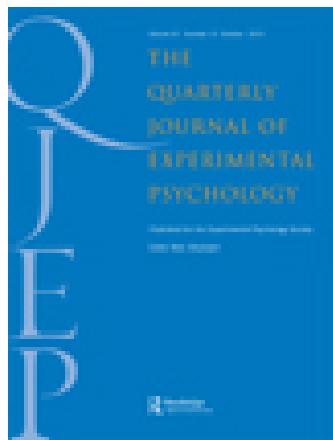


This article was downloaded by: [Harvard Library]

On: 11 March 2015, At: 10:11

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



[Click for updates](#)

The Quarterly Journal of Experimental Psychology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/pqje20>

Action-effect congruence during observational learning leads to faster action sequence learning

Jared C. Horvath^{ab}, Zachary Gray^b, Lukas Schilberg^{bc}, Ilya Vidrin^b & Alvaro Pascual-Leone^{bd}

^a Science of Learning Research Center, Graduate School of Education, University of Melbourne, Melbourne, VIC, Australia

^b Department of Neurology, Beth Israel Deaconess Medical Center, Berenson-Allen Center for Noninvasive Brain Stimulation, Division of Cognitive Neurology, Harvard Medical School, Boston, MA, USA

^c Faculty of Psychology and Neuroscience, Department of Cognitive Neuroscience, University of Maastricht, Maastricht, The Netherlands

^d Institut Guttmann de Neurorehabilitació Universitat Autònoma, Barcelona, Spain

Published online: 02 Mar 2015.

To cite this article: Jared C. Horvath, Zachary Gray, Lukas Schilberg, Ilya Vidrin & Alvaro Pascual-Leone (2015): Action-effect congruence during observational learning leads to faster action sequence learning, *The Quarterly Journal of Experimental Psychology*, DOI: [10.1080/17470218.2015.1012086](https://doi.org/10.1080/17470218.2015.1012086)

To link to this article: <http://dx.doi.org/10.1080/17470218.2015.1012086>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Action–effect congruence during observational learning leads to faster action sequence learning

Jared C. Horvath^{1,2}, Zachary Gray², Lukas Schilberg^{2,3}, Ilya Vidrin², and Alvaro Pascual-Leone^{2,4}

¹Science of Learning Research Center, Graduate School of Education, University of Melbourne, Melbourne, VIC, Australia

²Department of Neurology, Beth Israel Deaconess Medical Center, Berenson–Allen Center for Noninvasive Brain Stimulation, Division of Cognitive Neurology, Harvard Medical School, Boston, MA, USA

³Faculty of Psychology and Neuroscience, Department of Cognitive Neuroscience, University of Maastricht, Maastricht, The Netherlands

⁴Institut Guttmann de Neurorehabilitació Universitat Autònoma, Barcelona, Spain

Common coding theory suggests that any action (pressing a piano key) is intimately linked with its resultant sensory effect (an auditory musical tone). We conducted two experiments to explore the effect of varying auditory action–effect patterns during complex action learning. In Experiment 1, participants were assigned to 1 of 4 groups, watched a silent video of a hand playing a sequence on a piano keyboard with no auditory action effect (observation) and were asked to practise and perform the sequence on an identical keyboard with varying action effects (reproduction). During reproduction, Group 1 heard no auditory tones (identical to observed video), Group 2 heard typical scale–ascending piano tones with each key press, Group 3 heard fixed but out-of-sequence piano tones with each key press, and Group 4 heard random piano tones with each key press. In Experiment two, new participants were assigned to 1 of 2 groups and watched an identical video; however, the video in this experiment contained typical, scale–ascending piano sounds. During reproduction, Group 1 heard no auditory tones while Group 2 heard typical, scale–ascending piano tones with each key press (identical to observed video). Our results showed that participants whose action–effect patterns during reproduction matched those in the observed video learned the action sequence faster than participants whose action–effect patterns during reproduction differed from those in the observed video. Additionally, our results suggest that adding an effect during reproduction (when one is absent during observation) is somewhat more detrimental to action sequence learning than removing an effect during reproduction (when one is present during observation).

Keywords: Action–effect association; Observational learning; Action observation and execution; Stimulus–response compatibility; Common coding theory

With the proliferation of the internet and on-demand television, people are increasingly accessing didactic material presented via “video” format. A natural limitation of this detached

instructional tool is its lack of reciprocity: Because the teacher and learner cannot interact, information must be conveyed and mastered through demonstration and mimicry. Although this form of

Correspondence should be addressed to Jared C. Horvath, Melbourne School of Psychological Sciences, Redmond Barry Building #613, University of Melbourne, VIC–3010, Australia. E-mail: jch155@mail.harvard.edu

educational delivery may well suit the transmission of isolated information (e.g., dates or simple facts), it is unclear how conducive this format is to the conveyance of dynamic information (e.g., complex action learning). In this context, the term “dynamic” refers to skills or behaviours that generate unique and intelligible effects within the environment. For instance, when one is learning to play the piano, not only must each individual finger movement be generated and committed to memory (action), but also the resultant tone engendered by each key depression must be attended to and mastered (effect: for details, Prinz, 2010). In this example, although actions and sensory effects are encoded simultaneously, it remains somewhat unclear what effect each has on the overall learning of said action patterns and, accordingly, what should be present/absent in a “video” lesson.

There has been a long history of research exploring the link between motor and sensory integration. First suggested over 30 years ago, common coding theory (CCT) posits that actions are intimately linked with their resultant sensory effects (Prinz, 1984, 1990). For instance, the act of flipping on a light switch and the resultant sensation of visual illumination are not experientially exclusive events, rather they are co-relational and commonly encoded. Elsner and Hommel (2001) provided compelling support for this associative action–effect coding hypothesis. During a primary acquisition phase, participants heard task-irrelevant but consistent auditory tones in response to the pressing of two buttons. During a subsequent test phase, participants were presented with auditory tones and were asked to press one of the buttons in response. Elsner and Hommel reported faster motor response times when the tone was consistent with the earlier acquired action–effect correlation. This finding and variations thereof have since been replicated by several researchers (Drost, Rieger, Brass, Gunter, & Prinz, 2005; Elsner & Hommel, 2004; Herwig & Waszak, 2009; Hoffman, Sebald, & Stocker, 2001; Kunde, Hoffmann, & Zellmann, 2002).

As can be predicted, the majority of CCT research to date has been undertaken during conditions of *active* motor exploration. However, the

consequences of this theory have not been thoroughly examined with regards to *observational learning*: the reproduction of a novel pattern of action following the visual observation of one or more correct performances of said pattern (Doody, Bird, & Ross, 1985). As outlined above, online and on-demand videos are necessarily observational by nature. Possible mechanisms by which observation can lead to learning have recently been elucidated via imaging research, which has revealed common neural activation patterns during both action observation and action execution (Buccino et al., 2004; Gazzola & Keysers, 2009; Nelissen, Luppino, Vanduffel, Rizzolatti, & Orban, 2005; Rizzolatti & Sinigaglia, 2010). However, despite these findings, questions regarding action–effect relationships during observational learning still remain. Put differently, it remains unclear what role *observed* action effects versus *performed* action effects play in later skill reproduction and mastery.

Recently, Paulus, van Dam, Hunnius, Lindemann, and Bekkering (2011) explored an aspect of this question. During an observational learning phase, participants observed a model pressing two buttons, each producing a specific and consistent auditory tone. Later, during an execution phase, participants were presented with tones and were asked to press one of the buttons in response. Paulus et al. found that reaction times were shortened if the tone–button (action–effect) relationship was the same as that during the observation phase. Whereas this finding is certainly suggestive of action–effect transfer, it explores a somewhat simple reactionary (rather than *complex active*) reproduction of the observed actions. Therefore, it is difficult to extrapolate this finding to the learning and execution of a more complex observational action learning paradigm, such as playing a musical instrument.

To examine this concept, we conducted two experiments examining auditory effects on observational action learning. Over the course of several days, musically naïve subjects were alternately presented with a video of a skilled musician playing a 15-second action sequence on an electronic piano keyboard and were asked to reproduce said

sequence on an identical keyboard. In the first experiment, the observational video contained no auditory sounds. As such, auditory action–effect relationships were only present during the later sequence reproduction stages. In the second experiment, the video contained typical piano tones with each key press thereby creating an auditory action–effect relationship during the observational phase as well as during the reproduction phase. Experimental conditions varied according to the auditory tones (effect) generated following each keystroke (action) during sequence reproduction phases.

For our first experiment, as the observed video contained no auditory piano effects, we hypothesized that the auditory action–effect relationship experienced during the reproduction phase would dictate how the action sequence would ultimately be learned and encoded. More specifically, we anticipated that the addition of standard, scale-ascending piano tones during the reproduction phase would lead to faster learning than the addition of scale-independent, random, or no piano tones. We made this hypothesis as standard, scale-ascending piano tones represent an intuitive action–effect relationship congruent with likely expectations for what tones “should” have been present in the video (namely, as the piano key presses ascend step-wise, so too do the auditory tones produced). Conversely, the scale-independent, random, and no piano tones represent counterintuitive action–effect relationships incongruent with expectations and probably requiring additional effort to account for and successfully learn.

EXPERIMENT 1

Method

Subjects

Twenty-four male and 36 female subjects were enrolled and were randomly assigned to one of four conditions (outlined below). All subjects were musically naïve and had received no formal training in either a musical instrument or vocal choir (self-reported prior to enrolment). During

the study, two subjects dropped out (no auditory effects, \emptyset , and fixed-ordered, FO, conditions), and two subjects were not included in analysis due to recording errors in the computer system utilized (fixed-disordered, FD, and random, R, conditions).

Apparatus

Subjects sat at a table with a computer monitor positioned immediately to the left of an 88-key electronic piano keyboard. The monitor was used to present the action sequence video to the participants. Five keys in the middle of the keyboard were marked. The keyboard was connected to two external speakers (placed on either side of the subject) and an Apple computer. A MIDI computer program was used to record the sequence and timing of each button pressed during subject practice sessions. During practice, an external metronome set to 60 bpm was placed atop the keyboard. The metronome generated a constant, high-pitched electronic beep and a flashing red light.

Experimental conditions

Subjects were divided into four groups according to the auditory effects generated following each key press. It is worth noting here that the observational video that each participant watched and learned from (described below) contained *no* auditory effects following each key press. During the action sequence practice and performance stages, the first group heard no auditory effects (\emptyset condition—identical to the observational video), the second group heard fixed-ordered scale-ascending piano tones with each key press (FO condition—1:C, 2:D, 3:E, 4:F, and 5:G), the third group heard fixed-disordered, scale-independent piano tones with each key press (FD condition—1:D, 2:F, 3:E, 4:G, and 5:C), and the fourth group heard random (R condition) piano tones with each key press (Figure 1).

As noted above, we hypothesized that the action–effect relationship during the action sequence and performance stages would dictate how each person learned the sequence; more specifically, we assumed that standard,

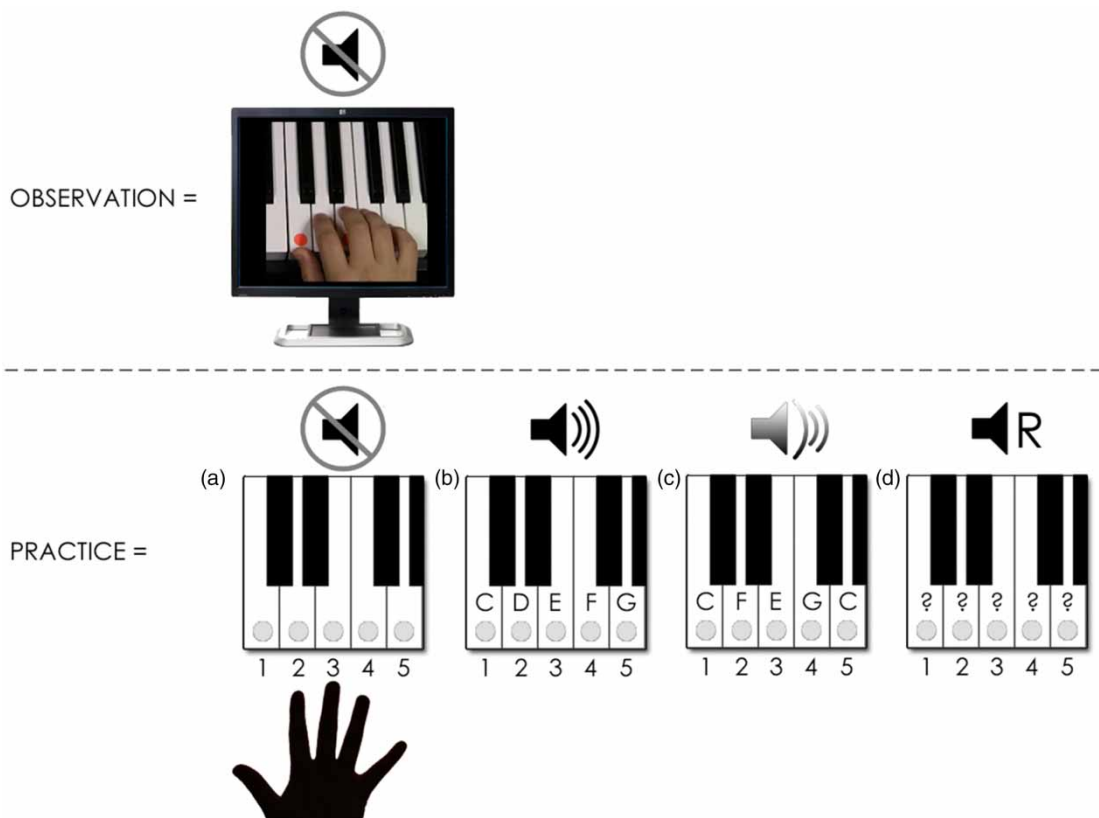


Figure 1. Experiment 1: Subjects observed a silent, first-person view of a musician playing an action sequence on the piano keyboard. When practising the sequence, subjects heard: (a) no auditory effect (\emptyset); (b) fixed-ordered (FO) scale-ascending piano tones; (c) fixed-disordered (FD) scale-independent piano tones; or (d) random (R) piano tones. To view this figure in colour, please visit the online version of this Journal.

scale-ascending piano tones would aid individuals in learning the sequence. We chose these four conditions to help elucidate this hypothesis. As the FO group experienced standard, scale-ascending piano tones, we expected these individuals to display faster learning. The FD group was included to determine whether learning effects could be attributed to any set of fixed auditory effects (rather than only the scale-ascending nature of the FO effects). The R group was included to determine whether learning effects could be attributed to the inclusion of any set of auditory effects, regardless of reliability or predictability. Finally, the \emptyset group was chosen as a baseline condition against which to compare any effects (though this was the only group to directly match the silent observational video).

Design and procedure

With the exception of the tones generated during each button press, each subject underwent an identical protocol. Participants came to the lab on three consecutive days for approximately 45 minutes each day. Each day consisted of alternating blocks of observing, practising, and performing the action sequence (Figure 2).

Video. The action sequence video showed a skilled musician playing a pattern using five marked keys in the middle of an electronic piano keyboard (the middle C, D, E, F, and G keys—hereafter referred to as 1, 2, 3, 4, and 5, respectively). Each key was exclusively pressed by either the thumb or one of the four digits. Although the key presses were silent, the video did contain a high-pitched

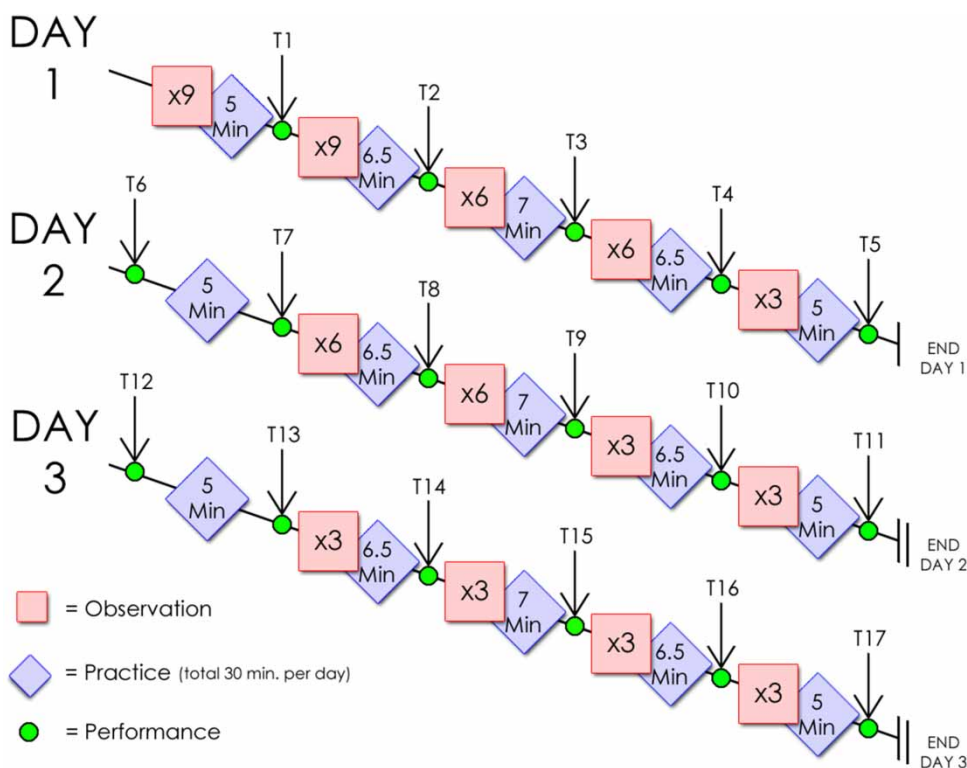


Figure 2. Experiment 1: Schematic of three-day training sequence. Red squares represent the observational phase (number of times consecutively viewing the video), blue squares represent the practice phase (in minutes), and green dots represent the reproduction phase (one “performance” per time point). In total, participants watched the action sequence video 63 times interleaved with 90 minutes of practice time and 17 “performances”. To view this figure in colour, please visit the online version of this Journal.

electronic beep playing at 60 bpm. The perspective of the video was first person from the view of the musician (Figure 1). During observation, subjects were asked to sit comfortably with their hands in their laps. They were not allowed to practise along with the video nor move their fingers whilst watching the video.

Sequence. The action sequence was approximately 15 s long and consisted of 29 key presses (1, 2, 3, 4, 5, 4, 3, 2, 1, 3, 5, 3, 4, 3, 2, 5, 1, 2, 3, 1, 2, 3, 4, 2, 5, 4, 3, 2, 1). Key presses were separated by exactly 500 ms.

Practice. Following action observation, subjects were given time to practise the sequence from memory. During practice, each button press generated

either no sound or auditory tones according to experimental condition (Figure 1). Subjects practised for a total of 30 min on each of the three days.

Performance. Following each practice block and at the start of Days 2 and 3, subjects were asked to perform the action sequence one time through as best they could. During this performance, the relevant tones continued to be produced, and the metronome remained on.

Data analysis

Over the course of three days, 17 performance sequences were recorded: one following each practice block and one at the start of Days 2 and 3 (Figure 2). Performances were graded and analysed separately on two dimensions: sequence and

timing. To grade each, a point system was developed. For the sequence score, subjects were granted one point for each note played in its correct sequential position (max sequence score = 29). For timing score, subjects were granted one point for each note played within 25 ms of the correct 500-ms interval (between 475 and 525 ms; max timing score = 28). We opted to analyse each dimension separately as research has demonstrated that, when learning a piano sequence, novices tend to focus primarily on sequence accuracy prior to sequence timing (Drake & Palmer, 2000; Gabriellson, 2003; Hallam, 1997). Accordingly, in order to differentiate any learning differences in these unique dimensions (and to avoid any plateau of accuracy impacting timing analysis), we have analysed each separately.

One-way analyses of variance (ANOVAs) were run to compare sequence and timing scores (separately) between the four conditions at T1 and at T17 to determine any differences in “starting” and “finishing” competency. A two-way mixed-model ANOVA (Group \times Time), Bonferroni-corrected independent-sample t -test was run to compare sequence accuracy and timing accuracy (separately) between groups. In order to better visualize performance dynamics, nonlinear regression curves were modelled and utilized to compare sequence and timing scores between groups across the entire experimental sequence; 95% confidence intervals (CIs) were determined for these regression curves.

Results

Sequence scores (Figure 3a)

A one-way ANOVA between groups for the first recording (T1) and the final recording (T17) showed no significant differences, suggesting that all groups began and ended the experiment at the same level of sequence competency [T1: $F(3, 52) = 0.652, p = .602$; T17: $F(3, 52) = 2.004, p = .125$]. A two-way ANOVA (Group \times Time) revealed a significant main effect of time, $F(16, 832) = 30.168, p < .001$, such that sequence scores significantly changed within groups across the study. This increase in scores can be understood as overall learning of the sequence. A significant main effect of group was also determined, $F(3, 52) = 4.018, p < .001$, suggesting a learning differential between conditions. The interaction effect was trending towards significance, $F(48, 832) = 1.422, p = .069$. Post hoc analysis showed a significant difference between the \emptyset and FO groups (Mean Difference = 2.4118, $p = .014$) such that the \emptyset group achieved higher scores than the FO group, but no other significant differences.

To determine which group learned the sequence fastest, a Michaelis–Menten nonlinear regression curve was fitted to each condition [$B_0 \times x / (B_1 + x)$; B_0 = maximum score, B_1 = point of half maximum score]. B_0 was constrained at 29 (maximum possible score), and standard error estimates were bootstrapped (Figure 4). A look at Table 1 shows a 5% trimmed CI differential

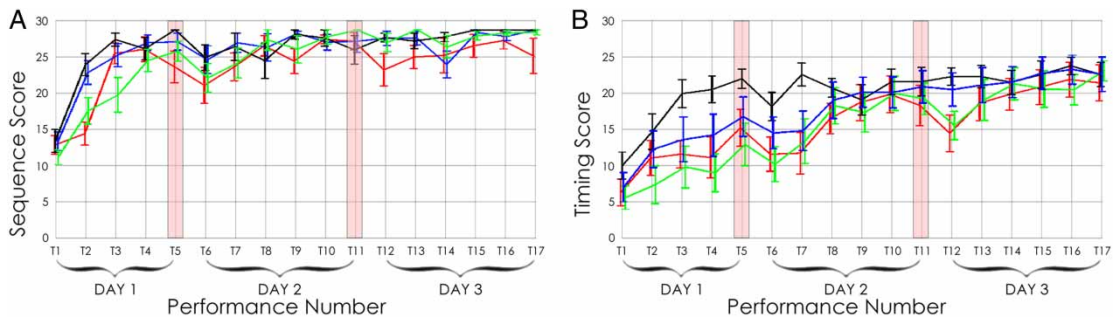


Figure 3. Experiment 1: (A) Average group sequence scores. (B) Average group timing scores. Black = \emptyset (no auditory effects); red = FO (fixed-ordered); green = FD (fixed-disordered); blue = R (random). Error bars represent ± 1 standard error. To view this figure in colour, please visit the online version of this Journal.

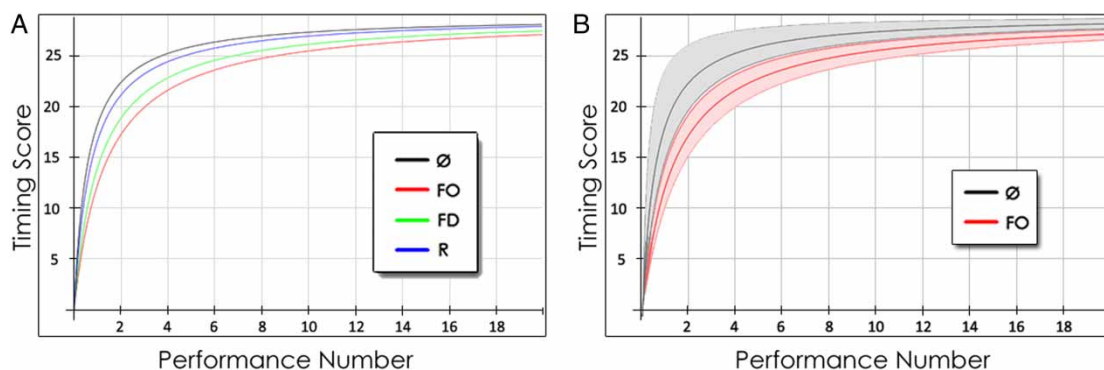


Figure 4. Experiment 1: Nonlinear regression curves for sequence scores; y axis = sequence score; x axis = performance number (time). \emptyset = no auditory effects; FO = fixed-ordered; FD = fixed-disordered; R = random. (A) Mean sequence score curves for all groups. (B) Mean sequence score curves showing confidence interval (CI) range comparing \emptyset and FO. To view this figure in colour, please visit the online version of this Journal.

between \emptyset and FO but not between any other group. This suggests that the \emptyset learned the sequence at a faster rate than the FO group.

To better elucidate differences in sequence learning speed, Michaelis–Menten nonlinear regression curves were fit to each individual (constraints as outlined above), and statistics were run between groups utilizing the B_1 value (the time point at which the learner reached half maximum possible score). A one-way ANOVA revealed a significant difference in B_1 values, $F(3, 52) = 4.562, p = .007$. Post hoc analysis showed a significant difference between the \emptyset and FO groups (MD = 0.847, $p = .011$) but no other significant differences (Figure 5).

Finally, uncorrected within-subjects t -tests were run between T5 and T6 and between T11 and T12

to explore sequence retention between study days. A significant difference was found between T11 and T12 for the FO group, $t(13) = 2.468, p = .028$. No other significant differences were found.

Timing scores (Figure 3b)

A one-way ANOVA between groups for the first recording (T1) and the final recording (T17) showed no significant differences, suggesting that all groups began and ended the experiment at the same level of timing competency [T1: $F(3, 52) = 0.625, p = .602$; T17: $F(3, 52) = 0.901, p = .447$]. A two-way ANOVA (Group \times Time) revealed a significant main effect of time, $F(16, 832) = 23.124, p < .001$, such that timing scores

Table 1. Nonlinear regression equations and statistics for sequence scores in Experiment 1

Group	Equation $B_0 \times x / (B_1 + x)$	B_1		R^2
		SE	5% trimmed CI interval	
\emptyset	$29x / (0.583 + x)$	0.160	0.218–0.921	.933
FO	$29x / (1.351 + x)$	0.199	1.004–1.784	.890
FD	$29x / (1.065 + x)$	0.134	0.752–1.264	.942
R	$29x / (0.735 + x)$	0.142	0.544–1.079	.942

Note: \emptyset = no auditory effects; FO = fixed-ordered; FD = fixed-disordered; R = random; CI = confidence interval; B_0 = maximum score, B_1 = point of half maximum score.

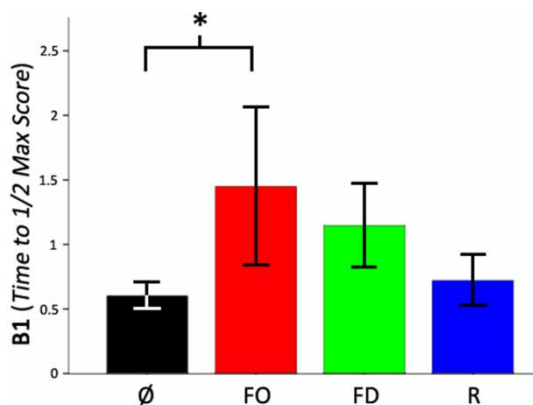


Figure 5. Experiment 1: Time needed to reach half the maximum sequence score (14.5). Error bars represent ± 2 standard errors. \emptyset = no auditory effects; FO = fixed-ordered; FD = fixed-disordered; R = random. To view this figure in colour, please visit the online version of this Journal.

significantly changed within groups across the study. This increase in scores can be understood as overall learning of the timing. Neither the main effect of group, $F(3, 52) = 1.928$, $p = .136$, nor the interaction effect, $F(48, 832) = 1.058$, $p = .369$, was significant.

Finally, uncorrected within-subjects t -tests were run between T5 and T6 and between T11 and T12 to explore timing retention between study days. No significant differences were found.

Discussion

In this experiment, we explored the effect of varying auditory action-effects during a reproduction stage on complex action learning following the observation of a video clip with no auditory action-effects. We anticipated that the auditory action-effect experienced during the reproduction stages would dictate the speed and accuracy of action learning. More specifically, we hypothesized that the FO group would learn both the sequence and timing faster than the other groups.

Interestingly, our data reveal that, whereas there were no significant differences in learning of the *timing* between groups, the \emptyset group (identical to the observed video) reached learning midpoint of the *sequence* significantly faster than the FO

group. This goes against our hypothesis and may suggest that what we anticipated to be a non-action-effect during the observational phase (in this case, actions with no tonal effect) may have, in fact, influenced later learning and performance. It is also interesting to note that there were no significant differences in the sequence learning curves of the \emptyset , FD, and R group. This may suggest that the state or quality of the effects during the practice and performance stages may impact learning speed.

There are two possible interpretations of these findings. The first is that the action effect encountered during the observational phase (even if the observed effect is silence) does have an effect on sequence learning. This supposition certainly appears to be supported by the learning curve slope differentials between the \emptyset and FO groups. Under this interpretation, the addition of auditory tones during the FO reproduction phases may have forced participants to “overwrite” the action-effect relationship encountered in the video (action movement/silence) and develop a secondary action-effect relationship reflective of their condition. This need to learn two unique action-effect relationships may have contributed to the longer learning duration. However, this interpretation does not fully explain why the FD and R groups did not learn at a significantly slower pace than the \emptyset group. A second interpretation suggests that the more predictable and consistent an action-effect relationship is (as with FO), the more negatively a learning curve may be impacted. Conversely, the more unpredictable and inconsistent an action-effect relationship is (as with R), the less a learning curve may be impacted. Most likely, these two interpretations are not mutually exclusive, and there is some contribution from both the observed action effect and quality of the later encountered action effect.

An important confound to consider may be auditory interference. Perhaps the addition of tones during the reproduction phase simply confused participants in the FO group, thereby interfering with their learning and generating a shallower learning curve. It is interesting to note, however, that this confound would most likely additionally impact timing scores—something our

results do not show. Also, simple auditory interference would most likely have similarly impacted the FD and R groups. Whereas their learning curves were, indeed, shallower than those of the \emptyset group, the differences did not reach significance.

EXPERIMENT 2

In order to determine whether the findings from Experiment 1 reflect observational action-effect transfer, practice/performance action-effect qualia, or auditory interference, we conducted a secondary study. In this experiment, we utilized the same observational action sequence video as that in Experiment 1, except this time we included the fixed, ordered, scale-ascending auditory piano tone effects in the video (FO observation, rather than silence). During the practice stages, one group heard tonal effects identical to the video (FO₂) whereas a second group heard no auditory effects (\emptyset_2). If the results of the previous experiment reflect observational action-effect transfer, we expected that the FO₂ group would show faster learning of the correct action timing. However, if the earlier results reflect an action-effect interaction between observation and practice/performance, then we should see no real difference between the FO₂ and \emptyset_2 groups, as neither group will encounter a secondary action-effect relationship. Finally, if the results of the previous experiment reflect auditory interference, then we expect that the FO₂ group will, again, learn slower than the \emptyset_2 group as there will be an identical level of interference to that in Experiment 1.

Method

Subjects

Twenty-four new subjects were enrolled (9 male, 15 female) and were randomly assigned to either the \emptyset_2 or the FO₂ group. All new subjects were musically naïve and had received no formal training in either a musical instrument or a vocal choir (self-reported prior to enrolment).

Apparatus

The apparatus was identical to that in Experiment 1 (see earlier).

Experimental conditions

Subjects were divided into two groups. During the action sequence practice stages, the first group heard no auditory effects (\emptyset_2 condition) while the second group heard fixed, ordered, scale-ascending piano tones with each key press (FO₂ condition—1:C, 2:D, 3:E, 4:F, and 5:G—Figure 6).

Design and procedure

With the exception of the tones generated during each button press, each subject underwent an identical protocol. Participants came to the lab only on two consecutive days for approximately 45 minutes each day. Each day consisted of alternating blocks

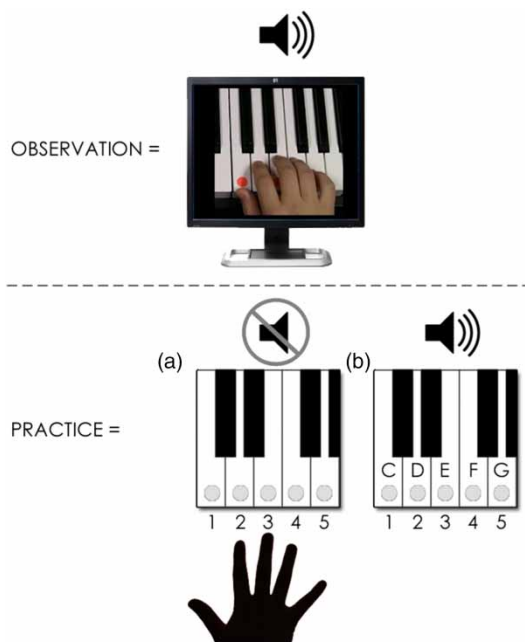


Figure 6. Experiment 2: Subjects observed a fixed-ordered auditory, first-person view of a musician playing an action sequence on the piano keyboard. When practising the sequence, subjects heard either: (a) no auditory effect (\emptyset_2), or (b) a fixed-ordered (FO₂) tonal scale. To view this figure in colour, please visit the online version of this Journal.

of observing, practising, and performing the action sequence (see first two days of Figure 2).

Video. The video was identical to the video utilized in Experiment 1 (see earlier) except, instead of no auditory effects, each key press generated a fixed, ordered, scale-ascending piano auditory tone effect (1:C, 2:D, 3:C, 4:D, and 5:E).

Sequence. The action sequence was approximately 15 s long and consisted of 29 key presses (see earlier). Key presses were separated by exactly 500 ms.

Practice. Following action observation, subjects were given time to practise the sequence from memory. During practice, each button press generated either no sound or auditory tones according to experimental condition (Figure 6). Subjects practised for a total of 30 min on each of the three days.

Performance. Following each practice block and at the start of Day 2, subjects were asked to perform the action sequence one time through as best they could. During this performance, the relevant tones continued to be produced, and the metronome remained on.

Data analysis

We utilized the same scoring scale and analysis as those in Experiment 1 (see earlier).

Results

Sequence scores (Figure 7a)

A one-way ANOVA between groups for the first recording (*T1*) and the final recording (*T17*) showed no significant differences, suggesting that each group began and ended the experiment at the same level of sequence competency [*T1*: $F(1, 26) = 1.423$, $p = .234$; *T17*: $F(1, 26) = 0.370$, $p = .548$]. A two-way ANOVA (Group \times Time) revealed a significant main effect of time, $F(10, 260) = 13.170$, $p < .001$, such that sequence scores significantly changed within groups across the study. This increase in scores can be understood as overall learning of the sequence. A significant main effect of group was also determined, $F(1, 26) = 8.281$, $p = .008$, such that the FO₂ group achieved higher scores than the \emptyset ₂ group, suggesting a learning differential between conditions. The interaction effect was nonsignificant, $F(10, 260) = 1.451$, $p = .158$.

To determine which group learned the sequence fastest, a Michaelis–Menten nonlinear curve was

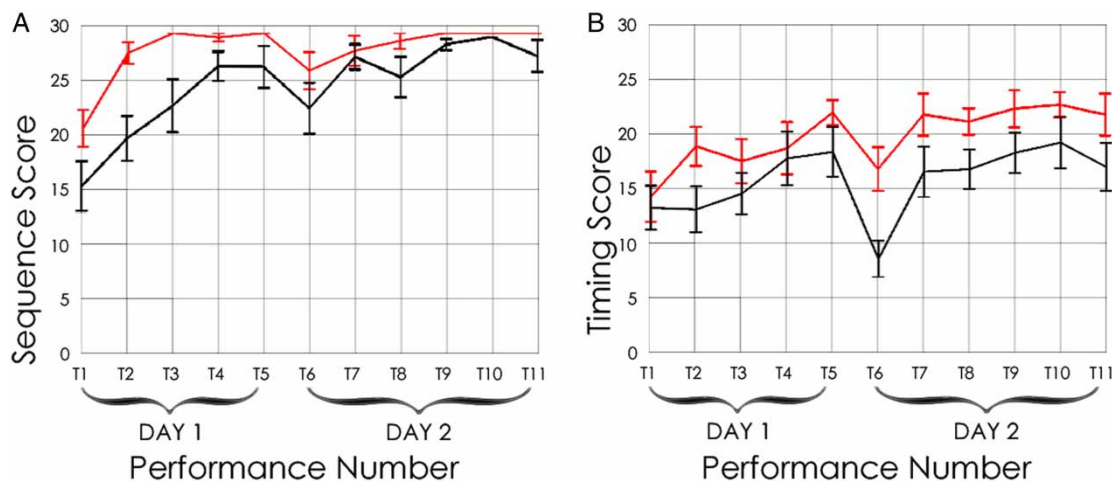


Figure 7. Experiment 2: (A) Average group sequence scores. (B) Average group timing scores. Black = \emptyset_2 (no auditory effect); red = FO₂ (fixed-ordered). Error bars represent ± 1 standard error. To view this figure in colour, please visit the online version of this Journal.

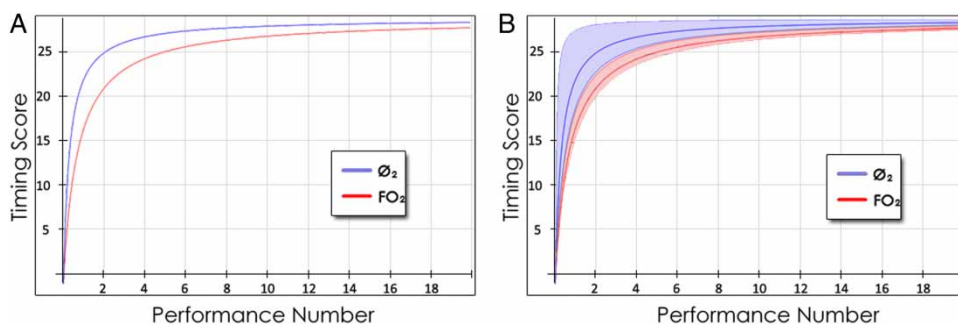


Figure 8. Experiment 2: Nonlinear regression curves for sequence scores; y axis = sequence score; x axis = performance number (time). (A) Mean sequence score curves. (B) Sequence score curves showing confidence interval (CI) range comparing \emptyset_2 (no auditory effect) and FO_2 (fixed-ordered). To view this figure in colour, please visit the online version of this Journal.

fitted to each condition [$B_0 \times x/(B_1 + x)$; B_0 = maximum score, B_1 = point of half maximum score]. B_0 was constrained at 29 (total possible score), and standard error estimates were bootstrapped (Figure 8). A look at Table 2 shows a 5% trimmed CI differential between \emptyset_2 and FO_2 . This suggests that the FO_2 learned the sequence at a faster rate than the \emptyset_2 group.

To better elucidate differences in sequence learning speed, Michaelis–Menten nonlinear regression curves were fitted to each individual (constraints as outlined above), and statistics were run between groups utilizing the B_1 value (the time point at which the learner reached half the maximum possible score). A one-way ANOVA revealed a significant difference in B_1 values, $F(1, 26) = 6.84$, $p = .015$ (Figure 9).

Finally, uncorrected within-subjects t -tests were run between T5 and T6 to explore sequence

retention between study days. No significant differences were found.

Timing scores (Figure 7b)

A one-way ANOVA between groups for the first recording (T1) and the final recording (T11) showed no significant differences, suggesting that each group began and ended the experiment at the same level of timing competency [T1: $F(1, 26) = 0.360$, $p = .554$; T11: $F(1, 26) = 2.167$, $p = .153$]. A two-way ANOVA (Group \times Time) revealed a significant main effect of time, $F(10, 260) = 9.786$, $p < .001$, such that timing scores

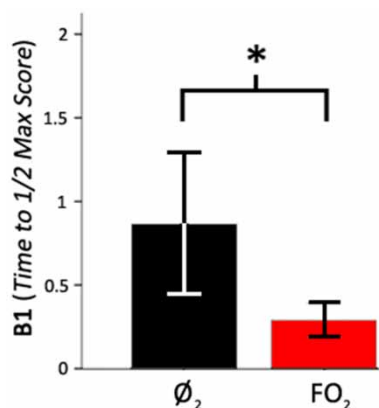


Figure 9. Experiment 2: Time needed to reach half maximum sequence score (14.5). Error bars represent ± 2 standard errors. \emptyset_2 = no auditory effect; FO_2 = fixed-ordered. To view this figure in colour, please visit the online version of this Journal.

Table 2. Nonlinear regression equations and statistics for \emptyset_2 and FO_2 sequence scores in Experiment 2

Group	Equation $B_0 \times x/(B_1 + x)$	B_1		R^2
		SE	5% trimmed CI interval	
\emptyset_2	$29x/(0.734 + x)$	0.068	0.558–0.869	.964
FO_2	$29x/(0.304 + x)$	0.102	0.045–0.515	.951

Note: \emptyset = no auditory effects; FO = fixed-ordered; CI = confidence interval; B_0 = maximum score, B_1 = point of half maximum score.

significantly changed within groups across the study. This increase in scores can be understood as overall learning of the timing. Neither the main effect of group, $F(1, 26) = 1.765, p = .196$, nor the interaction effect, $F(10, 260) = 0.929, p = .507$, was significant.

Finally, uncorrected within-subjects t -tests were run between T5 and T6 to explore timing retention between study days. Both groups displayed a significant difference in timing scores [\emptyset_2 : $t(13) = 3.526, p = .004$; FO $_2$: $t(13) = 3.269, p = .006$]. A between-subjects t -test showed no difference of the timing differential (T5 – T6) between groups, $t(26) = 1.119, p = .274$.

Discussion

Our data reveal that, similar to Experiment 1, the group whose motor reproduction action–effect condition matched that of the observed video (FO $_2$) learned the sequence at a significantly faster rate. Interestingly, both groups showed a significant decrease in timing scores between study days, thereby suggesting a lower timing retention than that seen in Experiment 1. Timing effects aside, these results appear to support the action–effect transfer hypothesis. If the earlier action–effect interaction or auditory interference hypotheses were correct, then we would have expected the \emptyset_2 group to perform equally to or better than the FO $_2$ group. This was not the case. To better tease out the relationship between observation and reproduction action–effect conditions, we decided to compare the \emptyset and FO groups from Experiment 1 with the \emptyset_2 and FO $_2$ groups from Experiment 2. As Experiment 2 was one day shorter than Experiment 1, we only utilized the first 11 performances from \emptyset and FO.

Combined analysis (groups \emptyset , FO, \emptyset_2 , and FO $_2$)

Sequence scores

A one-way ANOVA between groups for the first recording (T1) and the final recording (T11) showed no significant differences, suggesting that all groups began the experiment and ended the experiment at the same level of sequence

competency [T1: $F(3, 52) = 2.763, p = .051$; T11: $F(3, 52) = 0.733, p = .537$]. A two-way ANOVA (Group \times Time) revealed a significant main effect of time, $F(10, 520) = 13.286, p < .001$, such that sequence scores significantly changed within groups across the study. This increase in scores can be understood as overall learning of the sequence. A significant main effect of group was also determined, $F(3, 572) = 3.416, p = .024$, suggesting a learning differential between conditions. The interaction effect was nonsignificant, $F(30, 520) = 1.743, p = .170$. Post hoc analysis showed a significant difference between the FO and FO $_2$ groups (MD = 3.506, $p = .031$), the FO and \emptyset groups (MD = 4.701, $p = .004$), and the \emptyset_2 and FO $_2$ groups (MD = 2.487, $p = .013$).

To determine which group learned the sequence fastest, a Michaelis–Menten nonlinear curve was fitted to each condition [$B_0 \times x/(B_1 + x)$; B_0 = the plateau value, B_1 = curve midpoint]. B_0 was constrained at 29 (total possible score), and standard error estimates were bootstrapped (Figure 10a). To better elucidate differences in sequence learning speed, Michaelis–Menten nonlinear regression curves were fitted to each individual (constraints as outlined above), and statistics were run between groups utilizing the B_1 value. A one-way ANOVA revealed a significant difference in B_1 values, $F(3, 52) = 6.535, p = .001$. Post hoc analysis showed a significant difference between the FO and \emptyset groups (MD = 0.847, $p = .017$), and between the FO and FO $_2$ groups (MD = 1.151, $p = .001$), but no other significant differences (Figure 10b).

Timing scores

A two-way ANOVA (Group \times Time) revealed a significant main effect of time, $F(10, 520) = 9.786, p < .001$, such that timing scores significantly changed within groups across the study. This increase in scores can be understood as overall learning of the timing. Neither the main effect of group, $F(3, 52) = 2.631, p = .060$, nor the interaction effect, $F(30, 520) = 2.323, p = .086$, was significant.

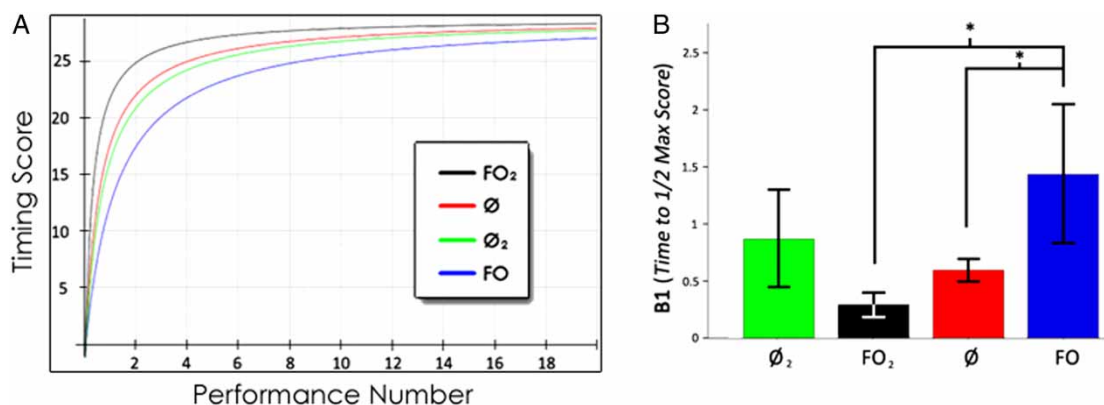


Figure 10. Combined experiments: (A) Average nonlinear regression curves for sequence scores; y axis = score; x axis = performance number (time). (B) Time needed to reach half the maximum score (14.5). Error bars represent ± 2 standard errors. To view this figure in colour, please visit the online version of this Journal.

GENERAL DISCUSSION

The studies undertaken have revealed some interesting features of action–effect transfer during observational learning. Although all subjects achieved comparable action sequence competency by the end of the study, the amount of time needed to achieve said competency differed across conditions in a manner that suggests that the action–effect relationship present during observation impacts sequence learning during reproduction. However, this impact appears to differ according to observational circumstance.

To begin, the \emptyset group from Experiment 1, which had an identical action–effect condition to that in the video, appeared to learn the sequence faster than the FO group. Interestingly, when auditory effects were added to the observational video, the FO₂ group appeared to learn the sequence faster than the \emptyset_2 group. Perhaps more interestingly, when compared across experiments, the groups that practised/performed utilizing the same action–effects as the observed video (\emptyset and FO₂) did not show any significant difference, nor did the two groups that practised/performed utilizing different action–effects from those in the observed video (\emptyset_2 and FO). These findings support the action–effect transfer during observational learning hypothesis, at least with regards to

action sequence learning (not action timing). Interestingly, the sequence scores for the \emptyset_2 group did not differ significantly from those for the \emptyset group. These findings suggest that *adding* an auditory effect during reproduction (when one is absent during observation) is somewhat detrimental to action sequence learning. However, *removing* an auditory effect from reproduction (when one is present during observation), although still detrimental to action sequence learning, appears to be less so.

However, additional findings from Experiment 1 appear to blur this simple correspondence somewhat. Rather than all unmatched action–effect relationships impacting sequence learning in the same way, it appears that the quality of the unique practice/performance action–effect is important. A look at the learning curves suggests that, with more power, it might be possible to obtain significant results showing a gradation between learning conditions, such that the less patterned and unpredictable the incongruent action–effects are (as in the FD and R groups), the less learning speed is negatively impacted. This idea is, of course, mere speculation at this point; however, this would certainly be an interesting concept to investigate in future research.

Interestingly, although to this point we have been discussing the \emptyset_2 group in light of an

action–effect being “removed” (from the observational stage), it is equally possible that the silence encountered during the reproduction stage was, in itself, a novel secondary action effect. Put differently, just as we argue that learning in the FO group probably suffered due to the introduction of a secondary, predictable action–effect relationship, so too may have the learning in the \emptyset_2 group suffered for a similar reason. The only difference might be that the added tones in the FO group required the relearning of five predictable action–effect relationships (each key mapped to C–G on the piano scale) whilst the removal of all tones in the \emptyset_2 group required the relearning of a single, predictable action–effect relationship (each key mapped to silence). Future experiments could tease this out by increasing the number of action–effect relationships utilized—say, by expanding the piano sequence to include 10 keys. With 10 keys, one would expect individuals in the FO condition to show a larger detriment than the one that we found (as they must now relearn 10 action–effect relationships) and members of the \emptyset_2 group to show a similar detriment to the one that we found (as they would still only have to learn one new action–effect relationship).

A likely contributing factor to our results is the transfer-appropriate processing (TAP) theory. First explored and elucidated in the 1970s (Craik & Lockhart 1972; Craik & Tulving, 1975; Morris, Bransford, & Franks, 1977), TAP proposes that memory retrieval is enhanced when undertaken in a manner identical to that experienced during encoding (Lee, 1988). Although typically discussed in the framework of future memory retrieval, integral to this theory is the assumption that all information that coincides with a task during the learning stage (e.g., action–effect congruence) will demonstrate common network encoding (Schendan & Kutas, 2007). The fact that, in this study, participants who practised and performed (memory retrieval) in a manner identical to the observed video (memory encoding) outperformed the other groups suggests that TAP was occurring. Interestingly, the majority of work exploring TAP to date has utilized active learning protocols (for review: Roediger, Gallo, & Geraci, 2002). Our

results suggest that, perhaps, TAP occurs in the observational sphere as well. More specifically, it is possible that memory retrieval is enhanced when it takes place in an environment/scenario identical to that encountered during observation. Future experiments can vary both the environment utilized in the observational video and the environment utilized for the practice/performance sections in order to determine how strong the TAP effect is in the observational learning domain.

The results obtained in these experiments hold many potential implications. First, with the proliferation of online classrooms and on-demand instructional videos, observational learning is quickly becoming a major tool for learners and teachers alike. Our data suggest that action–effect relationships are an important consideration, and, although we are uncertain how it may be leveraged to *enhance* learning, there are some things that online educators can certainly do to avoid hindering learning. Depending upon the type of lesson being conveyed, it would appear beneficial to include action effects that match those the learner will most likely encounter during practice. For instance, if one was designing a programme to aid stroke victims recover pinch strength in their fingers, the addition of unnecessary auditory cues or sound effects (typically not present when one makes a pinching motion) may lead to shallower learning curves than leaving the observed action silent. Similarly, if one is recording a video to teach novice athletes the correct form for a golf swing, the replacement of accurate action effects with narration may, again, hinder the speed with which the complex action is learned.

There are several limitations to this study that are worth noting. The first concerns sample size. As we recruited only 12–15 individuals per group, it is possible we may have missed several effects that would be present or overemphasized several effects that would diminish with more participants. Accordingly, our results must be interpreted with caution until similar research is undertaken. The next limitation concerns translatability. Although we have noted several potential applications of our findings, it is important to remember that the action learning task we utilized was relatively

simple. The dynamics of an action–effect relationship during observational learning of more complex behaviours and/or cognitive skills may present differently than here. Finally, as the standard error estimates in our regression analysis were bootstrapped, this carries the assumption that our data are correlated within each group but independent between each group. Though we have no obvious reason to question this assumption, all aforementioned results using this analysis method should be interpreted with caution.

Although interesting, this study is largely preliminary and generates many questions open to future investigation. For instance, how intimately must the action and resultant effect be linked in order to be coded together? Using our paradigm as an example, what if a training video contained a largely irrelevant but audible classical music piece in the background? Would the presence or absence of this classical track during practice affect learning and performance? Also, beyond action and auditory effect, what other sensory modalities are susceptible to common coding? Would flashing light effects elicit similar results? How about tastes or scents? Beyond simply determining that action–effects are transferred during observational learning, future research should focus on ways to leverage this transfer in a bid to increase and enhancing learning.

Original manuscript received 16 January 2014

Accepted revision received 11 January 2015

REFERENCES

- 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. *Neuron*, *42*, 323–334.
- Craik, F. I., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, *11*(6), 671–684.
- Craik, F. I., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294.
- Doody, S. G., Bird, A. M., & Ross, D. (1985). The effect of auditory and visual models on acquisition of a timing task. *Human Movement Science*, *4*, 271–281.
- Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: Relations between planning and temporal control. *Cognition*, *74*(1), 1–32.
- Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005). Action–effect coupling in pianists. *Research Psychologische Forschung*, *69*, 233–241.
- Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 229–240.
- Elsner, B., & Hommel, B. (2004). Contiguity and contingency in action–effect learning. *Psychological Research*, *68*, 138–154.
- Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of music*, *31*(3), 221–272.
- Gazzola, V., & Keysers, C. (2009). The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cerebral Cortex*, *19*(6), 1239–1255.
- Hallam, S. (1997). Approaches to instrumental music practice of experts and novices: Implication for education. In: H. Jørgensen & A. C. Lehmann (Eds.), *Does practice make perfect? Current theory and research on instrumental music practice* (pp. 89–107). Oslo: Norges musikkhøgskole.
- Herwig, A., & Waszak, F. (2009). Intention and attention in ideomotor learning. *The Quarterly Journal of Experimental Psychology*, *62*(2), 219–227.
- Hoffman, J., Sebal, A., & Stocker, C. (2001). Irrelevant response effects improve serial learning in serial reaction time tasks. *Journal of Experimental Psychology: Learning, memory, and Cognition*, *27*(2), 470–482.
- Kunde, W., Hoffmann, J., & Zellmann, P. (2002). The impact of anticipated action effects on action planning. *Acta Psychologica*, *109*, 137–155.
- Lee, T. D. (1988). Transfer-appropriate processing: A framework for conceptualizing practice effects in motor learning. *Advances in Psychology*, *50*, 201–215.
- Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior*, *16*(5), 519–533.
- Nelissen, K., Luppino, G., Vanduffel, W., Rizzolatti, G., & Orban, G. (2005). Observing others: Multiple action representation in the frontal lobe. *Science*, *310*, 332–336.

- Paulus, M., van Dam, W., Hunnius, S., Lindemann, O., & Bekkering, H. (2011). Action-effect binding by observational learning. *Psychonomic Bulletin & Review*, *18*(5), 1022–1028.
- Prinz, W. (1984). Modes of linkage between perception and action. In W. Prinz & A. F. Sanders (Eds.), *Cognition and Motor Processes* (pp. 185–193). Berlin: Springer.
- Prinz, W. (1990). *A common coding approach to perception and action*. Berlin: Springer.
- Prinz, W. (2010). Perception and action planning. *European Journal of Cognitive Psychology*, *9*(2), 129–154.
- Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*, 264–274.
- Roediger, H. L. III, Gallo, R., & Geraci, L. (2002). Processing approaches to cognition: The impetus from the levels-of-processing framework. *Memory*, *10*(5–6), 319–332.
- Schendan, H. E., & Kutas, M. (2007). Neurophysiological evidence for transfer appropriate processing of memory: processing versus feature similarity. *Psychonomic Bulletin & Review*, *14*(4), 612–619.