HAL
open science

Investigating three types of continuous auditory feedback in visuo-manual tracking

Eric Boyer, Frédéric Bevilacqua, Patrick Susini, Sylvain Hanneton

To cite this version:

Eric Boyer, Frédéric Bevilacqua, Patrick Susini, Sylvain Hanneton. Investigating three types of continuous auditory feedback in visuo-manual tracking. Experimental Brain Research, 2017, 235 (3), pp.691-701. 10.1007/s00221-016-4827-x . hal-01465822

HAL Id: hal-01465822

https://hal.science/hal-01465822v1

Submitted on 13 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Noname manuscript No.
(will be inserted by the editor)

Investigating three types of continuous auditory feedback in visuo-manual tracking

Éric O. Boyer · Frédéric Bevilacqua · Patrick Susini · Sylvain Hanneton

Received: date / Accepted: date

Abstract The use of continuous auditory feedback for motor control and learning is still under-studied and deserves more attention regarding fundamental mechanisms and applications. This paper presents the results of three experiments studying the contribution of task-, error-, and user-related sonification to visuo-manual tracking and assessing its benefits on sensorimotor learning. First results show that sonification can help decreasing the tracking error, as well as increasing the energy in participant's movement. In the second experiment, when alternating feedback presence, the user-related sonification did not show feedback dependency effects, contrary to the error and task-related feedback. In the third experiment, a reduced exposure of 50% diminished the positive effect of sonification on performance whereas the increase of the average energy with sound was still significant. In a retention test performed on the next day without auditory feedback, movement energy was still superior for the groups previously trained with the feedback. Although performance was not affected by sound, a learning effect was measurable in both sessions and the user-related group improved its performance also in the retention test. These results confirm that a continuous auditory feedback can be beneficial for movement training and also show an interesting effect of sonification on movement energy. User-related sonification can prevent feedback dependency and increase retention. Consequently, sonification of the user's own motion appears as a promising solution to support movement learning with interactive feedback.

Keywords Tracking · Auditory feedback · Sensorimotor learning · Sound · Interaction

Éric O. Boyer
Ircam - STMS CNRS UPMC - 1 pl. Igor Stravinsky, 75004 Paris
and
Laboratoire de Psychologie de la Perception UMR CNRS 8242, Université Paris Descartes - 45 rue des Saints Pères, 75006 Paris
Tel: +33144784884
Fax: +33144781545
E-mail: eric.boyer@ircam.fr

Frédéric Bevilacqua
Ircam - STMS CNRS UPMC - 1 pl. Igor Stravinsky, 75004 Paris

Patrick Susini
Ircam - STMS CNRS UPMC - 1 pl. Igor Stravinsky, 75004 Paris

Sylvain Hanneton
Laboratoire de Psychologie de la Perception UMR CNRS 8242, Université Paris Descartes - 45 rue des Saints Pères, 75006 Paris

1 Introduction

Continuous visuo-manual tracking tasks have been widely used in neuroscience research as an experimental paradigm to investigate human motor control and behavior (Craik, 1947; McRuer, 1980; Miall et al., 1993). During a tracking task, the participant has to pursue continuously a moving target that can follow a periodic (predictable) or noisy (non predictable) trajectory. This continuous task allows for well-controlled experiments with a limited number of parameters since generally only vision and proprioception are involved. Moreover, the movement regulation loop can be simply formalized by considering a single output command (the displacement of the hand on the interface) that responds and anticipates

to few sensory inputs (for instance the distance between the target and the pointer position). Studies reported that the tracking behavior can not be modeled by a linear input/output function (Hanneton et al., 1997). In other words, the movement of the participants contains frequencies that are not present in the target trajectory. Even if this task may be considered simple, this nonlinearity in the control makes difficult the development of satisfying models. The purpose of the present study is to investigate more particularly the action-perception coupling in a visuo-manual tracking task with various types of auditory feedback. Specifically, the effects of a real-time continuous auditory feedback on performance and learning are assessed.

Training and repetitions in a tracking task allow participants to improve their eye-hand coordination, the efficiency of sensory integration (vision and proprioception), tracking error interpretation and their ability to anticipate target motion. Changes in tracking gain, temporal lag or tracking error between target and pointer trajectories are measurable evidence of these improvements (Huang and Hwang, 2012). Visuo-manual coordination in tracking tasks is often considered to be controlled by a combination of feed-forward and feedback adaptive internal models (Neilson et al., 1988; Miall et al., 1993; Shams and Seitz, 2008). In this framework, training participates in the optimization of these internal models. Supplementary (multi-modal) sensory feedback is believed to support this process of optimization, and particularly the auditory modality due to its short time processing in the brain and a wide informational bandwidth (Robertson et al., 2009).

The use of augmented feedback (Hartveld and Hegarty, 1996) to enhance motor control has been largely studied over the past decades. Nevertheless, few studies have considered the auditory modality as augmented feedback. For example, Effenberg (Effenberg, 2004) proposed to use sound feedback to enhance movement perception. Kagerer and Contreras-Vidal describes particular phenomenons using auditory-motor transformations (Kagerer and Contreras-Vidal, 2009). They observed that the auditory-motor space can be affected by a newly formed internal model built consequently after a visuo-motor perturbation during a reaching task. Sigrist et al. (Sigrist et al., 2013) presented a review of augmented feedback techniques for motor learning and pointed out the high potential of auditory feedback. The relative simplicity and narrow range of published experiments and applications might be explained by technical issues and a lack of methodology regarding sound feedback design. Nevertheless, promising results showed that sound feedback, amongst others, can help learning specific movements (Sigrist et al., 2014; Boyer et al., 2014) or to control the orientation of an object (Rath and Schleicher, 2008). Overall, the underlying processes of the adaptation are rarely addressed in this field.

Providing extrinsic feedback while learning a motor task can lead to a potential dependency on the feedback (van Vliet and Wulf, 2006). This observation is sometimes referred to as the guidance hypothesis (Salmoni et al., 1984). The fact that the feedback is always present since the beginning of the learning can cause a decrease of performance when feedback is removed. In order to minimize the chance of this effect to appear, some authors suggested to reduce the rate or time of feedback presentation (Winstein and Schmidt, 1990; van Vliet and Wulf, 2006). But feedback dependency seems to be linked to the task and the context: guidance can develop during learning if the context is favorable to a spatial or visual representation of the action (Buchanan and Wang, 2012), suggesting that training with feedback always present is not necessarily detrimental. Very few studies examined this effect with auditory feedback though. Ronsse et al. (Ronsse et al., 2011) showed that sound feedback can be used to learn a bi-manual coordination pattern, and that auditory feedback did not lead to dependency on the feedback at the end of practice, unlike the use of a visual feedback (Avanzini et al., 2009).

In the context of robot-assisted rehabilitation, several studies showed that sound feedback is efficient to complement visual feedback. Rosati et al. reported studies on unidimensional tracking with a joystick, comparing the sonification of the target velocity and the tracking error (Rosati et al., 2012). They found that auditory feedback based on task parameters (features of the target or the setup) improved performance during an unpredictable task. Their results also showed that sonification of the tracking error did not have positive effect on performance and tended to deteriorate adaptation to a visual perturbation. Yet, this study did not include a control group and focused on input interface mapping.

Similarly to Rosati's study, we report here the comparison between three different sonification strategies in a visuo-manual tracking task. The three sonification strategies are based on the same sound synthesis system and exhibit identical acoustic features. As explained below, the difference resides only in the data that is being sonified. The first sonification is related to the instantaneous distance between the target and the pointer. It is *per se* an auditory augmentation of the available visual feedback and will be called 'error-related' feedback (*Error*). The second and third sonification strategies are related to instantaneous velocity signals: the 'task-related' sonification (*Target*) reflects the velocity of the

target, and the 'user-related' reflects the velocity of the manipulated pointer (*Pointer*). We emphasize that the *Pointer* sonification provides information only about the participants' movement, regardless of the task. To our knowledge, it is the first time that the effect of a continuous and user-related auditory feedback is tested in a tracking task.

What are the perceptual and neurophysiological mechanisms that could be behind the effectiveness of auditory feedback? Current findings in Psychology and Neurosciences can give insights about the co-processing of visual and auditory feedbacks by the CNS in perceptual and motor tasks (Effenberg et al., 2016). First, auditory feedback can be expected to have an influence at the perceptive level only: each different type of feedback may simply enhance the visual perception of the task through cross-modal processing. For instance, the target-related auditory feedback could increase the accuracy of the visual perception of the target. Indeed, audiovisual stimuli have been proved to be more efficient than unimodal stimuli in increasing the level of performance for stimulus detection (Vroomen and de Gelder, 2000; Frassinetti et al., 2002) or stimuli discrimination (Seitz et al., 2006; Giard and Peronnet, 1999). But the different types of feedback may also be processed differently by different neuronal populations in the sensorimotor system because it is fed with information of different nature. A feedback related to the velocity of the target conveys external information concerning the motion of an object in the environment. Visual motion and motion-related sounds could be processed together in the CNS. This is supported by the discovery of a cortical region in the posterior superior temporal sulcus (Bidet-Caulet et al., 2005) where motion-related sounds and visual (biological) motion seems to be co-integrated. The error-related feedback may participate in a regulation process able to minimize the velocity error during tracking. In the CNS, structures in the vicinity of the anterior cingulate cortex are involved in performance monitoring, particularly in detection of errors (Gehring and Fencsik, 2001). Online errors during eye movements to static or moving targets are coded in the superior colliculus (Orban de Xivry and Lefèvre, 2007) and in the posterior parietal cortex (Zhou et al., 2016). Neuroimaging revealed that activity in bilateral superior temporal cortex is related to auditory errors (real vs expected sound discrepancy) during speech production (Tourville et al., 2008). The pointer-related feedback conveys information about the motion of the participant and is not related to his level of performance. Exteroceptive by nature, it may act as a supplementary kinesthetic signal and not as an exteroceptive one. Auditory feedback has already been tested successfully as a substitute of kinesthetic vestibular signals in patients with vestibular loss (Dozza et al., 2005) and is thought to allow for kinesthetic and tactile exploration of objects (Boyer et al., 2015). Thus, we suggest that different neural circuitries could be involved in the processing of these different kinds of auditory feedback with various influences on motor control and learning. In this latter case, we should observe differences in the measured level of performance of participants reflecting the involvement of different neural processes for the different kinds of feedback.

In this study, we propose to evaluate the following hypotheses: a- real-time and continuous sonification, as an augmented feedback during a visuo-manual tracking task, can improve performance *i.e.* reduce the tracking error, b- the different types of feedback (related to the tracking error, the target or pointer velocity) would affect the level of performance in different ways, c- similarly, feedback dependency and learning retention may vary according to the type of feedback, and d- the presence of sonification can have an influence on the motion energy expended by the participants during tracking. The effect of the three sonifications were tested and compared in experiments 1 and 2. Experiment 2 allowed to test feedback dependency. A third experiment focused on the two velocity-related sonifications. In this experiment, the exposure to the feedback is reduced (i.e. not available all the time), and a 24-hours retention test was performed to observe learning stability. An additional group was added to assess the effect on the performance of a feedback that is not congruent with the task.

2 Materials and Methods

2.1 Participants

A total of one hundred and twenty participants volunteered for the study. They were aged from 18 to 70, with an average of 28.6 ± 10.2 years old. The gender balance was 41.7% female 58.3% male. Table 1 presents the detailed statistics about the participants in the different experiments. Every participant was involved in only one experiment and in one group exclusively. Exclusions criteria were: diagnosed hearing loss, physical impairment of the dominant arm and hand, color-blind condition or inability to distinguish the colored dots on the screen. All participants were healthy and had normal hearing. They were asked to rate their video games and sports practices on a five-point scale ranging from 1-never and 5-every day. This study was carried out in accordance with the Declaration of Helsinki and approved by the health research projects ethics commit-

tee of Paris Descartes University (International Review Board number 20142700001072). All participants gave written informed consent after reading the instructions. Participants in Experiment 3, which last over two days, received compensation.

Table 1 Age, gender, self-reported video games and sports practice for all the participants

	N	Age	Gender	Video games practice	Sports practice
Exp 1	36	30.8 ± 10.3	30.6% F	2.3 ± 1.3	2.6 ± 1.0
Exp 2	36	27.4 ± 9.6	44.4% F	1.9 ± 1.1	3.1 ± 1.2
Exp 3	36	29.3 ± 9.7	47.2% F	1.9 ± 1.1	2.3 ± 1.1
Control	12	26.4 ± 9.4	41.6% F	1.3 ± 0.9	2.6 ± 1.4
Total	120	28.6 ± 10.2	41.7% F	2.0 ± 1.2	2.6 ± 1.1

2.2 Experimental setup

The three experiments shared the same setup. Auditory feedback was delivered through headphones (AKG K271 MKII) that participants have in all the conditions. Experiments 1 and 2 were carried out in quiet offices at Ircam-Centre Pompidou and in the Sports Sciences department of Paris Descartes University, France. Experiment 3 was carried out in a double-walled sound-insulated booth at Ircam-Centre Pompidou.

In all the experiments, participants were seated in front of a desk with a graphic tablet on top (Wacom Intuos2, 304 x 228 mm, with XP-501E stylus). The visual environment was displayed on a Samsung SyncMaster 2053BW screen driven by an ATI Radeon HD2600XT 256 Mo graphic board. The computer used was a Mac Pro 2x2.8 GHz and 6 GB RAM running OSX 10.8.5. A program built under the Max environment (Cycling'74) allowed for the experiment control, real-time data processing and recording, as well as visual display and auditory feedback production. As the participants moved the stylus on the tablet surface, the (x, y) position data from the tablet translating the position of the cursor on the screen were analyzed and recorded at a sample rate of 100 Hz. The graphics were rendered using the *jit.jl* Max objects collection.

The overall latency of the system between the stylus moving on the tablet and the sound feedback generation was assessed by recording the contact sound of the stylus on the tablet surface with a microphone and collecting the time delays of the subsequent events in the processing chain. The total measure was 31 ms taking into account the audio driver latency.

2.2.1 Visual display

The target and pointer visuals were represented on the screen by respectively red and green 10.8 mm diameter plain circles on a black background. They were rendered at 60 Hz on 1680x1050 pixels (433.4 x 270.9 mm), the maximum resolution of the screen. The position of the target was generated with random numbers at 100 Hz and low-pass filtered to generate a reasonably smooth target trajectory, yet difficult enough to follow. Numbers are chosen between 0 and 1 with a step of 0.01, and then filtered using second-order recursive linear filters with 0.6 Hz cutoff frequency, unit normalized gain and q-factor. Three cascade filters were used to get a -36 dB/oct slope. Numbers are then scaled to fit the tablet space. In the end, the target exhibits an average velocity of 118 ± 65 mm.s^{-1}, ranging from 0 to 360 mm.s^{-1}. This target behavior is identical throughout the three experiments. For Experiment 3, three different trajectories have been pre-recorded corresponding to trials 1-2-3, so that the target trajectory of each respective trial is the same during both sessions (trial 1 of day 1 and trial 2 day 2, etc.).

2.2.2 Auditory feedback

The auditory feedback was generated in real-time using white noise filtered with a resonant filter (Max object *reson∼*) with a resonance factor of 23), which center frequency varied between 80 and 4000 Hz. This range is mapped to the minimum and maximum range of the varying parameter of each feedback type described below.

Three types of feedback were designed for the experiment sharing that same architecture. For the first one, the error-related feedback (*Error*), the filter frequency is modulated by the Euclidean distance between the target and the participant's pointer on the screen. The distance can be within 0-380 mm on the tablet and is mapped on a 80-4000 frequency range. The two others use the velocity of the target (*Target*) and the pointer (*Pointer*) to control the frequency of the filter. The same frequency range is mapped this time on 0-1200 mm.s^{-1} velocity values. For the incongruent feedback group, an incongruent sound was produced by modulating the same sound generator with a 0.6 Hz low-passed filtered noise. The resulting sound is then independent of the target motion and the movements of the participants but contains the same acoustical features as in the three feedback conditions.

2.3 Experimental procedure

Participants perform the task for a total time of 12 minutes in each experiment. In experiment 1 and 2, four trials of 3 minutes are performed, and three trials of 4 minutes are done in experiment 3. Participants took a 1 minute break between each trial. The experimental procedures in the different experiments tested the following feedback conditions: *NoAudio*, no auditory feedback; *Error*, sonification of the tracking error; *Target*, target velocity sonification; and *Pointer*, pointer velocity sonification.

Experiment 1 tested the three auditory feedback in different orders of presentation versus the *NoAudio* condition for 36 participants. In experiment 2, three groups of participants (one for each feedback) alternately received auditory feedback and the *NoAudio* condition. The additional incongruent feedback group (N=12) followed the same pattern but with the incongruent sound. Finally, in experiment 3, *Target* and *Pointer* conditions were used to train two groups of participants. On day 1 (*Training* session) the feedback was presented only 50% of the time during each trial of the task. A retention test is performed on day 2 (*Post* session) without feedback. In addition, a third group (control group) did the same two only in the the *NoAudio* condition. At the end of the session, the participants were asked to rate their perception of the task difficulty on a five-point scale from "very easy" to "very hard", and the auditory and visual fatigue (for experiment 1 and 2 only), from "none" to "maximum".

2.4 Data analysis

The recorded trajectories produced by the participants have been low-pass filtered at 8 Hz with a Gaussian filter before further analysis. The RMS tracking error (RMSE) was computed as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i}^{N} err(i)^2} \tag{1}$$

where $err(i)$ being the instantaneous Euclidean distance between the coordinates of the target (x_t, y_t) and the pointer (x_p, y_p) at time i:

$$err(i) = \sqrt{[x_p(i) - x_t(i)]^2 + [y_p(i) - y_t(i)]^2} \tag{2}$$

During the tracking task, participants expend energy to move the hand and follow the pointer. However, underlying neuronal processes involved in the control of movements face at least two constraints. The first general constraint is to minimize the energy expenditure. But, since the tracking of the target is not perfect, an additional amount of energy is necessary to produce corrective movements. Visuo-manual tracking has been shown to elicit an intermittent control of the hand movements, even in the case of very predictable target movements like periodic sinusoidal stimulations. It seems very difficult for humans to produce smooth pursuit hand movements to follow low-frequency target trajectories. Indeed, the hand trajectory exhibits rapid corrective movements.These saccadic-like movements contain higher frequencies (than the target trajectory) and they require energy expenditure. Consequently, the hand trajectory contains more energy than the energy required to perfectly follow the target. The ratio of expended energy over required energy could then be superior to one. We define the normalized energy in the movement as the amount of energy in the tangential velocity signal normalized by the same quantity for the target:

$$E = \frac{\sum(v_p(t))^2}{\sum(v_t(t))^2} \tag{3}$$

where $v_p(t)$ and $v_t(t)$ are the tangential velocity signal for the pointer and the target respectively.

The energy expenditure can be thought as to serve the performance of a participant if, for a set of trials, there is a significant negative correlation between the tracking error and the energy expenditure in the trials. In this case, the energy is expended to increase the level of performance of the participant. Instead, movement energy can be considered as wasted if no correlation is found, or a significant positive correlation between energy and tracking error.

2.5 Statistical analysis

Data from the first experiment were analyzed with ANOVAs on the RMSE error, the normalized energy E and the difficulty ratings as dependent variables. The following factors were used: condition (4 levels repeated measures factor - RMF) corresponding to the 4 different feedback conditions, and group (6 levels) testing the effect of order of presentation.

In the second experiment, a first analysis tested the effect of starting the experiment with or without sound (order factor). Analysis of variance is performed afterwards similarly on RMSE, E and difficulty ratings with the following factors: group (3 levels), sound during the trial (2 levels, on/off, RMF), and repetition of the feedback alternation (2 levels, RMF).

In the last experiment that investigates the link between feedback exposure and retention, data were ana-

lyzed through an ANOVA on RMSE and E values considering the following factors: session (2 levels RMF, a *Training* session followed the next day by the *Post* session), trials (3 levels RMF), and group (3 levels, Control, *Target* and *Pointer*).

For the incongruent feedback group, we searched for a significant effect of a sound on/off main factor in an ANOVA with two repeated measures factors, feedback on/off and repetition, and a single group main factor.

3 Results

3.1 Experiment 1

In experiment 1, the participants were asked to perform the task during four trials of 3 minutes. Each participant was exposed to the three feedback conditions: Error, Target, and Pointer. However, in order to test any possible order effect, each participant was assigned to one of the six different possible groups: E-T-P, E-P-T, P-E-T, P-T-E, T-E-P, T-P-E. All the groups began with a first trial in the NoAudio condition. Consequently, we defined a six-levels group factor, and a four-levels repeated measures condition factor (NoAudio, E, T, P).

The analysis of variance on the RMSE tracking error revealed no effect of the order of the sonification conditions (group factor), and no significant interaction between group and condition factors. The analysis showed a significant effect of condition on the tracking error ($F(3, 90) = 13.563, p < 0.005$), see left side of figure 1. Post-hoc comparisons showed that error decreases in the three sonification conditions compared to *NoAudio*were significant ($p < 10^{-5}$, Bonferroni post-hoc tests). The tracking error was reduced by 12.8%, 13.7% and 12.3% for *Error*, *Target* and *Pointer* conditions respectively. No significant difference was found between the three types of sonification regarding the tracking performance.

The amount of normalized energy in the movement E was superior to 1 for all conditions. The velocity signal exhibits more energy and more peaks than in the target signal (see figure 2). This example illustrates a typical behavior of target tracking. The pointer often overtakes and crosses the target trajectory, and probably with 'catch-up' saccades like at 7 seconds on the figure. The control of hand movements during tracking is not continuous, but intermittent (Hanneton et al., 1997). The trajectory of the hand is a combination of slow movements and saccadic-like faster movements. The presence of these saccadic-like movements explains the observation that there is more high-frequencies in the trajectory of the hand than in the trajectory of the target.

Feedback conditions had a significant effect on E ($F(3, 90) = 15.110, p < 10^{-5}$, right side of figure 1) and Bonferroni post-hoc tests confirmed that these conditions exhibited higher E values ($p < 10^{-4}$) than the *NoAudio* condition. Participant's movements were more energetic with sonification, nevertheless, no difference was found between the three sonification conditions either. As for RMSE, neither the main group factor nor group/condition interaction reached the significance level.

The ratings of the task difficulty were significantly affected by feedback conditions ($F(3, 90) = 19.490, p < 10^{-5}$), but neither by the group factor nor the interaction. The perceived difficulty was significantly reduced with sound 2.6 ± 0.9 (over 5) against 3.6 ± 1.0 ($p < 10^{-3}$ Bonferroni post-hoc test). No significant difference was found between the 3 types of feedback condition. Participants reported no particular auditory fatigue while hearing the auditory feedback (1.1 ± 0.4, 5 being maximum) and experienced medium visual fatigue (2.4 ± 1.2). A small number of participants reported hand fatigue or contraction by the end of the experiment and took advantage of the breaks between the trials to stretch and relax their hand.

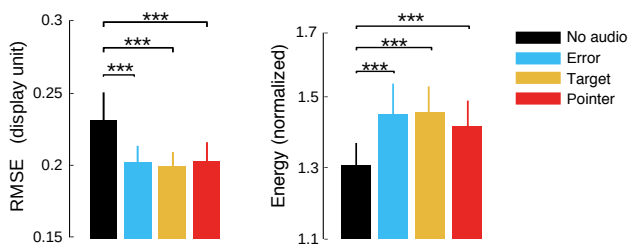

Fig. 1 Tracking error RMSE (left) and movement energy E (right) for each condition in experiment 1; the error bars indicate 95% confidence intervals; pairwise comparisons and significance levels are indicated above the bars; *** = $p < 10^{-3}$

3.2 Experiment 2

In the second experiment, the auditory conditions were tested separately with three different groups. Participants performed the task during 4 trials of 3 minutes, where sonification conditions, *Error* or *Target* or *Pointer* (12 participants each, mutually exclusive groups), alternated with the *NoAudio* condition. In each group half of the participants started with the *NoAudio* condition, half started with sound (order factor). Participants thus alternated twice between a sonification con-

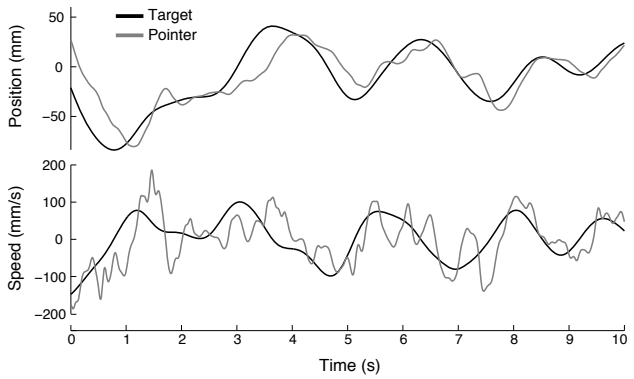

Fig. 2 Typical example of tracking motion (top) and velocity (bottom) along one dimension; the pointer (gray line) crosses the target trajectory (black) showing error corrections and higher energy.

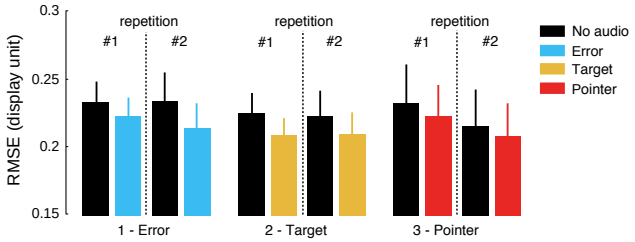

Fig. 3 Tracking error RMSE for the three feedback groups in experiment 2; the error bars indicate 95% confidence intervals.

dition and the *NoAudio* condition, which we call the repetition factor.

The order of presentation had no effect on the RMSE tracking error, and no significant interaction was found. Another ANOVA was carried out with sound (two levels), repetition (two levels) and group (three levels) factors, without the order factor.

The presence of sound during the trials had a significant effect on the RMSE error ($F(1,33) = 25.552, p < 10^{-4}$), and so did the repetition factor ($F(1,33) = 5.043, p < 0.05$). No interaction was significant. Post-hoc tests show that presence of sound significantly reduced the tracking error in the experiment ($p < 10^{-4}$ Bonferroni). However, as illustrated in figure 3, the *Pointer* feedback group seems to exhibit a specific behavior regarding the repetition effect. Whereas in the two other groups repetition seems to have no effect on RMSE (influence of sound remains the same), the RMSE seems to decrease during the last two trials. Consequently the interaction between group and repetition factors was tested with a contrast analysis where the *Error* and *Target* groups were opposed together to the *Pointer* group. The contrasted interaction between group and repetition was found significant ($F(1,33) = 4.745, p < 0.05$). Post-hoc analysis of the interaction showed that repetition improves the mean level of performance only in the *Pointer* group ($p < 0.025$ Bonferroni test).

The presence of the auditory feedback had a significant effect on the normalized energy E ($F(1,33) = 12.013, p < 0.002$). The energy was significantly increased by 7.3% on average with sound feedback. No other factor or interaction effect reached significance levels.

Concerning self-reports on perceived difficulty, participants generally reported that they felt more "comfortable" or "focused" with the sound but admitted being unable to tell if their performance actually improved. No factor had a significant effect on the difficulty ratings. Auditory and visual fatigue were rated similarly to Experiment 1 (2.6 ± 1.0 and 1.1 ± 0.5 respectively). The observations concerning hand fatigue are valid in this experiment too.

3.3 Experiment 3

Experiment 3 focused on the *Target* and *Pointer* sonification conditions, which are the two auditory feedback conditions based on velocity data. Participants performed the task during two sessions, a *Training* session on day 1, and a retention test (*Post* session) 24 hours later. Both sessions included 3 trials of 4 minutes. Contrary to the previous experiments, participants received auditory feedback 50% of the time of the trial during the training session: the sonification conditions alternated with the *NoAudio* condition every minute during a trial. On day 2, participants were asked to perform the same task without receiving auditory feedback. The participants were separated in three different groups corresponding to *Target* condition, *Pointer* condition and a control group that never received auditory feedback during the two sessions. Consequently, we defined a three-levels group factor (control, T, P), a three-levels repeated measures trial factor (trial 1, trial 2 and trial 3), and a two-level repeated measures session factor (day 1, day 2).

The ANOVA revealed that the trial factor (3 levels) had a significant effect on the RMSE tracking error ($F(2,66) = 67.314, p < 0.0001$). The first trial was always the less successful ($p < 0.0001$ for both sessions, Bonferroni post-hoc tests). The session factor (*Training* and *Post*) was also significant ($F(1,33) = 7.8895, p < 0.01$) with a decrease of the mean RMSE in the second session (see Figure 4). The trial*session interaction was significant ($F(2,66) = 7.106, p < 0.002$), certainly due to the smaller improvement of performance observed in the *Post* session. The group factor had no significant effect on RMSE, with no significant interaction neither.

However, during the *Post* session, the *Pointer* group is the only one that exhibited a significant improvement between the first and the last trials ($p < 0.0005$).

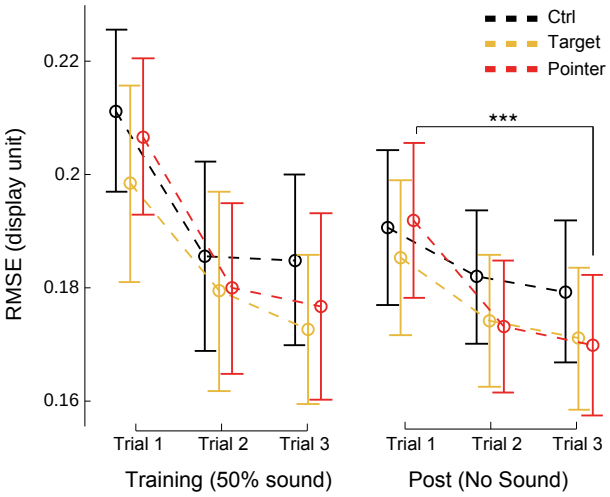

Fig. 4 Tracking errors RMSE for each trial in Experiment 3 during Training and Post sessions; error bars indicate 95% confidence intervals. In the Post session, only the *Pointer* sonification group significantly improved its level of performance ($p < 10^{-3}$).

Figure 5 shows the energy values and confidence intervals for experiment 3. Statistical analysis revealed a significant effect of the group factor on the energy ($F(2, 33) = 3.3966, p < 0.05$). The auditory feedback groups exhibited significantly more energy than the control group: +12.8% for *Target* and +15.9% for *Pointer* on average ($p < 0.05$ Bonferroni post-hoc tests). However, no significant difference was found between these two groups. The trial factor also had a significant effect on the energy values ($F(2, 66) = 8.8023, p < 0.005$), the last trial of each session appearing less energetic but in a non significant magnitude. No significant effect of the session factor was obtained, although there was no auditory feedback in the second session.

3.4 Incongruent feedback group

The ANOVA performed on the data from the incongruent group showed no significant effect of the main sound presence factor or the order factor. The RMSE was very similar with and without sound feedback: 0.24 (std=0.08) and 0.25 (std=0.08) respectively. We obtained a significant interaction between repetition and sound ($F(1, 33) = 6.256, p < 0.05$) but the post-hoc tests did not show any significant difference between the corresponding means.

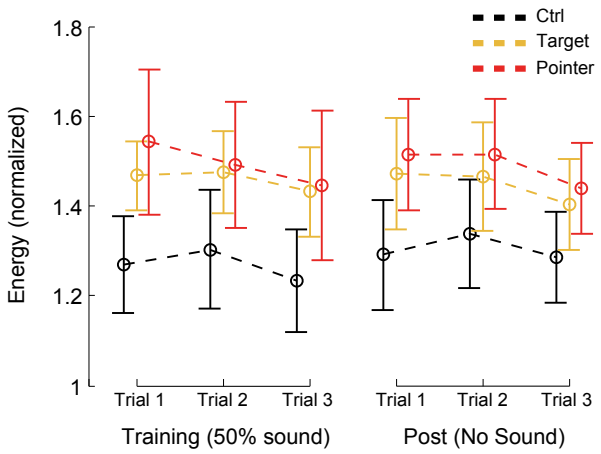

Fig. 5 Movement energy E for each trial in Experiment 3; error bars indicate 95% confidence intervals; the control group exhibits significantly less energy than the two velocity feedback groups ($p < 0.05$)

4 Discussion

4.1 Performance and learning

The first two experiments allowed us to test the first hypothesis. Taken together they show that the auditory feedback helped improving performance in the task, which confirms our first hypothesis. This result is in accordance with the feelings expressed by the participants that sound feedback can decrease the subjective difficulty of the task. However, the positive results of this experiment could be attributed to the prior training before exposure to auditory feedback. This is clarified in experiment 2 that showed positive benefits of the sound feedback are independent of the order of presentation, regardless of the type of data driving the sonification. Furthermore, additional results from the control group showed that an incongruent auditory feedback did not have a significant effect on the performance. This supports the idea that the benefits of the different sound feedback tested are not due other uncontrolled effects (like focus of attention for instance), but to the mapping between motion and sound in the auditory feedback.

The fact that *Target* as well as *Error* auditory feedback significantly improved performance differs from the observations by Rosati et al. (Rosati et al., 2012), where an error sonification did not prove to help for the tracking task. Our results also pointed out that the pointer-related (user-related) group improved its performance with repetition, contrary to the two other groups. Consequently, the user-related feedback group seems to be less sensitive to feedback removal trial after trial. As this auditory feedback is independent from both the target state and the performance, this result

addresses our second and third hypotheses: it suggests that in this context, the integration of an additional feedback does not require it to be task-oriented. Besides, sonification of the user's movement might have helped the participants to develop a learning process more robust to feedback removal.

The question of feedback dependency (guidance hypothesis (Buchanan and Wang, 2012; Ronsse et al., 2011)) is addressed in experiments 2 and 3. We can suppose that this effect explains the results observed for *Error* and *Target* groups in experiment 2. However, in experiment 3, when reducing feedback exposure (50% of the time in each trial), no benefit of sound on the level of performance is observed relatively to the no-sound control group. This shows that feedback dependency cannot be suppressed by a reduction of feedback exposure in this context.

Nevertheless, there is an indication that the *Pointer* feedback group might have a greater potential of improving its level of performance without sound during the retention test. This type of auditory feedback seems to stand out regarding the level of performance but also when observing the energy in the movement.

4.2 Movement energy

There is no straightforward relationship between the expenditure of energy and the level of performance of participants (tracking error). If saccadic-like corrective movements are accurate and anticipate the target motion, the supplementary energy is positively used to decrease the tracking error. On the contrary, a high amount of supplementary energy may be due to a poor pursuit, with many corrective but inaccurate movements. The values of the normalized energy, greater than unity for all the conditions, confirm the non-linear nature of the error regulation from the participants, who exhibited higher dynamics than the target trajectory (Hanneton et al., 1997). It is observable in experiments 2 and 3 that the energy contained in the motion is significantly increased with sonification, however, no significant difference was found between the three feedback tested. This result confirms our hypothesis that continuous sonification can modify movement features, here by increasing the average energy in the motion. A noticeable result is the fact that in experiment 3 the level of energy remains significantly higher in the retention test for the groups that received auditory feedback, even though the feedback is turned off during this session. This suggests that the effect of auditory feedback on motor control remains after feedback is removed.

In addition, in this experiment, the *Pointer* condition group tends to exhibit more energy than the *Target* group, both during training and retention test (although significance is not reached). The slope of the linear regression between energy and performance grows stronger for the *Pointer* group during the retention test. This can indicate that for this group the supplementary energy could be used to improve the level of performance as observed in the *Post* session. Although this cannot be fully addressed by our experiments, the results obtained here open important questions regarding the particular interest of this movement-related auditory feedback. The energy-related feature is a measure that is independent on the level of performance and that can serve to compare the effect of different types of augmented sensory feedback.

4.3 Towards an auditory-driven proprioception?

The sonification of the pointer provides the users with information on their hand kinematics while performing the task, which differs from the two other types of feedback tested here. This information could be interpreted by the sensory system as an augmentation of proprioception, more specifically of kinesthesia. The fact that participants interact with a tangible object on the tablet (the stylus) emphasizes the physical nature of the pointer feedback and the coherence with motion. The sensorimotor system could thus benefit from this richer sensory input, as it complements visual input, which is already available. As a result, the sensory feedback prediction (forward model) could be faster and more accurate (Miall and Wolpert, 1996; Wolpert and Ghahramani, 2000). Effects of this mechanism have been previously observed: auditory feedback can be used to enhance the movement consciousness (Vogt et al., 2009; Schmitz and Effenberg, 2012). The high energy measured in the participants' movement with the pointer feedback could also be the manifestation of such a mechanism.

Recent studies examined the link between proprioception and motor learning. Rosenkranz and Rothwell (Rosenkranz and Rothwell, 2012) showed that integration and modulation of proprioceptive input induced a positive effect on motor learning. Wong et al. (Wong et al., 2012) showed that a proprioceptive training could reduce position error and increase speed, leading to a larger learning in arm movement. These recent studies on sensory training support the theory of perceptual learning (Darainy et al., 2013), which describes sensorimotor learning entrained by perceptual changes. These results are in favor of our hypothesis, but further investigations are needed to address this particular question with continuous movement sonification. The high temporal granularity of auditory signals compared to vision

makes interactive movement sonification a good candidate to supplement proprioception.

4.4 Individual differences

None of the participants were excluded from the analysis based on performance criteria. However, we observed large inter-individual differences in the level of performance of participants, particularly in their response to the presence of auditory feedback. Some participants may be significantly less responsive to auditory-motor relationship, especially in this case where the attention is mainly focused by the visuo-manual regulation. We previously observed such heterogeneity in the case of a closer auditory-motor relationship, under the paradigm of *sound-oriented task* (Boyer et al., 2014). We believe that the diversity in participants responses to sonic interaction is not addressed enough in the related literature. Modifying the sonification mapping through the experiment and according to the participant baseline performance could be considered - machine learning algorithms for multimodal mappings can offer this solution (Françoise et al., 2014).

4.5 Perspectives

The task used in the present study does not offer a large room for learning since participants exhibited a decent initial level of performance. Our results can be magnified introducing, for instance, a visuo-motor disturbance between hand and pointer motion. Auditory feedback could then enhance the adaptation rate of participants to the disturbance. The analogous influence of auditory feedback has also to be confirmed in other tasks also used to study sensorimotor learning - for instance reaching tasks with prismatic goggles or workspace rotation paradigms (Krakauer and Mazzoni, 2011).

Varying the spectral components of the target trajectory can also be a way to evaluate whether the phenomenon observed here with auditory feedback operate at similar or different spectral ranges of motion. Adding or cutting higher frequencies in the target motion would make the trajectory respectively more or less difficult to follow, and would allow to observe frequency-dependent correction mechanisms. The integration of the auditory feedback may also vary with the subsequent difficulty of the task.

4.6 Conclusive remarks

The presence of continuous auditory feedback in a two-dimensional tracking task proved to have a significant effect on performance and learning. Error-related, as well as target and user-related sonifications, can improve both tracking performance and affect movement energy. Our observations show that although error sonification can help improving level of performance, auditory feedback design for motor control should consider providing user-related feedback on the performed action. We suggest the interest of using velocity-related modalities for continuous sonification, especially because of its potential link to energy in motion. Although many questions remain open, we argue that continuous sonification of movement features should be further employed, especially considering interactive scenarios where increasing energy is required. Physical rehabilitation, where the engagement of participants is crucial and sensory feedback often impaired can be a promising application.

4.7 Ethical approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

4.8 Acknowledgment

This work has been funded by ANR French National Research Agency, under the ANR-Blanc program 2011 (LEGOS project ANR-11-BS02-012) and additional support from Cap Digital.

References

[Avanzini et al., 2009]Avanzini, F., De Götzen, A., Spagnol, S., and Rodà, A. (2009). Integrating auditory feedback in motor rehabilitation systems. In *Proceedings of International Conference on Multimodal Interfaces for Skills Transfer (SKILLS09)*.

[Bidet-Caulet et al., 2005]Bidet-Caulet, A., Voisin, J., Bertrand, O., and Fonlupt, P. (2005). Listening to a walking human activates the temporal biological motion area. *Neuroimage*, 28:132–39.

[Boyer et al., 2014]Boyer, E. O., Pyanet, Q., Hanneton, S., and Bevilacqua, F. (2014). Learning Movement Kinematics with a Targeted Sound. In *Sound,*

Music & Motion, Lecture Notes in Computer Science Volume 8905, pages 218–233. Springer International Publishing.

[Boyer et al., 2015]Boyer, E. O., Vandervoorde, L., Bevilacqua, F., and Hanneton, S. (2015). Touching sounds: audio virtual surfaces. In *2015 IEEE 2nd VR Workshop on Sonic Interactions for Virtual Environments (SIVE)*, pages 1–5. IEEE.

[Buchanan and Wang, 2012]Buchanan, J. J. and Wang, C. (2012). Overcoming the guidance effect in motor skill learning: feedback all the time can be beneficial. *Experimental Brain Research*, 219(2):305–20.

[Craik, 1947]Craik, K. (1947). Theory of the human operator in control systems. I. The operator as an engineering system. *British Journal of Psychology*, 38(2):56–61.

[Darainy et al., 2013]Darainy, M., Vahdat, S., and Ostry, D. J. (2013). Perceptual learning in sensorimotor adaptation. *Journal of neurophysiology*, 110(9):2152–62.

[Dozza et al., 2005]Dozza, M., Chiari, L., and Horak, F. B. (2005). Audio-biofeedback improves balance in patients with bilateral vestibular loss. *Archives of Physical Medicine and Rehabilitation*, 86:1401–1403.

[Effenberg, 2004]Effenberg, A. O. (2004). Using Sonification to Enhance Perception and Reproduction Accuracy of Human Movement Patterns. In *Proceedings of the International Workshop on Interactive Sonification*, Bielefeld.

[Effenberg et al., 2016]Effenberg, A. O., Fehse, U., Schmitz, G., Krueger, B., and Mechling, H. (2016). Movement sonification: Effects on motor learning beyond rhythmic adjustments. *Frontiers in Neuroscience*, 10(219).

[Françoise et al., 2014]Françoise, J., Schnell, N., Borghesi, R., and Bevilacqua, F. (2014). Probabilistic models for designing motion and sound relationships. In *Proceedings of the 2014 International Conference on New Interfaces for Musical Expression*, pages 287–292.

[Frassinetti et al., 2002]Frassinetti, F., Bolognini, N., , and Làdavas, E. (2002). Enhancement of visual perception by crossmodal visuo-auditory interaction. *Experimental Brain Research*, 147:332–343.

[Gehring and Fencsik, 2001]Gehring, W. J. and Fencsik, D. E. (2001). Functions of the medial frontal cortex in the processing of conflict and errors. *Journal of Neuroscience*, 21(23):9430–9437.

[Giard and Peronnet, 1999]Giard, M. and Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: a behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, 11:473–490.

[Hanneton et al., 1997]Hanneton, S., Berthoz, A., Droulez, J., and Slotine, J. (1997). Does the brain use sliding variables for the control of movements? *Biological Cybernetics*, 77(6):381–393.

[Hartveld and Hegarty, 1996]Hartveld, A. and Hegarty, J. (1996). Augmented Feedback and Physiotherapy Practice. *Physiotherapy*, 82(8):480–490.

[Huang and Hwang, 2012]Huang, C.-T. and Hwang, I.-S. (2012). Eye-hand synergy and intermittent behaviors during target-directed tracking with visual and non-visual information. *PloS one*, 7(12):e51417.

[Kagerer and Contreras-Vidal, 2009]Kagerer, F. A. and Contreras-Vidal, J. L. (2009). Adaptation of sound localization induced by rotated visual feedback in reaching movements. *Experimental Brain Research*, 193(2):315–21.

[Krakauer and Mazzoni, 2011]Krakauer, J. W. and Mazzoni, P. (2011). Human sensorimotor learning: adaptation, skill, and beyond. *Current opinion in neurobiology*, 21(4):636–44.

[McRuer, 1980]McRuer, D. (1980). Human dynamics in man-machine systems. *Automatica*, 16:237–253.

[Miall et al., 1993]Miall, R. C., Weir, D. J., and Stein, J. F. (1993). Intermittency in human manual tracking tasks. *Journal of motor behavior*, 25(1):53–63.

[Miall and Wolpert, 1996]Miall, R. C. and Wolpert, D. M. (1996). Forward Models for physiological motor control. *Neural networks*, 9(8):1265–1279.

[Neilson et al., 1988]Neilson, P., Neilson, M., and O'Dwyer, N. (1988). Internal models and intermittency: a theoretical account of human tracking behavior. *Biological Cybernetics*, 58:101–112.

[Orban de Xivry and Lefèvre, 2007]Orban de Xivry, J.-J. and Lefèvre, P. (2007). Saccades and pursuit: two outcomes of a single sensorimotor process. *The Journal of Physiology*, 548(1):11–23.

[Rath and Schleicher, 2008]Rath, M. and Schleicher, R. (2008). On the Relevance of Auditory Feedback for Quality of Control in a Balancing Task. *Acta Acustica united with Acustica*, 94:12–20.

[Robertson et al., 2009]Robertson, J. V. G., Hoellinger, T., Lindberg, P., Bensmail, D., Hanneton, S., and Roby-Brami, A. (2009). Effect of auditory feedback differs according to side of hemiparesis: a comparative pilot study. *Journal of neuroengineering and rehabilitation*, 6:45.

[Ronsse et al., 2011]Ronsse, R., Puttemans, V., Coxon, J. P., Goble, D. J., Wagemans, J., Wenderoth, N., and Swinnen, S. P. (2011). Motor learning with augmented feedback: modality-dependent behavioral and neural consequences. *Cerebral cortex*

(New York, N.Y. : 1991), 21(6):1283–94.

[Rosati et al., 2012]Rosati, G., Oscari, F., Spagnol, S., Avanzini, F., and Masiero, S. (2012). Effect of task-related continuous auditory feedback during learning of tracking motion exercises. *Journal of neuroengineering and rehabilitation*, 9(1):79.

[Rosenkranz and Rothwell, 2012]Rosenkranz, K. and Rothwell, J. C. (2012). Modulation of proprioceptive integration in the motor cortex shapes human motor learning. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 32(26):9000–6.

[Salmoni et al., 1984]Salmoni, A. W., Schmidt, R. A., and Walter, C. B. (1984). Knowledge of Results and Motor Learning : A Review and Critical Reappraisal. *Psychological Bulletin*, 95(3):355–386.

[Schmitz and Effenberg, 2012]Schmitz, G. and Effenberg, A. O. (2012). Perceptual effects of auditory information about own and other movements. *Proceedings of the 18th International Conference on Auditory Display*, pages 89–94.

[Seitz et al., 2006]Seitz, A. R., Kim, R., and Shams, L. (2006). Sound facilitates visual learning. *Current biology : CB*, 16(14):1422–7.

[Shams and Seitz, 2008]Shams, L. and Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in cognitive sciences*, 12(11):411–7.

[Sigrist et al., 2014]Sigrist, R., Rauter, G., Marchal-Crespo, L., Riener, R., and Wolf, P. (2014). Sonification and haptic feedback in addition to visual feedback enhances complex motor task learning. *Experimental brain research.*

[Sigrist et al., 2013]Sigrist, R., Rauter, G., Riener, R., and Wolf, P. (2013). Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. *Psychonomic bulletin & review*, 20(1):21–53.

[Tourville et al., 2008]Tourville, J., Reilly, K., and Guenther, F. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage*, 39(3):1429–43.

[van Vliet and Wulf, 2006]van Vliet, P. M. and Wulf, G. (2006). Extrinsic feedback for motor learning after stroke: what is the evidence? *Disability and rehabilitation*, 28(13-14):831–40.

[Vogt et al., 2009]Vogt, K., Pirro, D., Kobenz, I., Höldrich, R., and Eckel, G. (2009). Physiosonic - Movement sonification as auditory feedback. In *6th International Symposium, CMMR/ICAD 2009*, pages 1–7, Copenhagen.

[Vroomen and de Gelder, 2000]Vroomen, J. and de Gelder, B. (2000). Sound enhances visual perception : cross-modal effects of auditory organization on vision. *Journal of Experimental Psychology:Human Perception & Performance*, 26:1583–1590.

[Winstein and Schmidt, 1990]Winstein, C. J. and Schmidt, R. A. (1990). Reduced frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(4):677–691.

[Wolpert and Ghahramani, 2000]Wolpert, D. M. and Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature neuroscience*, 3 Suppl(november):1212–1217.

[Wong et al., 2012]Wong, J. D., Kistemaker, D. a., Chin, A., and Gribble, P. L. (2012). Can proprioceptive training improve motor learning? *Journal of neurophysiology*, 108(12):3313–21.

[Zhou et al., 2016]Zhou, Y., Liu, Y., Lu, H., Wu, S., and M., Z. (2016). Neuronal representation of saccadic error in macaque posterior parietal cortex (ppc). *eLIFE*, 20(5).