also possible that particular chords were judged by specific diagnostic features. For example, the G chord involves a G note on the 3rd fret of the high E string (the highest pitch of all the chords), and the E chord involves an E note from the low open E string (the lowest pitch of all the chords).

Finally, it is interesting that nonguitarists demonstrated a small but reliable Stroop-like effect in accuracy. We assume that this Stroop-like effect was not driven by the pre-existing visual and auditory associations that drive the effect in guitarists, but instead assume that the Stroop-like effect in nonguitarists reflects visual and auditory associations that developed over the course of the experimental session. Our design involved 25 visual-auditory items with five congruent and 20 incongruent items. Congruent items were presented more frequently than incongruent items to ensure a balance of 50% congruent and incongruent trials. We suggest that the Stroop-like effect for nonguitarists stems from an item-specific learning process sensitive to differences in frequency between congruent and incongruent items (Logan, 1988; Schmidt, 2007) that improved performance for the high frequency congruent items over the low frequency incongruent items, and that the performance improvement was reflected in accuracy rather than RT.

General Discussion

The results of two experiments provide novel evidence that guitarists represent chords in a multimodal fashion incorporating visual, auditory, and kinesthetic representations. Experiment 1 demonstrated that visual representations of guitar chords are orientation specific and directly associated with their corresponding motor actions. Experiment 2 demonstrated that visual representations of guitar chords have bidirectional associations with their corresponding auditory sounds.

Although we argue that guitarists rely on vision partly for reasons specific to the instrument and the more informal nature of their training experience, we do not claim that reliance on visual cues is unique to guitar skill. Indeed, determining the role of vision across different instrumentalists is an interesting avenue for further research.

Musical skill in general is multimodal in nature, and the notion that musicians develop crossmodal associations throughout acquisition of musical skill has been a topic of recent interest (Drost et al., 2005a, 2005b, 2007). For example, the ideomotor hypothesis assumes that perception and action information are integrated together in units referred to as event files (Hommel, 1998). In a musical context, the ideomotor hypothesis predicts that repeated experience with particular actions (e.g., playing a C chord) and their consequent effects (e.g., a C chord is sounded) will lead to the formation of action-effect associations that are bidirectional in nature. For example, in experiments similar to our own, Drost et al. (2005a) demonstrated that skilled pianists develop bidirectional associations between chord actions and chord sounds. Our present findings, which demonstrate visual-motor and bidirectional visual-auditory associations, fit within the view that associative learning promotes integration of information across modalities in musical skill.

On the one hand, our evidence for visual-motor and visual-auditory associations could reflect a passive automatic learning process. On this view, repeatedly watching oneself play an instrument would cause associations to form between visual, audio, and kinesthetic aspects of chord production. These associations would develop automatically as a consequence of viewing the actions, even if guitarists do not actively rely upon them during performance. On the other hand, bidirectional auditory-action or visual-action associations may play a more central role in planning and guiding actions during performance. For example, Drost et al. (2005a) suggested that pianists use bidirectional associations between actions and sounds to prepare upcoming actions by mentally imaging upcoming auditory events. Similarly, violinists apparently make use of bidirectional visual-action associations when judging the visual onset of another violinist initiating their performance, and perhaps rely on these associations to facilitate synchronization when playing with others (Wöllner & Cañal-Bruland, 2010). Similarly, we suggest that guitarists may prepare for upcoming actions by relying on visual simulations of the actions required to produce the desired auditory effects.

It is worth considering the role of visual information in acquiring musical skill. Experiment 1 showed that visual format influences identification and production of guitar chords (e.g., letter RTs < chart RTs < photograph RTs). This result could easily be attributed to differences in visual complexity between the display types. At the same time, each display may offer particular affordances, especially for the novice guitarist. For example, the photographs depict how the fingers should be arranged to produce a chord on a fretboard. These displays are known to activate the mirror neuron system (Buccino et al., 2004; Vogt et al., 2007), potentially facilitating the transformation of the visual display into desired finger placements. Brain regions involved in action imitation are important because they allow one to specify some details of potential actions through observation. We suggest that learning musical sequences by observing another guitarist, either in real life or in video format, would allow the neural substrates involved in action imitation to be activated, and possibly facilitate acquisition of the musical sequence. Indeed it would be interesting to compare musical sequence learning between visual formats that activate the mirror neuron system (e.g., watching another guitarist), and those that do not (e.g., musical notation or tablature).

We assume that guitarists actively rely on visual information, but we must be cautious in generalizing our findings to the context of playing music on the guitar. Both Experiments 1 and 2 employed a choice RT procedure, which requires a response as quickly as possible after target onset. Playing music requires serially ordering responses at particular moments in time. These different requirements restrict our ability to generalize from our current findings to a normal performance context. Indeed, a topic for future research is to better understand how visual representations contribute to performance of more complex chord forms and sequences of notes and chords. We expect that visual information can provide feedback for corrective action, aid recognition of chord and note sequences played by other guitarists, and deliver important cues for synchronization while playing with other musicians.

Author Note

We would like to thank Bruno H. Repp and two anonymous reviewers for their insightful comments on previous version of this manuscript. We are grateful to Gibson Musical Instruments for donating the guitars used in the current research.

Correspondence concerning this article should be addressed to M. J. C. Crump (E-MAIL: mcrump@brooklyn.cuny.edu), Department of Psychology, Brooklyn College of CUNY, 2900 Bedford Avenue, Brooklyn, NY, 11210, or G. D. Logan (E-MAIL: gordon.logan@vanderbilt.edu), Department of Psychology, Vanderbilt University, Wilson Hall, 111 21st Avenue South, Nashville, TN, 37203.

References

BANTON, L. J. (1995). The role of visual and auditory feedback during the sight-reading of music. *Psychology of Music, 23,* 3–16.

BRAINARD, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436.

BUCCINO, G., VOGT, S., RITZL, A., FINK, G. R., ZILLES, K., FREUND, H. J., & RIZZOLATTI, G. (2004). Neural circuits underlying imitation learning of hand actions: An event-related fMRI study. *Journal of Cognitive Neuroscience, 16,* 114–126.

DIAMOND, R., & CAREY, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General, 115,* 107–117.

DROST, U. C., RIEGER, M., BRASS, M., GUNTER, T. C., & PRINZ, W. (2005a). Action-effect coupling in pianists. *Psychological Research, 69,* 233–241.

DROST, U. C., RIEGER, M., BRASS, M., GUNTER, T. C., & PRINZ, W. (2005b). When hearing turns into playing: Movement induction by auditory stimuli in pianists. *The Quarterly Journal of Experimental Psychology, 58A,* 1376–1389.

DROST, U. C., RIEGER, M., & PRINZ, W. (2007). Instrument specificity in experienced musicians. *The Quarterly Journal of Experimental Psychology, 60,* 527–533.

HOMMEL, B. (1998). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition, 5,* 183–216.

KORNBLUM, S., HASBROUCQ, T., & OSMAN, A. (1990). Dimensional overlap: Cognitive basis for stimulus-response compatibility–A model and taxonomy. *Psychological Review, 97,* 253–270.

LEHMANN, A. C., & ERICSSON, K. A. (1996). Performance without preparation: Structure and acquisition of expert sight-reading and accompanying performance. *Psychomusicology, 15,* 1–29.

LEHMANN, A. C., & MCARTHUR, V. (2002). Sight-reading. In R. Parncutt & G. E. McPherson (Eds.), *The science and psychology of music performance: Creative strategies for teaching and learning* (pp. 135–150). New York: Oxford University Press.

LOGAN, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

MACLEOD, C., & DUNBAR, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 126–135.

MELARA, R. D., & ALGOM, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review, 110,* 422–471.

MCGURK, H., & MACDONALD, J. (1976). Hearing lips and seeing voices. *Nature, 264,* 746–748.

MCPHERSON, G. E. (1994). Factors and abilities influencing sightreading skill in music. *Journal of Research in Music Education, 42,* 217–231.

MUSEN, G., & SQUIRE, L. R. (1993). Implicit learning of color-word associations using a Stroop paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 789–798.

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

Pelli, D. G. (1997). The *VideoToolbox* software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Pollack, I. (1952). The information of elementary auditory displays. *Journal of the Acoustical Society of America, 24*, 745–749.

Pollack, I., & Ficks, L. (1954). Information of elementary multi-dimensional auditory displays. *Journal of the Acoustical Society of America, 26*, 155–158.

Ronkainen, S., & Kussi, T. (2009). The keyboard as part of a visual, auditory, and kinesthetic processing in sight-reading at the piano. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, & P. S. Eerola (Eds.), *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music 2009* (pp. 453–458). Jyväskylä University Digital Archive.

Rossion, B., & Gauthier, I. (2002). How does the brain process upright and inverted faces. *Behavioral and Cognitive Neuroscience Reviews, 1*, 63–75.

Schmidt, J. R. (2007). Contingency learning without awareness: Evidence for implicit control. *Consciousness and Cognition, 16*, 421–435.

Sloboda, J. A. (1976). Visual perception of musical notation: Registering pitch symbols in memory. *Quarterly Journal of Experimental Psychology, 28*, 1–16.

Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003). Brain changes after learning to read and play music. *Neuroimage, 20*, 71–83.

Stewart, L., Walsh, V., & Frith, U. (2004). Reading music modifies spatial mapping in pianists. *Perception and Psychophysics, 66*, 183–195.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Van Selst, M., & Jolicoeur, P. (1994). A solution to the effect of sample size on outlier elimination. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 47*, 631–650.

Vogt, S., Buccino, G., Wohlschläger, A. M., Canessa, N., Shah, N. J., Zilles, K., et al. (2007). Prefrontal involvement in imitation learning of hand actions: Effects of practice and expertise. *NeuroImage, 37,* 1371–1383.

Wolf, T. (1976). A cognitive model of musical sight-reading. *Journal of Psycholinguisitic Research, 5*, 143–171.

Wöllner, C., & Cañal-Bruland, C. (2010). Keeping an eye on the violinist: Motor experts show superior timing consistency in a visual perception task. *Psychological Research, 74*, 579–585.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology, 81*, 141–145.

Zakay, D., & Glicksohn, J. (1985). Stimulus congruity and S-R compatibility as a determinants of interference in a Stroop-like task. *Canadian Journal of Experimental Psychology, 39*, 414–423.