

# The planning and execution of short auditory sequences

PETER E. KELLER

*Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany  
and University of Finance and Management, Warsaw, Poland*

and

IRING KOCH

*Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany  
and RWTH Aachen University, Aachen, Germany*

Action planning, but not action execution, in speeded tasks is typically faster when responses and their effects are compatible than when they are incompatible. We tested whether response–effect compatibility (REC) affects the execution of music-like sequential actions that require temporal regularity rather than rapidity. Musicians responded to metronomic visual stimuli by producing sequences of three taps at a specific tempo on three vertically aligned keys. Each tap triggered a tone. Key-to-tone mapping was either compatible or incompatible in terms of spatial height and pitch height. The results indicated that tap timing was more accurate with compatible than with incompatible mappings, both for taps produced before (Tap 1) and after (Taps 2 and 3) the onset of auditory feedback. Thus, the observed influence of REC on action execution was not due exclusively to actual auditory feedback. The anticipation of distal action effects may be involved in planning the dynamics of temporally precise movements.

The planning and execution of action sequences is a part of daily life, no matter how sublime (playing music) or mundane (washing dishes). Most such actions are more or less voluntary and intended to give rise to certain desired effects (enchanting sounds or sparkling kitchenware). Voluntary action planning is often studied in the laboratory with speeded reaction time (RT) tasks in which participants respond to stimuli by making prespecified movements as quickly as possible. Research employing the speeded RT paradigm has revealed that action planning involves the anticipation of action effects: RTs are typically faster when movements are compatible with the ensuing action effects than when the movements and effects are incompatible. For example, Kunde (2001; Kunde, Koch, & Hoffmann, 2004) has demonstrated that soft keypresses are initiated more rapidly when they trigger quiet tones than when they trigger loud tones, whereas the reverse is true for forceful keypresses. Thus, movement initiation times can be affected by auditory effects that are not yet physically present, supporting the ideomotor hypothesis that action effects are imagined during action planning (Greenwald, 1970; Hommel, Müssele, Aschersleben, & Prinz, 2001; James, 1890; Prinz, 1987). Similar response–

effect compatibility (REC) effects have been observed for various types of discrete actions, as well as for action sequences (for reviews, see Hoffmann, Stoecker, & Kunde, 2004, and Koch, Keller, & Prinz, 2004).

The present article is concerned with the planning and execution of action sequences with music-like auditory effects. In a recent speeded RT study of auditory sequence planning, Keller and Koch (2004) required participants to respond to each of four arbitrary (color patch) stimuli by producing a unique sequence of three taps on three vertically aligned keys. Action effects consisted of tones that were triggered by the taps. Planning in this task involved selecting which movement sequence to execute in response to a given stimulus and setting the movement control parameters required to execute this sequence with maximum speed.

In a series of experiments, Keller and Koch (2004) varied the compatibility of the mapping between response location (key) and effect pitch (tone) between blocks. Note that participants would have been unable to anticipate action effects reliably if key-to-tone mapping varied within blocks. In the compatible condition, taps on the top, middle, and bottom keys triggered high-, medium-, and low-pitched tones, respectively. Across several incompatible conditions, this key-to-tone mapping was reversed or scrambled (while remaining fixed within blocks) or taps on different keys triggered the same tone. This compatibility manipulation was based on the assumption that there is a “natural” (preexisting, instrument independent) cross-modal correspondence between pitch height and spatial height (see Mudd, 1963). Keller and Koch found that ac-

---

We thank Silvija Mikerevic and Heike Mittmann for running the experiment, and Arthur Markman plus three anonymous reviewers for comments on an earlier version of the manuscript. Correspondence relating to this article may be sent to P. E. Keller, Department of Psychology, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstrasse 1a, D-04103 Leipzig, Germany (e-mail: keller@cbs.mpg.de).

tion planning was most efficient (i.e., RTs were fastest) for compatible key-to-tone mappings. This pitch height REC effect was observed most strongly for participants with formal music training.

In general, REC effects on planning in speeded RT tasks can be explained by the mutual priming of “codes” representing a future action’s proximal (proprioceptive, tactile) and distal (auditory) effects (Kunde et al., 2004). According to this account, the code activation threshold at which an associated motor program is triggered is reached sooner when distal and proximal effects share similar features (e.g., an ascending keypress sequence producing an ascending melodic contour) than when they are dissimilar.

In contrast to action planning, action execution—for example, the speed with which movements are carried out once initiated—is usually not susceptible to REC effects in speeded RT tasks. This may be because rapid movements are fully preplanned and then controlled in an open-loop fashion (see Schmidt & Lee, 1999). Accordingly, peak force was not affected by the compatibility between tone loudness and required force in the study by Kunde et al. (2004), and the duration of intertap intervals was not influenced by the various key-to-tone mappings in the study by Keller and Koch (2004). The latter finding squares with Lashley’s (1951) assertion that sensory feedback does not play a crucial role in the execution of rapid, well-learned sequential movements.

The aim of the present study was to test whether REC affects movement execution when the task requires short auditory sequences to be produced at a specific tempo—as is typical in music performance—rather than as quickly as possible. The requirement to be regular rather than rapid might be expected to elicit REC effects on movement execution, because sensory feedback appears to play a role in online motor control when temporal precision is required. It has been demonstrated that the addition of auditory or visual feedback can improve performance on motor tasks that require precise temporal control (Aschersleben & Prinz, 1995; Mechsner, Kerzel, Knoblich, & Prinz, 2001), and, conversely, skilled performance can be disrupted by eliminating or altering the sensory feedback that a performer is accustomed to (see, e.g., Pfordresher, 2003; Proteau, Marteniuk, & Lévesque, 1992).

The present study bridged the issues addressed in speeded RT studies of action planning, on one hand, and studies on the role of sensory feedback in skilled motor performance, on the other, by examining the effects of anticipated and actual action effects within a single experimental paradigm. To this end, the requirements of the task employed by Keller and Koch (2004) were amended so that now participants were instructed to respond to isochronously flashing visual stimuli by making sequences of three taps at a rate that continued the tempo set by the stimuli. Similar to the situation in the Keller and Koch task, each tap triggered a tone, and REC was manipulated to create one compatible and two incompatible (reversed or scrambled) conditions, in terms of the relationship between the spatial height of the three response keys and the pitch height of the three tones.

If REC affects the execution of sequential actions that require temporal precision rather than rapidity, tap timing should be more accurate (i.e., close to the target tempo) with the compatible key-to-tone mapping than with the incompatible mappings. Finding that REC affects timing accuracy for the first tap of each sequence would provide evidence for action effect anticipation, because auditory feedback was not presented until the finger made contact with the response key. Assuming sufficient time for response selection and movement preparation in our task (see the Method section below), such a result would suggest that anticipated auditory feedback affects the dynamics of action execution (perhaps in addition to affecting planning speed) when a precise temporal goal must be met. Effects of REC on timing accuracy for the second and third taps would be attributable to the influence of actual auditory feedback on online motor control and/or to a persisting influence of anticipated feedback.

## METHOD

### Participants

Eighteen formally trained, practicing musicians—experts at producing regular action sequences—were paid to participate in the experiment (12 females, 6 males; mean age = 24.2 years, range = 18–32 years). The musicians had an average of 10 years’ experience (range = 5–16 years) performing on a variety of instruments. All musicians were right-handed.

### Apparatus and Stimuli

The experiment took place in a sound-attenuating room. A response box with three vertically aligned metal plates, which served as response keys, was positioned in front of a computer at a comfortable height for tapping. The response box itself was 250 mm high and 150 mm wide. The metal plates embedded within its surface measured 30 × 30 mm and were separated by a vertical distance of 15 mm. The box was angled so that its surface was off-vertical (by 30°) to avoid excessive fatigue in the participants’ tapping arms. Imperative stimuli consisted of four patches of color (pink, yellow, green, or blue, measuring 16 × 16 mm and viewed from a distance of approximately 750 mm) presented in the center of the black computer screen. A white color patch with similar dimensions and presentation location served as a neutral stimulus (see the Design and Procedure section below).

The tones used as response effects were presented through PRO-10 earphones in a marimba timbre at a comfortable loudness level (60 dBA). Three different pitches were used: high = A<sub>4</sub> (440 Hz), medium = G<sub>4</sub> (392 Hz), and low = F<sub>4</sub> (349 Hz). All tones had a sharp onset followed by natural decay (lasting approximately 500 msec if not interrupted by the presentation of another tone). Stimulus presentation and response registration were controlled by the Experimental Run Time System software (Berisoft), and tones were generated by a SoundBlaster ISA soundcard.

### Design and Procedure

We employed a single-factor repeated measures design in which REC was manipulated by varying the relationship between response key location and feedback tone pitch to form three conditions: compatible, incompatible reversed, and incompatible scrambled.

The participants were tested individually, and each received both written and oral instructions before beginning the task. The experiment consisted of 13 (1 practice plus 12 test) blocks of 20 trials and typically took 30–40 min to complete (plus approximately 15 min of instruction time, during which the participants memorized the stimulus–response sequence mappings). At the start of each trial, a

fixation cross appeared at the center of the computer screen. When the participant was ready, he or she depressed the space bar on the computer keyboard, which immediately triggered the following sequence of events. First, one of the four imperative stimulus color patches appeared in place of the fixation cross and flashed three times with a 600-msec interonset interval (IOI; i.e., 200 msec of color patch plus 400 msec of blank screen). Then the neutral (white square) stimulus flashed three times in the same manner as the imperative stimulus (200 msec on, 400 msec off). Thus, a six-element visual pacing signal composed of five IOIs of 600 msec was presented (see Figure 1).

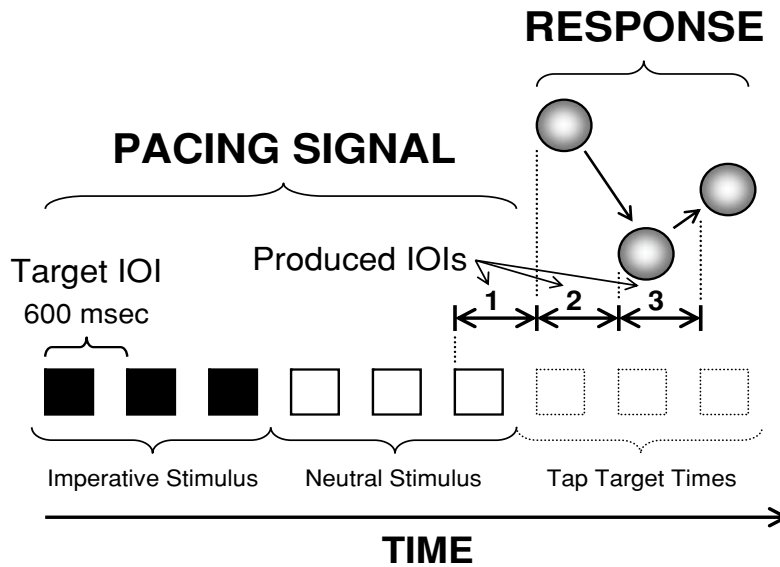
The participant was required to respond to the pacing signal by tapping the three response keys in one of four prescribed orders, depending on the color of the imperative stimulus, at the tempo set by the pacing signal. Perfect performance would entail tapping with a 600-msec IOI, starting 600 msec after the onset of the final neutral stimulus in the pacing signal (see Figure 1). For the *pink* stimulus, the correct tap sequence was top → middle → bottom; for *yellow*, the correct order was bottom → middle → top; for *green*, it was top → bottom → middle; and for *blue*, it was bottom → top → middle. The trial ended 500 msec after the third tap was registered by the computer. If a tap was made out of order, a feedback message (*Fehler*, German for “error”) appeared at the center of the screen and remained for 500 msec. The screen then went blank for 500 msec before the fixation cross that signaled the start of the next trial appeared (or the block ended if all trials were complete). A blank screen also intervened for 500 msec between correctly performed trials.

Throughout the 12 test blocks, each tap on a key triggered the presentation of a tone of high, medium, or low pitch. (The practice block, which served to allow the participant to rehearse the required

movements, was conducted without auditory feedback tones.) Three different types of key-to-tone mapping were employed in the test blocks, yielding one compatible and two incompatible (reversed and scrambled) REC conditions (see Table 1). In the compatible condition, a tap on the top key triggered the presentation of the high-pitched tone, a tap on the middle key triggered the medium tone, and a tap on the bottom key triggered the low tone. In the incompatible reversed condition, taps on the top, middle, and bottom keys triggered the low, medium, and high tones, respectively. In the incompatible scrambled condition, taps on the top, middle, and bottom keys were associated with medium, low, and high tones, respectively (producing a situation similar to lagged feedback). The condition (key-to-tone mapping) changed every four test blocks but remained fixed within each group of four blocks. The order in which the conditions were run was fully counterbalanced (i.e., six condition orders) across participants. The participants were told that even though the relationship between keys and tones would change twice during the experiment, the task remained unchanged: To respond to the color of the stimulus by tapping on the keys in the prescribed order as regularly and accurately as possible at the paced tempo. The tones were described as being irrelevant to the task.

**RESULTS**

Timing accuracy in each of the three REC conditions was assessed by examining mean produced IOIs and average absolute IOI error (i.e., the absolute value of the difference between produced IOIs and target IOIs) at each of the three tap positions within the response sequences.



**Figure 1.** An example of the experimental procedure in a single trial. A visual pacing signal with 600-msec interonset intervals (IOIs) was presented, wherein three flashes of a colored imperative stimulus (pink, yellow, green, or blue) were followed by three flashes of a neutral (white) stimulus. Participants responded to the pacing signal by making three consecutive finger taps on three vertically aligned response keys at a rate that continued the tempo set by the signal (target tap times are indicated by the dotted squares in the figure). The response keys were tapped in four different orders (only one is shown in the figure), with each order being associated with a different colored imperative stimulus. Each tap triggered a tone, and the key-to-tone mapping was manipulated between blocks to be either compatible or incompatible in terms of spatial height and pitch height. Tap timing accuracy was assessed by examining the three produced IOIs in terms of their closeness to the 600-msec target IOI.

**Table 1**  
**Movement Sequences and Auditory Effect Sequences From**  
**Compatible and Incompatible (Reversed and Scrambled)**  
**Response–Effect Compatibility Conditions**

Stimulus Color	Movement Sequence	Effect Sequence		
		Compatible	Reversed	Scrambled
Pink	1 → 2 → 3	1 → 2 → 3	3 → 2 → 1	2 → 3 → 1
Yellow	3 → 2 → 1	3 → 2 → 1	1 → 2 → 3	1 → 3 → 2
Green	1 → 3 → 2	1 → 3 → 2	2 → 3 → 1	2 → 1 → 3
Blue	3 → 1 → 2	3 → 1 → 2	2 → 1 → 3	1 → 2 → 3

Note—1, top key or high tone; 2, middle key or medium tone; 3, bottom key or low tone.

Absolute IOI error is potentially the most informative measure because it reflects deviations from target tempo. Trials in which the keys were tapped in the incorrect order (3% of trials) or in which any of the produced IOIs were longer than 900 msec (a further 7% of trials) were not analyzed. The remaining data from all but the first block in each REC condition were subjected to ANOVAs with the criterion for statistical significance set at  $\alpha = .05$  and the Greenhouse–Geisser correction applied.

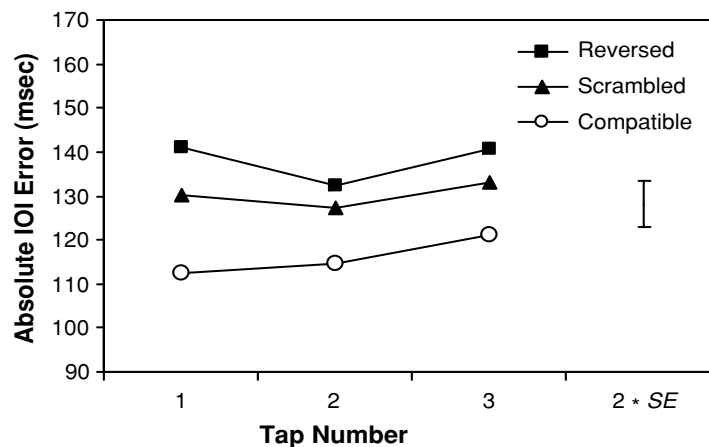
A  $3 \times 3$  ANOVA on mean IOIs with the factors REC (compatible, reversed, and scrambled) and tap number (1, 2, and 3) revealed that only the main effect of tap number was statistically significant [ $F(2,34) = 20.64, p < .001, \eta_p^2 = .55$ ]. The first IOI (693 msec) was longer than the second and third IOIs (485 and 478 msec, respectively). The REC main effect and the REC  $\times$  tap number interaction were not significant ( $ps > .1$ ). The significant main effect of tap number on mean IOIs may reflect a start-up cost (lengthening of the first IOI) associated with beginning to tap after the cessation of the pacing sequence, followed by an overly zealous attempt to compensate for this start-up cost by shortening the IOIs on subsequent taps.

Figure 2 shows average absolute IOI error across the three tap positions in each REC condition. Here it can be seen that absolute IOI error was generally smaller in

the compatible condition than in the incompatible conditions. This was confirmed by a significant main effect of REC in a  $3$  (REC)  $\times 3$  (tap number)  $\times 4$  (response sequence)  $\times 3$  (block) ANOVA on these data [ $F(2,34) = 9.89, p < .001, \eta_p^2 = .37$ ]. Pairwise comparisons revealed that the increase in absolute IOI errors from compatible to scrambled conditions was significant [ $F(1,17) = 9.16, p < .01, \eta_p^2 = .35$ ], but the increase from scrambled to reversed was not reliable ( $p > .1$ ). The tap number main effect and the REC  $\times$  tap number interaction were both far from significance ( $ps > .7$ ). Thus, the advantage enjoyed in the compatible condition was present across the three taps within the response sequences. Indeed, separate single-factor ANOVAs conducted on absolute IOI errors at each tap number all yielded a significant effect of REC ( $ps < .05$ ).

The ANOVA on absolute IOI errors revealed no significant main effects of response sequence or block and no significant interactions involving REC, response sequence, or block. Importantly, the REC  $\times$  block interaction was far from significance [ $F(4,68) = 0.37, p = .83, \eta_p^2 = .02$ ], which indicates that the effects of REC on absolute IOI error were not affected reliably by increasing exposure to the particular key-to-tone mapping within each condition. This confirms that the basis of the observed REC effect lies in a preexisting association between pitch and spatial height, rather than in specific key-to-tone associations formed during the experimental session. When we reran the ANOVA with condition order added as a between-groups factor, both the main effect of condition order and the interaction of condition order and REC condition were not significant ( $ps > .4$ ). Thus, the observed effects of REC on absolute IOI error are unlikely to be carryover effects.

Finally, a correlation analysis was conducted to explore the relationship between the effects of REC before and after the onset of auditory feedback. The degree of correlation between absolute IOI errors at Tap Positions 1



**Figure 2.** Average absolute interonset interval (IOI) error (i.e., |produced IOI – target IOI|) across the three sequential taps for each REC condition. The error bar on the right (labeled “ $2 * SE$ ”) represents double the standard error.

and 2 and at Tap Positions 2 and 3 was estimated separately for each participant across trials within each of the three REC conditions. The resulting Pearson correlation coefficients were then converted to Fisher  $z'$  scores before being entered into a  $3 \times 2$  ANOVA with the factors REC condition and tap pair (1 and 2 vs. 2 and 3). This analysis revealed that correlations between Taps 2 and 3 (mean  $r = .48$ ) were reliably stronger than those between Taps 1 and 2 (mean  $r = .08$ ) [ $F(1,17) = 23.77, p < .001, \eta_p^2 = .58$ ]; the main effect of REC and the REC  $\times$  tap pair interaction were not significant ( $ps > .4$ ). This result was corroborated by a nonparametric analysis conducted after  $z'$  scores were converted back to correlation coefficients following averaging across the three REC conditions. The degree of correlation between absolute IOI errors at Tap Positions 1 and 2 was lower than at Tap Positions 2 and 3 for 15 of the 18 participants (binomial test  $p < .01$ ).

## DISCUSSION

The present study investigated the execution of movement sequences with music-like auditory effects under conditions that required the sequences to be produced at a specific tempo rather than as quickly as possible. The results indicated that timing accuracy was better—that is, produced IOIs were closer to their target—when the mapping between response key location and effect tone pitch was compatible in terms of “height” than when it was incompatible (irrespective of the type of incompatibility, reversed or scrambled). This influence of REC on timing accuracy was observed across the three taps within sequences.

The presence of an REC effect on the first tap—which was made before the appearance of auditory feedback—is consistent with the results of work using speeded RT tasks showing that voluntary action is mediated by a mechanism that brings action effects to mind before they actually appear (see Hoffmann et al., 2004; Koch et al., 2004). The present findings extend the results of previous research by demonstrating that action effect anticipation not only plays a role in the rapid selection and initiation of speeded choice responses, it also plays a role in controlling the timing of movement so as to meet a precise temporal goal (i.e., to make a forthcoming action effect occur “on time”).

The presence of REC effects on the second and third taps may reflect a persisting influence of action effect anticipation, the influence of actual auditory feedback from the preceding tap(s), or a combination of these factors (see Greenwald, 1970). An exclusive influence of action effect anticipation is argued against by the fact that our correlation analyses revealed little evidence that REC effects before and after the onset of auditory feedback were related to one another, even though they were similar in magnitude. Thus, anticipated feedback and actual feedback may have affected movement timing differently. Although this finding seems to buttress claims that action planning and online control are distinct processes that rely on different types of information (see Glover, 2004), the results of our correlation analysis may be due at least in part to a com-

mon “temporal compensation” process driving the second and third taps (which were rushed) that was not present in generating the first tap (which was typically late).

Overall, the results of the present study demonstrate that the execution of auditory sequences is influenced by the compatibility between movements and their auditory effects when emphasis is placed on temporal precision rather than swiftness. The finding that the effects of REC extend beyond response selection and initiation suggests that compatibility between proximal and distal action effects affects more than the speed with which response execution thresholds are reached: Such compatibility may also influence the dynamics of an unfolding action, perhaps by affecting the parameters that control movement trajectories (see Erlhagen & Schöner, 2002). Schubö, Aschersleben, and Prinz (2001) found that the trajectories of sinusoidal arm movements can be biased by the amplitude of simultaneously presented sinusoidally moving visual stimuli. Our finding that REC affected movement timing before the onset of auditory feedback implies that such biases may be introduced not only by the perception of distal stimuli during action execution, but also by imagining such stimuli during planning. It may be that the movement control parameters that specify trajectory curvature, amplitude, and velocity adopt values that are more conducive to temporal precision (see Balasubramaniam, Wing, & Daffertshofer, 2004) when anticipated proximal and distal action effects are compatible rather than incompatible. Thus, incompatible mappings may come to disrupt timing accuracy by warping movements away from their optimal trajectories. We intend to use motion capture technology to pursue this idea in future research.

## REFERENCES

- ASCHERSLEBEN, G., & PRINZ, W. (1995). Synchronizing actions with events: The role of sensory information. *Perception & Psychophysics*, *57*, 305-317.
- BALASUBRAMANIAM, R., WING, A. M., & DAFFERTSHOFER, A. (2004). Keeping with the beat: Movement trajectories contribute to movement timing. *Experimental Brain Research*, *159*, 129-134.
- ERLHAGEN, W., & SCHÖNER, G. (2002). Dynamic field theory of movement preparation. *Psychological Review*, *109*, 545-572.
- GLOVER, S. (2004). Separate visual representations in the planning and control of action. *Behavioral & Brain Sciences*, *27*, 3-78.
- GREENWALD, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review*, *77*, 73-99.
- HOFFMANN, J., STOECKER, C., & KUNDE, W. (2004). Anticipatory control of actions. *International Journal of Sport & Exercise Psychology*, *2*, 346-361.
- HOMMEL, B., MÜSSELER, J., ASCHERSLEBEN, G., & PRINZ, W. (2001). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral & Brain Sciences*, *24*, 849-878.
- JAMES, W. (1890). *Principles of psychology*. New York: Holt.
- KELLER, P. E., & KOCH, I. (2004). *Action planning in sequential skills: Relations to music performance*. Manuscript submitted for publication.
- KOCH, I., KELLER, P. E., & PRINZ, W. (2004). The ideomotor approach to action control: Implications for skilled performance. *International Journal of Sport & Exercise Psychology*, *2*, 362-375.
- KUNDE, W. (2001). Response-effect compatibility in manual choice reaction tasks. *Journal of Experimental Psychology: Human Perception & Performance*, *27*, 387-394.
- KUNDE, W., KOCH, I., & HOFFMANN, J. (2004). Anticipated action ef-

- fects affect the selection, initiation, and execution of actions. *Quarterly Journal of Experimental Psychology*, **57A**, 87-106.
- LASHLEY, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior: The Hixon Symposium* (pp. 112-146). New York: Wiley.
- MECHNER, F., KERZEL, D., KNOBLICH, G., & PRINZ, W. (2001). Perceptual basis of bimanual coordination. *Nature*, **414**, 69-73.
- MUDD, S. A. (1963). Spatial stereotypes of four dimensions of pure tone. *Journal of Experimental Psychology*, **66**, 347-352.
- PFORDRESHER, P. Q. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 949-964.
- PRINZ, W. (1987). Ideo-motor action. In H. Heuer & A. F. Sanders (Eds.), *Perspectives on perception and action* (pp. 47-76). Hillsdale, NJ: Erlbaum.
- PROTEAU, L., MARTENIUK, R. G., & LÉVESQUE, L. (1992). A sensorimotor basis for motor learning: Evidence indicating specificity of practice. *Quarterly Journal of Experimental Psychology*, **44A**, 557-575.
- SCHMIDT, R. A., & LEE, T. D. (1999). *Motor control and learning: A behavioral emphasis*. Champaign, IL: Human Kinetics.
- SCHUBÖ, A., ASCHERSLEBEN, G., & PRINZ, W. (2001). Interactions between perception and action in a reaction task with overlapping S-R assignments. *Psychological Research*, **65**, 145-157.

(Manuscript received June 1, 2005;  
revision accepted for publication December 12, 2005.)