Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

# Motor resonance in observed digital musical interactions: a time-resolved grip force study

**Olivier Capra**

olivier.capra@univ-catholille.fr

Université Catholique de Lille

**Florent Berthaut**

Centre de Recherche en Informatique, Signal et Automatique de Lille

**Research Article**

**Keywords:** motor resonance, action observation, grip force, digital interactions, digital musical instruments, motor simulation, Bayesian time-series analysis

**Posted Date:** June 22nd, 2026

**DOI:** https://doi.org/10.21203/rs.3.rs-9995159/v1

**License:** This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

**Additional Declarations:** Competing interest reported. O. Capra serves as a co-editor of the special issue in which this manuscript is submitted. F. Berthaut declares no competing interests.

# Motor resonance in observed digital musical interactions: a time-resolved grip force study

Olivier Capra[1*] and Florent Berthaut[2]

[1*]Neurethic-Lab – ETHICS EA 7446, Université Catholique de Lille, Lille, F-59000, France.
[2]CRIStAL, CNRS, Univ. Lille, Lille, F-59000, France.

*Corresponding author(s). E-mail(s): olivier.capra@univ-catholille.fr;
Contributing authors: florent.berthaut@univ-lille.fr;

**Abstract**

Motor resonance during action observation is well documented. However, its dynamics during the observation of digital interactions, in which gestures and outcomes are not necessarily constrained by direct physical causality, remain largely unexplored. This question is particularly salient in performances with digital musical instruments, where the audience relies on tight coherence between gesture and sound.

In this study, 42 participants observed a musician triggering a sound modulation on a digital musical instrument. The gesture was either direct or preceded by a larger ancillary movement that had no effect on the sound modulation. Grip force microvariations were recorded continuously and analyzed using a time-resolved Bayesian approach.

Grip force dynamics were temporally structured and aligned with key phases of the interaction. The larger ancillary gesture produced a local increase and reshaped the temporal profile of the signal, but did not yield a reliably larger peak response. Moreover, when the signal was realigned to the triggering event, the response to the same contact event was smaller after the ancillary gesture.

These findings indicate that motor resonance during digital interaction observation follows the temporal structure of the action and is shaped by gestural context rather than movement amplitude alone. More broadly, they support continuous grip force as a time-resolved index of motor resonance and position digital musical interactions as a promising framework for studying action–outcome relations under controlled yet ecologically meaningful conditions.

**Keywords:** motor resonance, action observation, grip force, digital interactions, digital musical instruments, motor simulation, Bayesian time-series analysis

## Public Significance Statement

Digital interactions challenge a brain that evolved in a physical world, where gestures typically have direct and predictable outcomes. Today, a single gesture on a digital device can carry many meanings. By measuring tiny involuntary changes in grip force while participants watched digital musical interactions, this research clarifies how motor engagement unfolds during the observation of digitally mediated actions.

## Introduction

Motor resonance during action observation is well documented. However, its dynamics during the observation of digital interactions, in which an action is performed through a digital device, remain largely unexplored. In these interactions, gestures and their consequences are no longer necessarily constrained by physical causality. The same gesture can produce different outcomes, sometimes unrelated in nature and scale, and the same outcome can result from different gestures. This tension is particularly acute in live performances involving digital musical instruments, where the spectator experience relies on tight multimodal coherence between gesture and sound despite the absence of a physically constrained gesture–outcome link.

This decoupling raises a fundamental question for motor simulation: when the amplitude of an observed gesture is decoupled from its outcome, to what extent does motor engagement follow movement amplitude, and to what extent is it shaped by the temporal and contextual structure of the interaction?

To address this question, we first review evidence on motor resonance and the many factors that modulate it. We then consider how digital interactions decouple gesture from outcome, and why digital musical interactions provide a controlled yet naturalistic context for observation. Finally, in the present study, we examine this issue using continuous grip force recordings, which allow us to monitor motor engagement over time, and we analyze its temporal dynamics using a time-resolved Bayesian approach.

### Motor resonance during action observation

Extensive convergent data support the existence of motor resonance during action observation (Craighero, 2024; Di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Fadiga, Fogassi, Pavesi, & Rizzolatti, 1995; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996). Often interpreted as a manifestation of the mirror neuron system (Rizzolatti & Craighero, 2004), it is proposed to arise from a motor simulation mechanism whereby observed actions are mapped onto the observer's own motor system (Barsalou, 2009; Jeannerod, 2001). Following the taxonomy proposed by Stockner, Mazzoni, Perrin, and Ianì (2025), we use *motor resonance* throughout this work to refer to the measurable engagement of the observer's motor system during action observation (*the result*) and *motor simulation* to refer to the underlying mechanism proposed to generate it (*the process*).

Early studies operationalized motor resonance through measures of corticospinal excitability, showing increased excitability during action observation indexed by motor

evoked potentials (MEPs) elicited with transcranial magnetic stimulation (TMS) (Fadiga et al., 1995).

Since then, research has expanded considerably, identifying multiple factors influencing motor resonance (for reviews, see Kemmerer, 2021). In the context of the present study, we focus on four that are particularly relevant: the physical properties of the observed gesture, the attentional engagement of the observer, the temporal dynamics of motor resonance, and the integration of contextual cues.

Gangitano, Mottaghy, and Pascual-Leone (2001) showed that the amplitude of the observed action modulates MEP magnitude, indicating that motor resonance is sensitive to the physical properties of the observed movement. Beyond this general variation with amplitude, motor resonance also shows effector specificity; greater MEP amplitudes are observed when the recorded muscle corresponds to the effector involved (Cavallo, Becchio, Sartori, Bucchioni, & Castiello, 2012; Fadiga et al., 1995).

Motor resonance is also sensitive to attentional state. Reduced attention to the observed action attenuates this response, indicating that it is not fully automatic but depends on the observer's engagement with the observed agent (Baldissera, Cavallari, Craighero, & Fadiga, 2001; Puglisi et al., 2017).

Motor resonance tracks the temporal unfolding of observed actions (Nishitani & Hari, 2000). Corticospinal excitability shows an early, non-muscle-specific facilitation around 90 ms after action onset, followed by a later muscle-specific phase around 200 ms that closely reflects the unfolding muscle activity of the observed agent (Lepage, Tremblay, & Théoret, 2010; Naish, Houston-Price, Bremner, & Holmes, 2014). Strikingly, motor resonance can even anticipate the observed action, with peaks preceding the corresponding movement of the agent (Borroni, Montagna, Cerri, & Baldissera, 2005), pointing to a predictive rather than strictly reactive simulation mechanism.

Finally, motor resonance is modulated by contextual information beyond the observed kinematics. At later latencies, around 400 ms, modulations reflect the contextual congruence of the observed action (Amoruso, Finisguerra, & Urgesi, 2016). The work of Alaerts, Swinnen, and Wenderoth (2010) further supports the influence of inferred contextual cues by reporting greater MEPs when an observed grasping action involved heavier objects, a modulation that cannot be explained by direct kinematic matching alone, since object weight is not directly visible but must be inferred from contextual cues.

Together, these findings indicate that motor resonance is not driven by direct kinematic matching alone, but is shaped by multiple factors, including attentional engagement, temporal dynamics, and the integration of contextual cues.

## Digital musical interactions and the decoupling of gesture and outcome

Unlike physical interactions, in which an agent's gestures and their consequences are mechanically linked, digital interactions mediate gesture and outcome through a device. Depending on its configuration, the same gesture can produce different consequences, sometimes unrelated in nature and scale, and the same outcome can result

from different gestures. For an observer, this decoupling between action and outcome raises a challenge for how motor resonance is organized.

Very few studies have specifically examined motor resonance in observers of digital interactions. Most related work has focused on virtual reality contexts, often aimed at motor imagery training (e.g., Choi, Kim, Huh, & Jo, 2020). Decloe and Obhi (2013) showed that motor cortical processing during observation of typing actions on a cell phone was causally involved in subsequent object recognition, suggesting that motor simulation accompanies observation of digital device use. Other work has leveraged digital and virtual environments primarily for therapeutic purposes, using action observation within brain-computer interfaces (BCI) feedback loops to activate sensorimotor circuits (Syrov, Yakovlev, Miroshnikov, & Kaplan, 2023). More directly relevant to the present study, Matamala-Gomez et al. (2025) reported greater corticospinal excitability when participants observed first-person virtual hand movements connected to a virtual upper limb compared to an isolated hand, supporting the idea that the visual coherence of the observed agent modulates motor simulation. This also illustrates how digital setups allow such coherence to be systematically manipulated.

Digital musical instruments instantiate a gesture–outcome decoupling within the rich context of musical performance. Observers bring strong prior expectations regarding gesture–sound mappings (Huron, 2008; Jensenius & Wanderley, 2010), rooted in the sensitivity of the motor system to action–sound contingencies (Kohler et al., 2002; Zatorre, Chen, & Penhune, 2007). Because gesture–sound relations are digitally mediated, these expected mappings can be fulfilled, violated, or rendered ambiguous.

Behavioral studies show that such ambiguity leads observers to underestimate the musician's contribution (Capra, 2020; Capra, Berthaut, & Grisoni, 2020), attributing greater agency to the digital system. Conversely, enhancing the visibility of this relationship can restore the attributed agency to the musician. These findings indicate that observers actively interpret the structure of the interaction, rather than passively mapping observed kinematics to outcomes.

Musical performance has been proposed as an ideal framework for studying motor cognition, offering a productive balance between experimental control and ecological validity (D'Ausilio, Novembre, Fadiga, & Keller, 2015). Digital musical instruments extend this framework further by allowing gesture and outcome to be manipulated independently while preserving the rich expectation structure that characterizes musical experience. They also provide fine temporal control over the interaction: sensor events can be timestamped, and the sound-generation process can be designed, tracked, and replicated, allowing auditory salience and gesture–sound timing to be controlled experimentally.

Digital musical interactions may thus offer a particularly useful setting to study motor resonance during action observation. They preserve an ecologically meaningful gesture–sound relation while allowing gesture kinematics, auditory outcome, and temporal structure to be experimentally controlled and selectively dissociated. This combination is difficult to obtain in standard action-observation paradigms, where movement kinematics, inferred effort, and action outcome are usually physically coupled.

## Grip force microvariations as a continuous measure of motor engagement

A growing body of research supports the use of grip force as a sensitive readout of motor involvement in perceptual and cognitive processes, from action observation (Blampain, Ott, & Delevoye-Turrell, 2018) and language processing (Aravena et al., 2012; Frak, Nazir, Goyette, Cohen, & Jeannerod, 2010; Nazir et al., 2017) to number processing (Michirev, Lindemann, Kühne, Fischer, & Raab, 2024; Miklashevsky, Fischer, & Lindemann, 2022; Miklashevsky, Lindemann, & Fischer, 2021). This approach has also been extended to low-level phenomena such as saliency detection (Novembre, Chavane, & Ross, 2018; Rangel, Novembre, & Wessel, 2024). By continuously recording spontaneous grip force microvariations (i.e., involuntary low-amplitude fluctuations) at millisecond resolution and millinewton sensitivity, this approach makes it possible to monitor the dynamic unfolding of motor engagement (Frak et al., 2010).

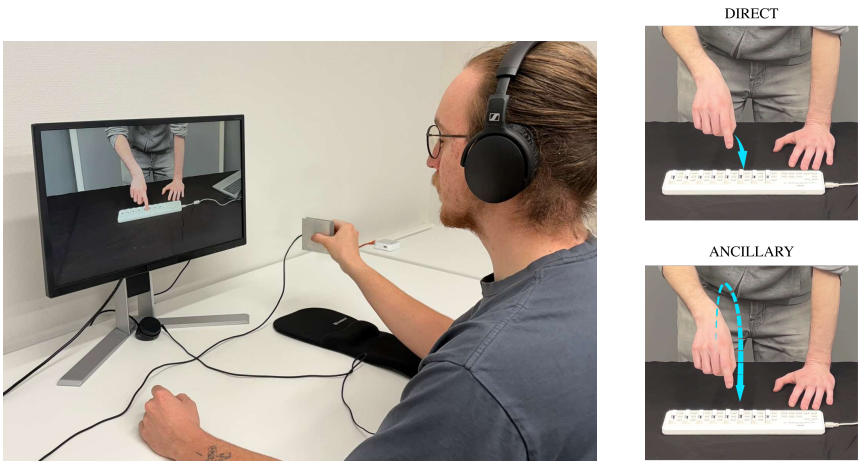
Historically, the study of motor resonance has relied predominantly on transcranial magnetic stimulation (TMS), which, despite allowing precise timing of corticospinal excitability modulations (Guidali, Picardi, Franca, Caronni, & Bolognini, 2023; Lepage et al., 2010), does not provide continuous measurement. Continuous methods such as electroencephalography (EEG) and magnetoencephalography (MEG) have complemented this approach. EEG captures the temporal unfolding of motor simulation (Avanzini et al., 2012) at the cost of spatial resolution, and MEG offers improved spatial resolution (van Schie et al., 2008) but under highly constrained experimental conditions. These approaches demonstrate that motor resonance unfolds over time, but they either rely on discrete sampling (TMS) or impose strong experimental constraints (EEG/MEG). Continuous GF recordings provide a complementary approach, capturing the full temporal evolution of motor engagement in more flexible and ecological settings.

Furthermore, the peripheral nature of GF does not preclude meaningful parallels with neural signals: studies combining grip force with EEG have reported close correspondences between GF dynamics and cortical motor activity (Novembre et al., 2018, 2019). This does not mean, however, that GF maps onto a single neural locus: as a continuous peripheral readout, it reflects the integrated output of a cortical and subcortical motor network (Caligiore et al., 2017; Kornhuber, 1971).

In action observation specifically, Blampain et al. (2018) established GF recordings as an index of motor resonance and showed that GF variations scale with the intensity of the observed action. Yet, as in most action-observation paradigms, movement intensity and action outcome were naturally coupled, making it difficult to dissociate their contributions. Digital musical interactions provide a way to address this limitation. This combination makes continuous GF recording particularly suited to the present aim: examining how motor engagement unfolds across a temporally structured digital interaction.

## The present study

Two research questions guided this study. First, does the observation of digital musical interactions give rise to motor resonance that follows the temporal unfolding of the




(a) Experimental setup (b) Ancillary gesture

**Fig. 1**: (a) A participant holds the grip force sensor (ATI Mini40, fitted with custom aluminum contact plates) between thumb, index, and middle fingers, with the forearm resting on a cushioned support at approximately 45°, while observing a digital musical interaction on screen. (b) Illustration of the two gesture conditions: in the *Direct* condition (top), the performer presses the button directly from a resting position approximately 2.5 cm above the controller; in the *Ancil* condition (bottom), the hand first rises from the same resting position to an apex of approximately 22 cm

observed action? Second, does an ancillary gesture, a larger preparatory movement with no effect on the sound, reshape this motor response beyond a simple scaling with movement amplitude?

To address these questions, participants observed a musician triggering a sound modulation via a digital controller while their grip force was recorded continuously. The contact gesture was either direct or preceded by the ancillary gesture (Fig. 1b). Critically, the digital instrument was configured so that the button press acted as a simple on/off trigger, producing the same sound modulation in both gesture conditions.

Each stimulus was repeated five times consecutively, both to increase the amount of data collected and to allow exploratory analysis of the intra-trial dynamics across repetitions.

Our analyses used a time-resolved Bayesian approach to capture the dynamic evolution of evidential strength across conditions. As an exploratory complement, post-trial ratings documented participants' explicit judgments of the interaction, in particular the contribution they attributed to the musician relative to the computer system.

# Methods

## Participants

Forty-two participants (15 male, 27 female; mean age = 22.2 years, SD = 3.7; age range 18–39) were retained for analysis. The target sample size was set to 50 participants, anticipating participant loss due to signal-quality exclusions, which are common in continuous force recordings. Of the 50 recruited, 49 completed the experiment; one participant withdrew after arrival. Seven additional participants were excluded during artifact screening (see Artifact screening and exclusion). Participants were screened for normal or corrected-to-normal vision and hearing, reading ability, absence of trauma or injury to the right upper limb, and handedness (pre-experiment questionnaire; Online Resource 3). Thirty-nine participants were right-handed and three were left-handed. Familiarity with electronic music and the experimental instrument was assessed via post-experimental questionnaire (Online Resource 3) and scored on a normalized 10-point scale (M = 2.1, SD = 2.0, range: 0–7.5).

After being informed about the 45-minute session and the voluntary nature of participation, all participants provided written informed consent. Participants were compensated 12 € via bank transfer. All participants were recruited through the first author's lab electronic participation system. Recruitment and data processing complied with the European General Data Protection Regulation (GDPR) and were approved by the Data Protection Officer (DPO) and the local ethics committee.

## Stimuli

The stimuli consisted of 5-second video clips depicting a simple musical interaction in which a performer presses a button on a digital controller to trigger a change in sound volume (Figure 2, top). Each clip was repeated five times consecutively, yielding a single trial of approximately 25 seconds. Only two base video recordings served as source material, one per ancillary gesture condition. From these visual bases, a post-production process independently manipulated three levels of auditory transparency and three levels of visual augmentation, generating a full set of 18 stimuli ($2 \times 3 \times 3$ stimulus matrix).

The stimulus set is publicly available on the Open Science Framework (OSF) at https://osf.io/yda5r/.

The auditory stream consisted of a continuous sound present from stimulus onset, which changed in volume at the moment of the button press. Using a continuous rather than onset-triggered sound was designed to minimize auditory artifacts that could temporally align with the visual press event.

### Interaction sequence

The interaction in all stimuli followed an identical temporal sequence divided into seven phases (Figure 2, top). Durations below are approximate and averaged across the two base recordings (see Appendix Table 3):

1. gray fade-in (500 ms),
2. initial idle period with the performer's hand above the controller (∼1,000 ms),

3. gesture phase (∼400 ms): direct press in the *Direct* condition, upward preparatory movement followed by press in the *Ancil* condition,
4. contact with the controller (∼700 ms),
5. return gesture toward the starting position (∼275 ms),
6. second idle period (∼1,500 ms),
7. gray fade-out (500 ms).

## Experimental factors

The main experimental manipulation reported in this work concerned the absence or presence of an ancillary (preparatory) gesture during a simple interaction in which a musician pressed a button on a digital musical instrument. In the *Direct* condition, the press was executed directly, without preparatory motion. In the *Ancil* condition, the performer executed a short upward hand movement before pressing the button, starting from a resting position approximately 2.5 cm above the controller, shared with the *Direct* condition, and reaching an apex of 22 cm before returning to press (Fig. 1b). Importantly, this ancillary gesture had no impact on the auditory outcome. Non-sonic gestures of this kind are common in musical performance and are known to influence both performer expressivity and audience perception (Jensenius & Wanderley, 2010; Wanderley, 2002).

The full experimental design followed a $3 \times 3 \times 2$ factorial structure, comprising three factors: ancillary gesture (*Direct* vs. *Ancil*), interaction transparency (three levels), and visual augmentation (three levels).

The present report focuses on two complementary questions: whether GF dynamics reveal structured motor engagement during the observation of digital interactions, and whether this temporal profile is reshaped by the ancillary gesture beyond a simple scaling with movement amplitude. The transparency and visual-augmentation factors address distinct research questions and are not analyzed here; they are briefly described below.

Interaction transparency varied the degree of gesture–sound coupling, ranging from a salient and temporally synchronous sound modification (high transparency) to a barely perceptible and temporally offset modification (low transparency), with an intermediate level. Visual augmentation consisted of dynamic graphic overlays reflecting the gesture–sound mapping, deployed at three levels: absent, congruent, or incongruent with the interaction.

Data were therefore collapsed across transparency and visual-augmentation levels for the present analyses. Exploratory inspection of the corresponding time courses did not indicate that transparency or visual augmentation modulated the ancillary-gesture contrast reported here. A high- versus low-transparency comparison is provided as a robustness check in Appendix Fig. 9.

## Procedure

The experiment took place in a temperature-controlled room at the first author's laboratory. Room temperature and humidity were recorded throughout the session;

full environmental and setup details are provided in Supplementary Material (Online Resource 1) for transparency and reproducibility.

Upon arrival, participants provided informed consent, then removed personal items before being seated in front of the stimulus display. The grip force device was disinfected in their presence, and participants sanitized their hands with an alcohol-based gel. They then completed a brief questionnaire (Online Resource 3) assessing handedness, vision, hearing, and reading ability. Headphones with disposable ear covers were provided, and participants adjusted playback volume to a comfortable level.

Participants were instructed to adopt a standardized posture with the grip force device held in the right hand (Fig. 1). All participants, including the three left-handers, used their right hand. A rolling partition was then placed between the participant and the experimenter, concealing the control screen from view.

Following the final trial, participants completed a post-experiment questionnaire assessing their musical background, concert attendance, electronic music experience, and familiarity with the experimental instrument. The pre- and post-experiment questionnaires are provided in Online Resource 3.

## Main task procedure

### Individual Pretrial Target (IPT)

Prior to the main task, a stable individual baseline pressure, referred to as the Individual Pretrial Target (IPT), was established for each participant. Participants held the device and applied a natural, comfortable amount of force for approximately one minute; the average grip force over this period served as the target pressure at the onset of each trial. Whereas previous grip force studies often used fixed target forces (typically $\sim$1.5 N), this individualized approach accounts for inter-individual variability in grip force and sensor-specific calibration.

Participants then completed a brief familiarization phase consisting of three representative stimuli. Following this phase, they received task instructions and began the main task.

At the onset of each trial, the experimenter guided the participant back to their IPT level, verified via real-time monitoring; once reached and maintained for 3 seconds, the video stimulus began. Each trial consisted of five consecutive repetitions of the same stimulus. Participants were instructed to observe the interaction attentively throughout. After each trial, they set the device down and answered a brief series of statements probing their subjective representation of the interaction, grouped into two categories: attributed agency and error detection. Each statement was answered yes/no and followed by a confidence rating on an unmarked slider ranging from “answered at random” to “fully certain”; this rating was used to weight the corresponding response. The trial-by-trial statements are listed in Appendix Table 2.

The 18 stimuli were presented in randomized order for each participant. Participants were not informed about the specific manipulation under study (the ancillary gesture). A short break was offered halfway through the session, though most participants declined. The total session lasted approximately 45 minutes, of which roughly 20 minutes corresponded to active grip force recording.

## Equipment

### Grip force setup

Grip force was recorded using an ATI Mini40 force sensor (SI-40–2 calibration) fitted with custom aluminum contact plates. The sensor was connected to the experimental computer through a National Instruments NI-6221 acquisition card. Stimulus presentation was controlled in PsychoPy (v2023.2.2), and force data were recorded in parallel via a custom Python acquisition script.

To synchronize stimulus events with force recordings, digital event markers were converted into analog pulses using an Arduino Uno microcontroller. A secondary synchronization stream was also recorded through Lab Streaming Layer (LSL; Kothe et al., 2025), providing redundancy and offline verification of temporal alignment.

### Digital musical instrument configuration

A minimal digital musical instrument (DMI) was used to create a controlled real-time gesture–sound interaction. It consisted of a standard MIDI controller (Korg nanoKONTROL) connected to a laptop, which handled both sound synthesis and gesture-to-sound mapping implemented in Pure Data (Puckette, 1996). Sound synthesis relied on amplitude modulation (a 440 Hz sine wave amplitude-modulated at 40 Hz). The button registered presses as binary on/off events, independent of the force applied to it; each press triggered a change in sound volume, whose perceptual salience and temporal coupling to the press were manipulated to produce the three transparency levels (see Experimental Factors).

An extended version of the setup is available in Supplementary Material (Online Resource 1).

## Data acquisition and processing

### Acquisition

Grip force data were acquired following the protocol established by Nazir et al. (2017). The sensor's maximum sampling frequency is 4.9 kHz; the three force components (Fx, Fy, Fz) were recorded at 200 Hz to match previous studies, and only the force applied perpendicularly to the sensor (Fz) was analyzed.

### Preprocessing

The continuous signal was low-pass filtered using a fourth-order Butterworth filter (15 Hz cutoff), applied bidirectionally via `filtfilt` (Virtanen et al., 2020) to eliminate phase distortion. Filtering was applied prior to trial segmentation to minimize edge artifacts at trial boundaries.

An additional Savitzky–Golay (SG) smoothing filter (order = 1, window = 15 samples, 75 ms) was applied after low-pass filtering to reduce point-to-point variability ahead of the time-resolved analyses, while preserving temporal modulations at the scale of the observed interaction. Robustness checks are provided in Appendix Fig. 6; they show comparable temporal profiles, albeit with increased noise.

### *Artifact screening and exclusion*

Following Nazir et al. (2017), grip force signals were screened for outlier patterns at the level of individual repetitions. The exclusion threshold for excessive variation was partially individualized relative to each participant's IPT, using a weighted combination of the standard fixed threshold (200 mN) and an IPT-scaled value, with the fixed threshold as a lower bound; the exact computation is provided in Online Resource 1. Rapid changes exceeding 100 mN within 100 ms were excluded regardless of IPT. Trials with no valid repetitions were discarded. Participants with fewer than 80% valid trials were excluded from further analysis.

Across participants, IPT averaged 2.36 N (SD = 1.10 N; range: 0.84–5.78 N). Participants excluded due to excessive signal variability did not exhibit markedly higher IPT values than those retained ($BF_{10}$ = 0.347). From the initial sample (n = 49), the IPT-weighted criterion excluded 31% of repetitions across all participants (84% of trials retained at least one valid repetition). Seven participants were then removed for having fewer than 80% valid trials. Within the resulting sample of 42 participants, 2,901 of 3,780 repetitions were retained, i.e., 23% repetition-level exclusion. The grand mean signal computed on the full initial sample (n = 49), prior to artifact rejection, displayed comparable temporal dynamics (Appendix Fig. 7).

### *Temporal alignment*

Since stimuli were derived from two base video recordings with slightly varying phase durations, grip force data were resampled within each phase to match the mean duration across conditions, normalizing the temporal structure across trials while preserving within-phase dynamics. After resampling, the *Ancil* and *Direct* conditions shared identical phase durations and aligned event markers in the analyzed time base. Pointwise comparisons were therefore performed at homologous moments of the interaction sequence rather than at unmatched absolute video times. Normalized phase durations are reported in Appendix Table 3; raw-signal plots without temporal alignment are also provided in the Appendix as robustness checks.

### *Baseline correction*

A baseline correction was applied to each individual repetition within a trial, following the general approach of Nazir et al. (2017). Unless otherwise specified, the baseline was computed as the mean grip force over a 100 ms window centered on the *fadeIn_end* marker (50 ms before, 50 ms after), and subtracted from the entire time series, centering the signal relative to the onset of the visible interaction.

## Statistical analysis

The analysis pipeline was designed to combine predefined phase-level summaries with time-resolved characterization of the grip force signal. Phase-level models were used to test windows defined by the temporal structure of the observed interaction. Time-resolved Bayesian analyses were used to describe the temporal evolution of evidential strength within and across these phases. Finally, temporally contiguous evidential

clusters identified in these time-resolved analyses were evaluated using a permutation-based cluster procedure and summarized with linear mixed models fitted over the corresponding cluster windows.

### Phase-level mixed models

For the overall grip force response, phase-level means were compared against the baseline (0 mN). The phase windows were defined from the temporal structure of the stimulus sequence: fade-in, initial idle period, movement ignition (direct or ancillary), button press (trigger), return gesture, second idle period, and fade-out. For the *Ancil* vs. *Direct* contrast, linear mixed models were fitted separately for each phase, with condition as a fixed effect and participant as a random intercept. The *Ancil* condition served as the reference level, such that negative estimates for *Direct* indicate lower grip force in the *Direct* condition than in the *Ancil* condition. Since these phase windows were defined a priori by the interaction sequence, no correction for multiple comparisons was applied across phases.

### Time-resolved Bayesian Evidence Display

Time-resolved analyses were conducted within a Bayesian framework using paired-sample $t$-tests with a weakly informative Cauchy prior centered at zero (scale $r = 0.707$), following standard recommendations for default Bayesian hypothesis testing (Rouder, Speckman, Sun, Morey, & Iverson, 2009). This prior imposes no directional bias on effect size. Bayes Factors were used to quantify the strength of evidence for or against differences between conditions. At each time point, a paired Bayesian $t$-test was computed on participant-level condition means, capturing the dynamic unfolding of evidential strength over the course of the interaction. All Bayesian pointwise analyses were conducted using the `Pingouin` Python package (v0.5.5; Vallat, 2018).

Resulting Bayes Factors were transformed into decibans ($10 \times \log_{10}(\mathrm{BF}_{10})$), hereafter $\mathrm{dBF}_{10}$. Under this form, they provide an additive and interpretable measure of evidential strength (Jeffreys, 1961; Wagenmakers et al., 2018): positive values indicate evidence for a difference between conditions, whereas negative values indicate evidence for the absence of a difference. Following Jeffreys (1961), we adopt the thresholds of 5, 10, 15, and 20 decibans as indicating moderate, strong, very strong, and decisive evidence, respectively. Deciban thresholds and their correspondence to evidential strength are illustrated in Appendix Fig. 12.

Deciban values were plotted over time as an evidential trace, with color and gradient opacity encoding the strength of evidence. We refer to this visualization as the Time-resolved Bayesian Evidence Display (TBED), following Yavuz et al. (2026) (preprint).

### Permutation-based evaluation and mixed-model summary of TBED clusters

To assess whether temporally contiguous evidential clusters identified in the TBED exceeded what could be expected under label permutation, we applied a permutation-based cluster procedure to the $\mathrm{dBF}_{10}$ time series adapted from Maris and Oostenveld

(2007). Clusters were defined as contiguous time points exceeding the moderate-evidence threshold for $H_1$ ($dBF_{10} > 5$ dBan). For each observed cluster, its temporal extent (the number of contiguous time points exceeding threshold) was used as the cluster-level test statistic.

The null distribution was generated by randomly permuting condition labels within participants and recomputing the full time-resolved $dBF_{10}$ trace for each permutation. For each permuted dataset, the maximum cluster extent was retained, yielding a permutation distribution of cluster sizes under the null hypothesis. Observed clusters were then evaluated against this distribution (2000 permutations).

Clusters supported by the permutation procedure were then characterized descriptively with a linear mixed model fitted on grip force values averaged within the cluster bounds, with condition as a fixed effect and participant as a random intercept. This provided an interpretable estimate of the localized condition effect in mN, with its standard error, rather than an independent confirmatory test of the cluster. Complete phase-level models, condition-level descriptive and nonparametric statistics, and the full TBED permutation null distributions are reported in Online Resource 2, for the overall signal and for both baselines (*fadeIn_end* and *trigger_start*).

All data and analysis scripts are publicly available on the Open Science Framework (OSF) https://osf.io/yda5r/.

## Main analyses

The first analysis examined the baseline-corrected grip force time series throughout each trial in relation to the temporal structure of the stimulus (Fig. 2). We assessed whether grip force tracked motor engagement during digital interaction observation by comparing the full-trial signal against the baseline defined at the onset of the visible interaction (*fadeIn_end*).

The second analysis assessed whether and to what extent grip force dynamics were modulated by gestural context, through two contrasts. The first compared GF time courses between the *Ancil* and *Direct* conditions, baseline-corrected at *fadeIn_end* and analyzed over the full stimulus duration. The second rebaselined both conditions at the moment of contact with the instrument (*trigger_start*), to examine the response specifically associated with the trigger event independently of preceding cumulative amplitude differences. This transformation does not remove the influence of the preceding gesture, but rather re-expresses it as a difference in the motor state at the time of the trigger event. Under this baseline, differences between conditions reflect state-dependent modulation of the response to a shared event, rather than differences in signal amplitude accumulated prior to contact.

Repetition-by-repetition analyses and post-trial subjective ratings were treated as exploratory.

For discrete comparisons outside the TBED framework, Bayes Factors are reported as ratios rather than in decibans.

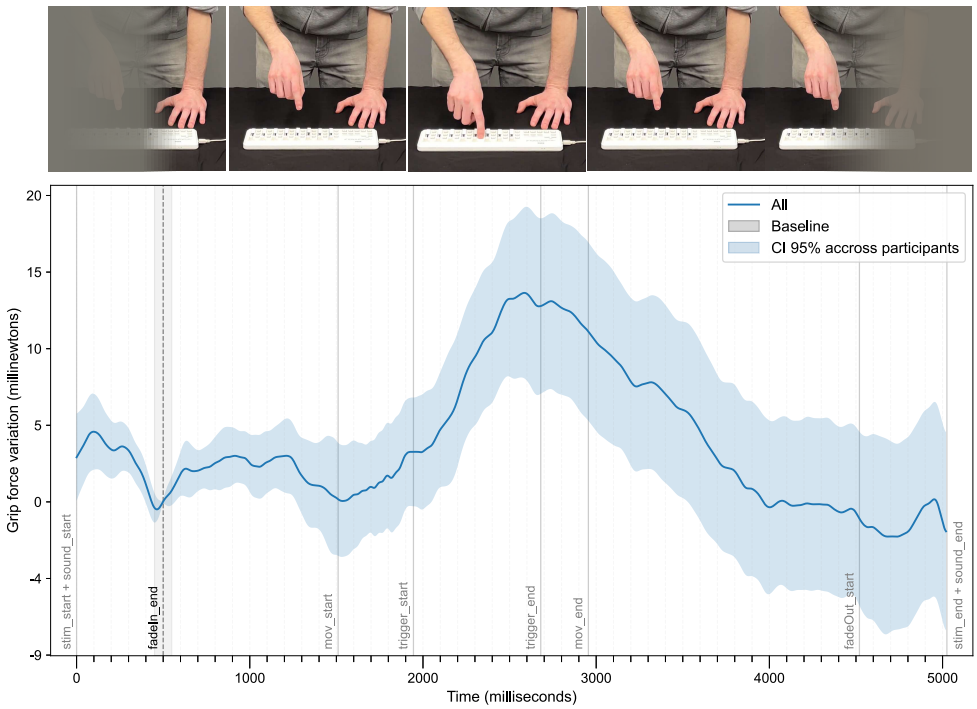



**Fig. 2**: Top panel: Representative video frames illustrating five key moments of the observed interaction—fade-in, initial idle period, button press (trigger), second idle period, and fade-out. Bottom panel: Grand mean grip force signal (n = 42), low-pass filtered and smoothed, temporally aligned to the mean phase structure, and baseline-corrected at *fadeIn_end* (dashed vertical line). The gray shaded area indicates the 100 ms baseline window; the blue shaded area represents the 95% confidence interval across participants

# Results

## Grip force dynamics during digital interaction observation

Across all valid trials, the average grip force (GF) recorded during the observation period was 2.29 N (SD = 1.07 N), with individual values ranging from 0.70 to 6.01 N, reflecting substantial interindividual variability. Baseline-corrected trial-wise force variation averaged 3.62 mN (SD = 11.65 mN), with values ranging from -26.64 to 44.19 mN.

The grand mean grip force signal, averaged across all conditions and repetitions, revealed a structured phase-locked modulation broadly following the temporal segmentation of the observed interaction (Fig. 2): GF decreased throughout the fade-in period, reached a local minimum around *fadeIn_end*, then rose, remained relatively stable through the idle period preceding movement onset, before rising markedly through the gesture and contact phases, reaching its maximum of 13.64 mN during the

**Table 1**: Phase-level one-sample tests against zero for the grand mean grip force signal across all conditions and repetitions. Temporal windows correspond to the phases of the observed interaction sequence. Values are baseline-corrected grip force variations averaged within each window; positive values indicate grip force above baseline. Rows shaded in yellow indicate windows with $p < .05$

| One-sample t-test vs 0 — All | | | | | |
|---|---|---|---|---|---|
| **Phase** | $M$ (mN) | $SD$ | $t$ | $p$ | $BF_{10}$ |
| Stim Start–Fade-in end | +2.77 | 4.42 | +4.01 | < .001 | 105.1 |
| Fade-in end–Mov. Start | +2.04 | 4.96 | +2.63 | 0.012 | 3.4 |
| Mov. Start–Trigger Start | +1.28 | 10.52 | +0.78 | 0.440 | 0.2 |
| Trigger Start–Trigger End | +8.96 | 13.85 | +4.14 | < .001 | 149.8 |
| Trigger End–Mov. End | +12.44 | 18.02 | +4.42 | < .001 | 328.8 |
| Mov. End–Fade-out Start | +3.63 | 17.45 | +1.33 | 0.190 | 0.4 |
| Fade-out Start–Sound End + Stim End | -1.42 | 19.38 | -0.47 | 0.642 | 0.2 |

trigger phase. Approximately 550 ms after the end of the return gesture (*mov_end*), the signal returned toward baseline. These modulations relative to baseline were supported by the phase-by-phase analysis (Table 1) and by pointwise Bayesian analysis (see Appendix Fig. 6).

Finer-grained inspection revealed recurring short-latency inflection patterns around several phase transitions. As an exploratory analysis on time-aligned data, two windows motivated by the known temporal architecture of motor resonance during action observation were overlaid on the grand mean grip force signal: an early window (0–100 ms) associated with non-specific corticospinal facilitation (Lepage et al., 2010), and a later window (200–400 ms) associated with muscle- and context-specific modulation (Amoruso et al., 2016; Lepage et al., 2010). These windows are indicative rather than exact: temporal alignment normalizes phase durations across conditions, so absolute timings are approximate. The resulting visualization suggested recurring inflection patterns across several phase transitions, as depicted in Appendix Fig. 10.

## Sensitivity to the ancillary gesture

Grip force curves for the *Ancil* and *Direct* conditions, baseline-corrected at *fadeIn_end* and averaged across all repetitions on the same event-aligned time base, were largely similar across the observed sequence (Fig. 3). This similarity was reflected in TBED values predominantly favoring $H_0$, frequently below −5 dBan. Both curves reached comparable peak amplitudes ($BF_{10} = 0.262$). The *Ancil* curve peaked at 15.33 mN, the *Direct* curve peaked at 12.81 mN.

The *Ancil* curve began to rise around the apex of the ancillary gesture, followed by a brief inflection at the moment of contact (*trigger_start*) and a resumption of its rise, at which point it rejoined the trajectory of the *Direct* condition. The condition divergence became evidentially supported from *trigger_start* onward, as reflected in the phase-by-phase LMM for the *trigger_start–trigger_end* window ($\hat{\beta} = -4.262$, SE

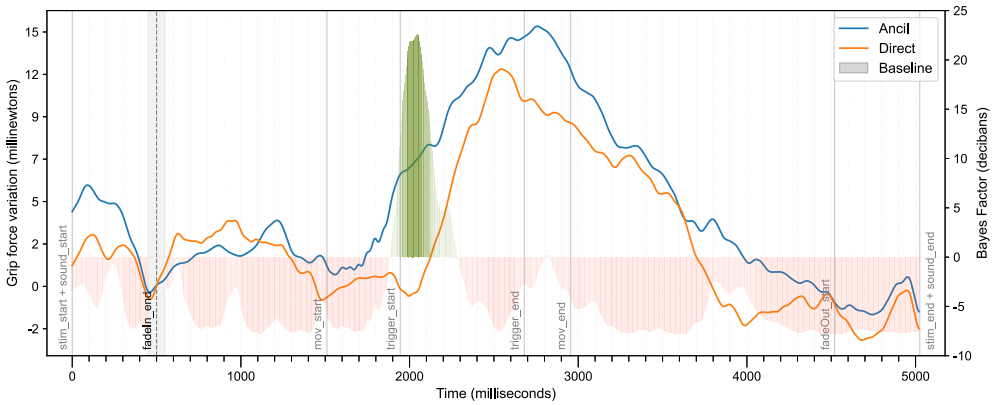



**Fig. 3**: Mean grip force signal for the *Ancil* (blue) and *Direct* (orange) conditions (n = 42), low-pass filtered and smoothed, temporally aligned, and baseline-corrected at *fadeIn_end* (dashed vertical line). The gray shaded area indicates the 100 ms baseline window. The TBED trace on the right axis shows pointwise Bayesian evidence for the condition contrast in decibans; green bars indicate evidence for $H_1$, red bars indicate evidence for $H_0$, and opacity scales with evidential strength. Phase durations were normalized to their across-condition mean for visualization and pointwise comparison. A raw, non-time-aligned, pre-artifact-rejection version is provided in Appendix Fig. 8

= 1.920, $z = -2.22$, $p = .026$). A TBED cluster of 52 samples was detected in the 1920–2175 ms aligned-time window and was supported by the permutation analysis ($p = .004$). The follow-up LMM fitted on the cluster window showed lower GF in the *Direct* condition ($\hat{\beta} = -6.930$, SE = 1.863, 95% CI [−10.58, −3.28]).

Complete phase-level analyses, condition-level statistics, and permutation results for both baselines are provided in Online Resource 2.

## Motor response at the trigger event

To examine the response specifically associated with the moment of button contact, grip force signals were rebaselined at *trigger_start*. This transformation expresses the signal relative to the motor state immediately preceding the trigger event, independently of prior cumulative amplitude differences. From this common baseline, the two conditions showed a clear divergence, with the *Direct* condition reaching a higher relative amplitude than the *Ancil* condition within the *trigger_start–trigger_end* window. The phase-level LMM showed a significant effect ($\hat{\beta} = 2.062$, SE = 1.009, $z = 2.04$, $p = .041$), and a 32-sample TBED cluster (2480–2635 ms) was supported by the permutation analysis ($p = .015$). The follow-up LMM fitted on the TBED-localized cluster showed higher GF in *Direct* ($\hat{\beta} = +4.633$, SE = 1.416, 95% CI [+1.86, +7.41]). Deviations outside this window were not interpreted as trigger-locked responses and were therefore not analyzed in this trigger-centered contrast (see Appendix Fig. 11 for the full time course).

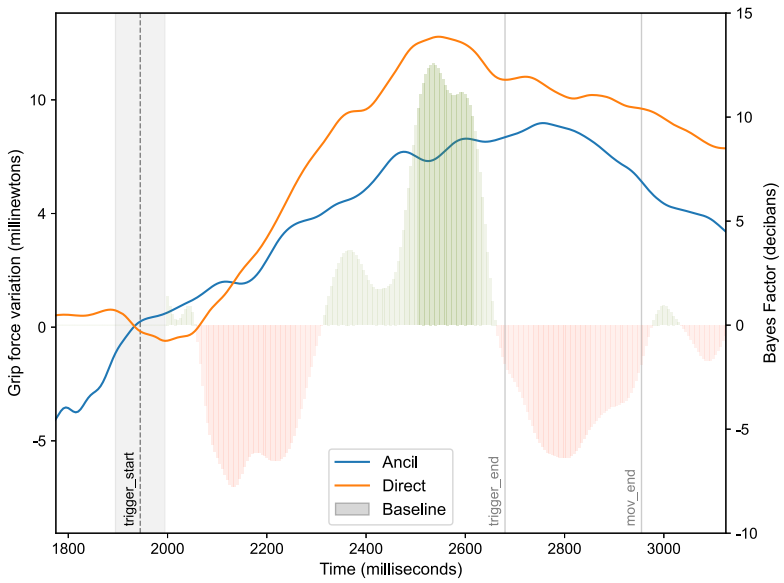



**Fig. 4**: Mean grip force signal for the *Ancil* (blue) and *Direct* (orange) conditions (n = 42), low-pass filtered and smoothed, temporally aligned, and rebaselined at *trigger_start*. The plot shows a zoomed view of the trigger window. Vertical lines indicate button contact (*trigger_start*), contact release (*trigger_end*), and the end of the return gesture (*mov_end*). The gray shaded area indicates the 100 ms baseline window. The TBED trace on the right axis shows pointwise Bayesian evidence for the condition contrast in decibans; green bars indicate evidence for $H_1$, red bars indicate evidence for $H_0$, and opacity scales with evidential strength. The full time course is provided in Appendix Fig. 11

## Intra-trial dynamics across repetitions

To characterize how the signal evolved within a trial, grip force dynamics were examined separately for each of the five repetitions (Fig. 5). Rep 1 showed the largest global response, with a mean value of 13.7 mN and a peak of 27.1 mN. Rep 2, Rep 3, and Rep 4 showed comparable attenuated dynamics, with a mean value of 0.9 mN ($SD = 0.9$ mN) and a mean peak of 10.4 mN ($SD = 2.5$ mN). Despite this amplitude reduction, the overall temporal profile of the response remained broadly preserved across repetitions. Phase-by-phase LMMs showed significant differences between Rep 1 and Rep 2–4 across the main interaction phases.

A permutation cluster test further delimited the temporal extent of this repetition-related modulation: a large cluster was identified from shortly after the beginning of the interaction to shortly before its end (705–4235 ms; 707 samples; $p = .002$), while the fade-in and fade-out phases remained outside the significant cluster.

Rep 5 showed a distinct temporal profile, starting near baseline at *fadeIn_end* and rising progressively through the interaction. A secondary TBED comparing Rep 5 with Rep 1 showed values ranging from evidence for $H_0$ to near-neutral evidence in

the early phases (−6 to −2 dBan), followed by divergence in the later interaction window, where Rep 5 remained lower than Rep 1. This later divergence is supported by an LMM fitted on the *trigger_end–mov_end* window ($\hat{\beta} = -13.88$, $p = .004$).

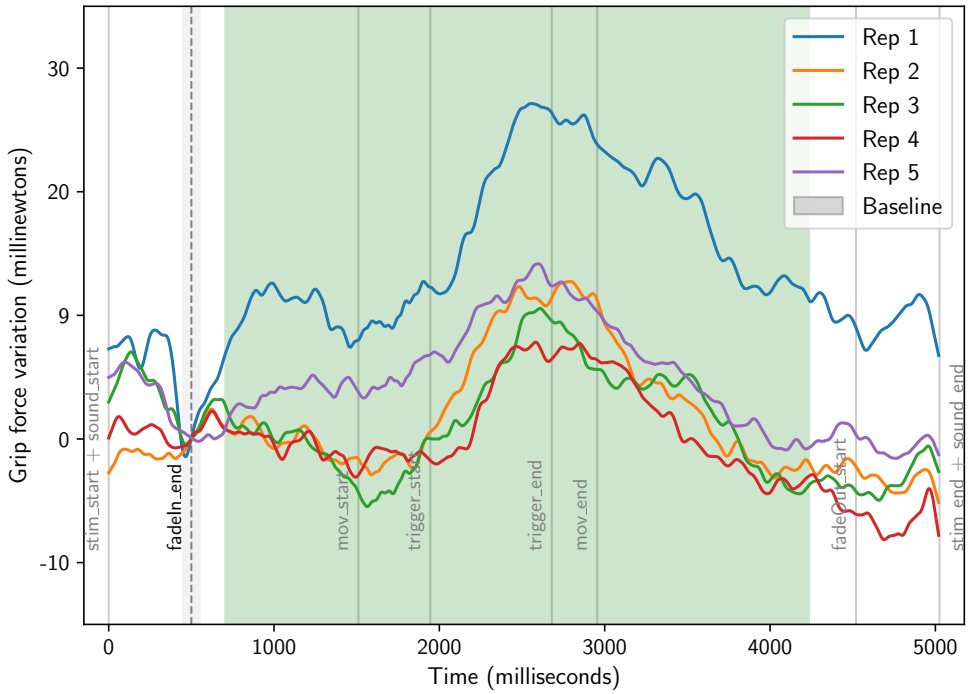



**Fig. 5**: Mean grip force signal for each of the five repetitions (n = 42), low-pass filtered and smoothed, temporally aligned, and baseline-corrected at *fadeIn_end*. Vertical lines indicate key phase transitions. Repetition 1 (blue) shows a higher global amplitude than subsequent repetitions, whereas Repetition 5 (purple) shows a distinct early profile, starting near baseline and rising progressively. The green shaded area indicates the significant cluster identified by the permutation cluster test (705–4235 ms; $p = .002$)

## Post-trial subjective ratings

Post-trial questions were grouped into two categories: attributed agency and error detection. Neither category showed a difference between *Ancil* and *Direct* conditions, with Bayes factors favoring no difference for both measures (attributed agency: $BF_{10} = 0.09$; error detection: $BF_{10} = 0.08$).

# Discussion

This study used continuous grip force (GF) recordings and time-resolved Bayesian analysis to examine motor resonance during the observation of digital interactions. Participants observed a musician triggering a sound modulation on a digital musical instrument, either directly or after an ancillary gesture, a larger preparatory movement with no effect on the sound. By embedding this manipulation within a musical-performance context, we asked whether motor resonance follows the temporal structure of the action, and whether it is shaped by gestural context beyond a simple scaling with movement amplitude.

Together, the results show that GF dynamics were temporally structured and phase-locked to the interaction sequence. The larger ancillary gesture produced a local increase and reshaped the temporal profile of the signal, but did not yield a reliably larger peak response. Moreover, when the signal was realigned to the triggering event, the response to the same contact event was smaller after the ancillary gesture. These findings suggest that motor resonance during digital interaction observation is sensitive to the temporal and gestural organization of the observed action, rather than to movement amplitude alone.

## Motor resonance in digital interaction observation

The grip force signal averaged across all conditions revealed a structured phase-locked modulation that broadly followed the temporal segmentation of the observed interaction. This pattern, supported by both phase-level models and pointwise Bayesian analysis (see Appendix Fig. 6 for raw-signal robustness analyses), provides evidence consistent with motor resonance during the observation of digital interactions.

As depicted in Figure 2, GF decreased during the fade-in and then rose immediately after its offset, while the performer's hand was already suspended above the controller but before any overt movement toward it occurred. This early modulation suggests that the signal was not driven by observed movement alone, but also by the perceptual onset of an action-ready scene and by attentional preparatory processes. This pattern is consistent with an attentional preparatory component of motor resonance (Baldissera et al., 2001; Puglisi et al., 2017).

### Temporal sensitivity

The GF signal increased before contact when the trigger was preceded by the ancillary gesture, whereas in the *Direct* condition the corresponding increase was concentrated around the contact event. Thus, GF captured a temporal redistribution of motor engagement across the interaction sequence: in the *Ancil* condition, part of the response emerged during the preparatory movement, whereas in the *Direct* condition it was concentrated around the contact event.

Short-latency inflections were also observed around several phase transitions, approximately 0–100 and 200–400 ms after key events, in both the grand-mean and per-condition signals (Appendix Fig. 10). These latencies are broadly consistent with the early non-specific and later muscle- and context-specific phases of motor resonance documented with TMS and EMG (Amoruso et al., 2016; Lepage et al., 2010; Naish

et al., 2014). Although post hoc and awaiting replication, these observations further illustrate the fine-grained temporal sensitivity of GF around aligned event transitions.

### Beyond movement amplitude

The ancillary gesture reshaped the temporal profile of the response without producing a reliably larger peak amplitude, despite its markedly greater physical extent (apex at ∼22 cm vs. ∼2.5 cm in the *Direct* condition). This dissociation indicates that GF dynamics were not governed by movement amplitude alone.

A similar contrast appears within the *Direct* condition, between the trigger gesture and the subsequent return movement. The two movements were of comparable amplitude (∼2.5 cm, traversed in opposite directions) and similar duration (233 versus 300 ms; Appendix Table 3), and both corresponded to symmetrically designed sound changes, switching the modulation on and off. Yet they elicited markedly different GF profiles: the trigger gesture initiated the sound modulation and was accompanied by a clear GF increase, whereas the return gesture, despite initiating the end of the modulation, left no comparable deflection and was largely absorbed into the post-trigger descent (Fig. 3). Together with the comparable temporal profiles observed across transparency levels (Appendix Fig. 9), this within-condition asymmetry suggests that motor engagement was shaped by the temporal and gestural organization of the sequence rather than by movement amplitude alone.

We interpret this pattern as compatible with sensitivity to the gesture's role within the unfolding gesture–sound sequence, consistent with evidence that motor simulation is sensitive to contextual and consequence-related information beyond observable kinematics (Alaerts et al., 2010; Gangitano, Mottaghy, & Pascual-Leone, 2004).

### Trigger-centered modulation

Rebaselining the GF signal at the moment of contact (*trigger_start*) revealed a further asymmetry (Fig. 4). When expressed relative to the motor state indexed immediately preceding contact, the *Direct* condition showed a greater relative GF increase than the *Ancil* condition, despite the auditory outcome being identical across conditions. This suggests that the response to the trigger event was state-dependent, reflecting the level of motor engagement established by the preceding gestural context.

Within this framework, the reduced response observed in the *Ancil* condition could correspond to a form of saturation or normalization, whereby part of the available dynamic range had already been recruited during the ancillary gesture. This interpretation, which assumes a limited dynamic range available for involuntary motor engagement, is compatible with the low-intensity modulations typically reported in studies of motor resonance (Loporto, Holmes, Wright, & McAllister, 2013).

An alternative interpretation is that the ancillary gesture reduced uncertainty about the upcoming contact event, resulting in a smaller response at contact (Huron, 2008). Further work will be required to distinguish these possibilities.

## Grip force and TBED as tools for continuous monitoring of motor resonance

The present study extends the grip force paradigm to the observation of digital interactions and applies the Time-resolved Bayesian Evidence Display (TBED) to characterize the temporal dynamics of motor engagement. Together, they provide a methodological contribution that complements existing approaches to motor resonance.

### Grip force as a continuous index of motor resonance

Continuous GF recording has established itself as a sensitive peripheral readout of motor involvement across a range of cognitive tasks (Aravena et al., 2012; Miklashevsky et al., 2022; Nazir et al., 2017), and as an index of motor resonance in action observation specifically (Blampain et al., 2018). GF recordings complement TMS by providing temporal continuity: rather than probing corticospinal excitability at discrete time points, they capture the unfolding of motor engagement across an observation sequence. The present results extend this sensitivity in a digital interaction context, showing that GF tracked the phase structure of the observed interaction and responded to a subtle gestural manipulation.

The exploratory repetition analysis further illustrates this temporal sensitivity. Although the first repetition elicited the largest response, subsequent repetitions preserved a broadly similar temporal profile at lower amplitude. This pattern suggests that repeated exposure may reduce the magnitude of motor engagement while partly preserving its temporal organization. Whether this attenuation reflects attentional decline, predictive suppression, or task-related anticipation remains an open question.

### Time-resolved Bayesian evidence and phasic decomposition

The TBED, introduced in Yavuz et al. (2026) (preprint), extends the logic of continuous measurement to statistical inference by providing a pointwise evidential trace that evolves in parallel with the GF signal. This representation is particularly useful for temporally structured signals, where the timing and duration of effects are theoretically meaningful. In the present study, visually salient TBED excursions were evaluated with a permutation procedure adapted from cluster-based testing (Maris & Oostenveld, 2007) and summarized with mixed models fitted over the corresponding windows. This combination allowed the TBED to guide signal interpretation while keeping localized effects tied to formal evaluation.

The rebaselining approach applied here illustrates a complementary methodological principle. Decomposing a continuous interaction into phasic segments allows motor engagement to be evaluated at theoretically motivated time points, in addition to the analysis of its global evolution from an early reference event such as *fadeIn_end*. By analogy with continuous language paradigms, where a sentence can be analyzed both as a sequence of discrete words and as an integrated unit carrying global meaning, an observed interaction can be approached both as a succession of discrete gestural events and as a temporally organized whole. This choice of analytical scope also raises the question of temporal granularity: characterizing slow evolution across an

entire trial may tolerate stronger smoothing or low-pass filtering, whereas resolving brief event-related modulations requires preserving faster temporal components and therefore using less restrictive preprocessing.

### Digital musical interactions as a paradigm for motor resonance research

Digital musical interactions provide a useful setting for motor resonance research because they preserve an ecologically meaningful gesture–sound relation while allowing gesture kinematics, auditory outcome, and temporal structure to be experimentally controlled and selectively dissociated. This combination is difficult to obtain in standard action-observation paradigms, where movement kinematics, inferred effort, and action outcome are usually physically coupled.

The present findings illustrate this potential. The ancillary gesture altered the motor state from which the trigger event was processed: in the *Ancil* condition, GF had already increased before contact, whereas in the *Direct* condition the response was centered on the contact event. These two patterns suggest different temporal organizations of motor engagement with the same auditory event: one contextually embedded by the preceding gesture, the other contact-centered. In this sense, ancillary gestures may act as contextual cues for the observer, shaping motor engagement before the sound-triggering action occurs, consistent with their communicative and expressive role in musical performance (Jensenius & Wanderley, 2010; Wanderley, 2002).

The post-trial ratings provided strong evidence for the absence of an *Ancil–Direct* difference in explicit judgments of agency and error detection, contrasting with the GF modulations observed in the same comparison. This suggests that GF captured motor dynamics not reflected in these selected explicit judgments.

## Limitations and future directions

This study examined motor resonance in a digital interaction context where gesture and outcome are not necessarily physically constrained. The interaction observed here instantiated this decoupling within a musical performance setting; yet it remained a simplified musical situation rather than a fully artistic performance, and the ancillary gesture is not a common feature of digital interactions. It is therefore a hybrid interaction that may not fully conform to canonical digital or musical contexts, which limits the generalizability of the findings. It nevertheless provides a useful first step, laying the groundwork for future work on motor resonance in digitally mediated interactions.

A related interpretive limitation concerns the level at which the observed effect of gestural context arises. The present data show that motor engagement is sensitive to the temporal and gestural organization of the observed interaction, but they do not determine whether this modulation reflects perceived action–outcome structure directly, or whether it is mediated by attentional orienting and predictive processes.

While no dedicated silent condition was included, the robustness check reported in Appendix Fig. 9 suggests that the auditory outcome was unlikely to be the sole driver of the reported modulations: notably, the lower-salience condition produced

comparable or higher GF amplitudes than the higher-salience condition, which would not be expected if auditory salience were the primary driver.

Looking ahead, several methodological directions merit attention. Exclusion rates in continuous grip force paradigms can lead to substantial data loss. For this reason, approaches in the literature have adapted the fixed threshold to stimulus duration (Miklashevsky et al., 2021), or indexed exclusion criteria on individual comfort force levels relative to the historical 1.5 N reference (Yavuz et al., 2026). The present study adopts an individualized, alpha-weighted threshold to make the exclusion procedure as transparent and adjustable as possible for future studies. All these approaches are principled, but the question of which force-variation threshold should be applied, and within what bounds, cannot be fully resolved until the consequences for signal quality are systematically characterized.

### Conclusion

This study investigated motor resonance during the observation of digital interactions using continuous grip force recordings. Situating the paradigm within a digital musical interaction allowed movement amplitude to be dissociated from auditory outcome in a controlled yet more naturalistic observational setting.

Time-resolved Bayesian analysis revealed structured temporal patterns in the signal, including modulations aligned with key interaction events. The larger ancillary gesture produced a local increase and reshaped the temporal profile of the signal, but did not yield a reliably larger peak response. Moreover, when the signal was realigned to the triggering event, the response to the same contact event was smaller after the ancillary gesture. These patterns indicate that GF dynamics were not governed by observed movement amplitude alone, but were shaped by the temporal and gestural organization of the observed interaction.

These findings are consistent with the view that motor resonance during action observation is shaped not only by movement kinematics, but also by the temporal and contextual role of an observed movement within the broader interaction sequence. They also support the use of grip force as a continuous and sensitive measure for capturing fine-grained motor dynamics. More broadly, they position digital musical interactions as a useful framework for future investigations of motor resonance across controlled and naturalistic settings, including musical performance.

## Statements and Declarations


### Funding

The authors would like to thank the Région Hauts-de-France for their financial support in the completion of this study.


### Competing interests

O. Capra serves as a co-editor of the special issue in which this manuscript is submitted. F. Berthaut declares no competing interests.

## Ethics approval

This study was approved by the ethics committee of the Université Catholique de Lille. All procedures were conducted in accordance with the ethical standards of the 1964 Declaration of Helsinki and its later amendments.

## Informed consent

Informed consent was obtained from all individual participants included in the study.

## Data availability

All stimuli, data, and analysis scripts are available on the Open Science Framework (OSF) at https://osf.io/yda5r/. The authors had full control of all primary data and agree to allow the journal to review the data if requested.

## Author contributions

Olivier Capra: Conceptualization, Methodology, Software, Formal analysis, Validation, Visualization, Writing—original draft, Writing—review & editing, Supervision, Funding acquisition, Project administration.

Florent Berthaut: Conceptualization, Methodology, Software, Writing—review & editing.

# References


Alaerts, K., Swinnen, S.P., Wenderoth, N. (2010). Observing how others lift light or heavy objects: which visual cues mediate the encoding of muscular force in the primary motor cortex? *Neuropsychologia*, *48*(7), 2082–2090,

Amoruso, L., Finisguerra, A., Urgesi, C. (2016). Tracking the time course of top-down contextual effects on motor responses during action comprehension. *Journal of neuroscience*, *36*(46), 11590–11600,

Aravena, P., Delevoye-Turrell, Y., Deprez, V., Cheylus, A., Paulignan, Y., Frak, V., Nazir, T. (2012). Grip force reveals the context sensitivity of language-induced motor activity during "action words" processing: Evidence from sentential negation. *PloS one*, *7*(12), e50287,

Avanzini, P., Fabbri-Destro, M., Dalla Volta, R., Daprati, E., Rizzolatti, G., Cantalupo, G. (2012). The dynamics of sensorimotor cortical oscillations during the observation of hand movements: an eeg study. *PloS one*, *7*(5), e37534,

Baldissera, F., Cavallari, P., Craighero, L., Fadiga, L. (2001). Modulation of spinal excitability during observation of hand actions in humans. *European Journal of Neuroscience*, *13*(1), 190–194,

Barsalou, L.W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical transactions of The Royal Society B: biological sciences*, *364*(1521), 1281–1289,

Blampain, J., Ott, L., Delevoye-Turrell, Y. (2018). Seeing action simulation as it unfolds: The implicit effects of action scenes on muscle contraction evidenced through the use of a grip-force sensor. *Neuropsychologia*, *114*, 231–242,

Borroni, P., Montagna, M., Cerri, G., Baldissera, F. (2005). Cyclic time course of motor excitability modulation during the observation of a cyclic hand movement. *Brain research*, *1065*(1-2), 115–124,

Caligiore, D., Pezzulo, G., Baldassarre, G., Bostan, A.C., Strick, P.L., Doya, K., ... others (2017). Consensus paper: towards a systems-level view of cerebellar function: the interplay between cerebellum, basal ganglia, and cortex. *The cerebellum*, *16*(1), 203–229,

Capra, O. (2020). *The experience of spectators of digital interactions: Benefits of visual augmentations and the role of attributed agency in electronic music performances* (Unpublished doctoral dissertation). Université de Lille.

Capra, O., Berthaut, F., Grisoni, L. (2020). Have a seat on stage: Restoring trust with spectator experience augmentation techniques. *Proceedings of the 2020 acm designing interactive systems conference* (pp. 695–707).

Cavallo, A., Becchio, C., Sartori, L., Bucchioni, G., Castiello, U. (2012). Grasping with tools: corticospinal excitability reflects observed hand movements. *Cerebral Cortex*, *22*(3), 710–716,

Choi, J.W., Kim, B.H., Huh, S., Jo, S. (2020). Observing actions through immersive virtual reality enhances motor imagery training. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *28*(7), 1614–1622,

Craighero, L. (2024). Motor resonance: neurophysiological origin, functional role, and contribution of the motivational, moral, and social aspects of action. *The routledge handbook of embodied cognition* (pp. 442–451). Routledge.

D'Ausilio, A., Novembre, G., Fadiga, L., Keller, P.E. (2015, March). What can music tell us about social interaction? *Trends in Cognitive Sciences*, *19*(3), 111–114, https://doi.org/10.1016/j.tics.2015.01.005

Decloe, R., & Obhi, S.S. (2013). Motor cortical processing is causally involved in object recognition. *BMC neuroscience*, *14*(1), 155,

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental brain research*, *91*(1), 176–180,

Fadiga, L., Fogassi, L., Pavesi, G., Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *Journal of neurophysiology*, *73*(6), 2608–2611,

Frak, V., Nazir, T., Goyette, M., Cohen, H., Jeannerod, M. (2010, March). Grip Force Is Part of the Semantic Representation of Manual Action Verbs. *PLoS ONE*, *5*(3), e9728, https://doi.org/10.1371/journal.pone.0009728

Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, *119*(2), 593–609,

Gangitano, M., Mottaghy, F.M., Pascual-Leone, A. (2001). Phase-specific modulation of cortical motor output during movement observation. *Neuroreport*, *12*(7), 1489–1492,

Gangitano, M., Mottaghy, F.M., Pascual-Leone, A. (2004). Modulation of premotor mirror neuron activity during observation of unpredictable grasping movements. *European Journal of Neuroscience*, *20*(8), 2193–2202,

Guidali, G., Picardi, M., Franca, M., Caronni, A., Bolognini, N. (2023). The social relevance and the temporal constraints of motor resonance in humans. *Scientific Reports*, *13*(1), 15933,

Huron, D. (2008). *Sweet anticipation: Music and the psychology of expectation*. MIT press.

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage*, *14*(1), S103–S109,

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.

Jensenius, A.R., & Wanderley, M.M. (2010). Musical gestures: Concepts and methods in research. *Musical gestures* (pp. 24–47). Routledge.

Kemmerer, D. (2021, October). What modulates the Mirror Neuron System during action observation?: Multiple factors involving the action, the actor, the observer, the relationship between actor and observer, and the context. *Progress in Neurobiology*, *205*, 102128, https://doi.org/10.1016/j.pneurobio.2021.102128

Kohler, E., Keysers, C., Umiltà, M.A., Fogassi, L., Gallese, V., Rizzolatti, G. (2002). Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons. *Science*, *297*(5582), 846–848,

Kornhuber, H.H. (1971). Motor functions of cerebellum and basal ganglia: the cerebellocortical saccadic (ballistic) clock, the cerebellonuclear hold regulator, and the basal ganglia ramp (voluntary speed smooth movement) generator. *Kybernetik*,

*8*(4), 157–162,

Kothe, C., Shirazi, S.Y., Stenner, T., Medine, D., Boulay, C., Grivich, M.I., ... Makeig, S. (2025). The lab streaming layer for synchronized multimodal recording. *Imaging Neuroscience*, *3*, IMAG.a.136, https://doi.org/10.1162/IMAG.a.136

Lepage, J.-F., Tremblay, S., Théoret, H. (2010). Early non-specific modulation of corticospinal excitability during action observation. *European Journal of Neuroscience*, *31*(5), 931–937,

Loporto, M., Holmes, P.S., Wright, D.J., McAllister, C.J. (2013). Reflecting on mirror mechanisms: motor resonance effects during action observation only present with low-intensity transcranial magnetic stimulation. *PloS one*, *8*(5), e64911,

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of eeg-and meg-data. *Journal of neuroscience methods*, *164*(1), 177–190,

Matamala-Gomez, M., Frisco, F., Guidali, G., Lega, C., Beacco, A., Bolognini, N., Maravita, A. (2025). Virtual body continuity during action observation affects motor cortical excitability. *Scientific reports*, *15*(1), 13364,

Michirev, A., Lindemann, O., Kühne, K., Fischer, M.H., Raab, M. (2024). Spontaneous grip force fluctuations mirror semantic numerical magnitude processing: Larger numbers elicit larger forces. *Acta Psychologica*, *249*, 104468,

Miklashevsky, A., Fischer, M., Lindemann, O. (2022, November). Spatial-numerical associations without a motor response? Grip force says 'Yes'. *Acta Psychologica*, *231*, 103791, https://doi.org/10.1016/j.actpsy.2022.103791

Miklashevsky, A., Lindemann, O., Fischer, M.H. (2021). The force of numbers: Investigating manual signatures of embodied number processing. *Frontiers in Human Neuroscience*, *14*, 590508,

Naish, K.R., Houston-Price, C., Bremner, A.J., Holmes, N.P. (2014). Effects of action observation on corticospinal excitability: Muscle specificity, direction, and timing of the mirror response. *Neuropsychologia*, *64*, 331–348,

Nazir, T.A., Hrycyk, L., Moreau, Q., Frak, V., Cheylus, A., Ott, L., ... Delevoye-Turrell, Y. (2017). A simple technique to study embodied language processes: the grip force sensor. *Behavior research methods*, *49*(1), 61–73,

Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences*, *97*(2), 913–918,

Novembre, G., Chavane, F., Ross, J. (2018). Grip force as a temporal proxy of cerebral activity. *NeuroImage*, *173*, 123–135,

Novembre, G., Pawar, V.M., Kilintari, M., Bufacchi, R.J., Guo, Y., Rothwell, J.C., Iannetti, G.D. (2019). The effect of salient stimuli on neural oscillations, isometric force, and their coupling. *Neuroimage*, *198*, 221–230,

Puckette, M. (1996). Pure data: another integrated computer music environment. *Proceedings of the second intercollege computer music concerts*, 37–41,

Puglisi, G., Leonetti, A., Landau, A., Fornia, L., Cerri, G., Borroni, P. (2017). The role of attention in human motor resonance. *PloS one*, *12*(5), e0177457,

Rangel, B.O., Novembre, G., Wessel, J.R. (2024). Measuring the nonselective effects of motor inhibition using isometric force recordings. *Behavior research methods*, *56*(5), 4486–4503,

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, *27*(1), 169–192,

Rouder, J.N., Speckman, P.L., Sun, D., Morey, R.D., Iverson, J.G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16*(2), 225–237, https://doi.org/10.3758/PBR.16.2.225

Stockner, M., Mazzoni, G., Perrin, D., Ianì, F. (2025). The same label, different processes: What lies behind the term "mental simulation" in the embodied cognition literature? *Brain and Cognition*, *188*, 106335,

Syrov, N., Yakovlev, L., Miroshnikov, A., Kaplan, A. (2023). Beyond passive observation: feedback anticipation and observation activate the mirror system in virtual

finger movement control via p300-bci. *Frontiers in Human Neuroscience*, *17*, 1180056,

Vallat, R. (2018). Pingouin: statistics in python. *Journal of Open Source Software*, *3*(31), 1026, https://doi.org/10.21105/joss.01026 Retrieved from https://doi.org/10.21105/joss.01026

van Schie, H.T., Koelewijn, T., Jensen, O., Oostenveld, R., Maris, E., Bekkering, H. (2008). Evidence for fast, low-level motor resonance to action observation: an meg study. *Social neuroscience*, *3*(3-4), 213–228,

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., SciPy 1.0 Contributors. (2020). Scipy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, *17*, 261–272, https://doi.org/10.1038/s41592-019-0686-2

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., ... Morey, R.D. (2018). Bayesian inference for psychology. part i: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57, https://doi.org/10.3758/s13423-017-1343-3

Wanderley, M.M. (2002). Non-sonic gestures in musical performance and their impact on audience perception. *Journal of New Music Research*, *31*, 19–29,

Yavuz, M., Vandeville, F., Mignard, C., Jayez, J., Nazir, T.A., Capra, O. (2026). *Grip-force modulation during action word reading: An RSVP paradigm with Bayesian temporal analysis.* Preprint. Research Square.

Zatorre, R.J., Chen, J.L., Penhune, V.B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nature reviews neuroscience*, *8*(7), 547–558,

# Appendix

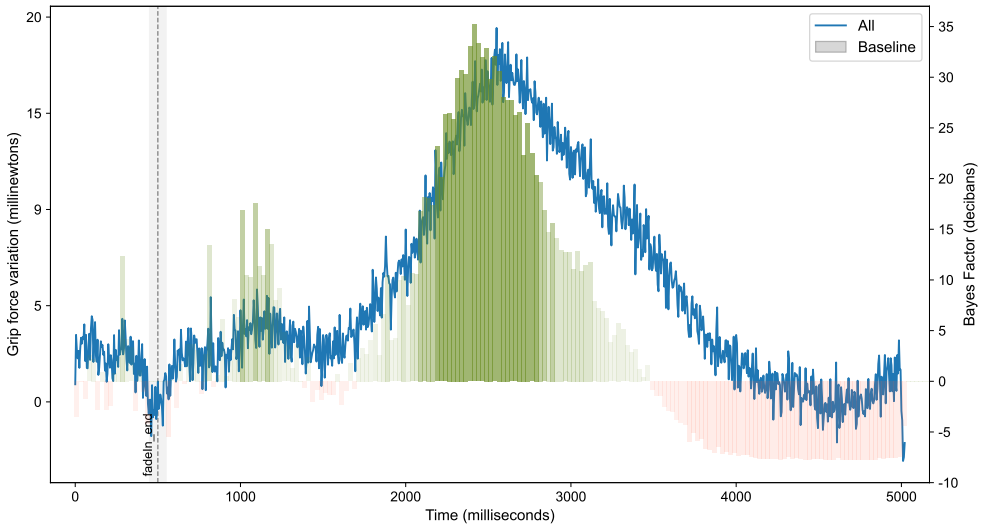



**Fig. 6**: Grand mean grip force signal across all conditions, repetitions, and participants (n = 42), displayed as a raw signal without filtering or temporal alignment, and baseline-corrected at *fadeIn_end* (dashed vertical line). The gray shaded area indicates the 100 ms baseline window. The TBED trace on the right axis shows pointwise Bayesian evidence for the baseline contrast in decibans; green bars indicate evidence for $H_1$ (difference from baseline), red bars indicate evidence for $H_0$ (no difference from baseline), and opacity scales with evidential strength. To reduce point-to-point noise in the unfiltered display, $dBF_{10}$ values are averaged within 25 ms bins

**Table 2**: Trial-by-trial questions. The table presents the original French items and their English translation. After each trial, four questions were drawn pseudo-randomly from this pool to document attributed agency and error detection; each statement was answered yes/no and weighted by an accompanying confidence rating (unmarked slider)

| **French** | **English** |
|---|---|
| Dans cette vidéo, c'est l'ordinateur qui contrôle le son. | In this video, the computer is the one controlling the sound. |
| Dans cette vidéo, je perçois bien comment le musicien contrôle le son. | In this video, I can clearly perceive how the musician controls the sound. |
| Dans cette vidéo, je peux voir quand le musicien intervient sur le son. | In this video, I can see when the musician acts on the sound. |
| Dans cette vidéo, il est facile de savoir si le musicien joue vraiment ou pas. | In this video, it is easy to tell whether the musician is actually playing or not. |
| Dans cette vidéo, il est évident que le musicien a modifié le son. | In this video, it is obvious that the musician has modified the sound. |
| Dans cette vidéo, le musicien bouge mais c'est l'ordinateur qui contrôle le son. | In this video, the musician moves, but it is the computer that controls the sound. |
| Dans cette vidéo, si le musicien ne faisait rien, il y aurait quand même des modifications sonores. | In this video, even if the musician did nothing, there would still be sound modifications. |
| Dans cette vidéo, si l'instrument ne fonctionnait pas bien, je m'en rendrais compte. | In this video, if the instrument were not working properly, I would notice. |
| Dans cette vidéo, si le musicien faisait une erreur, je m'en apercevrais. | In this video, if the musician made a mistake, I would realize it. |
| Dans cette vidéo, si le musicien se trompait, je m'en rendrais compte. | In this video, if the musician made a mistake, I would notice. |
| Dans cette vidéo, si l'instrument réagissait mal, je le remarquerais. | In this video, if the instrument responded poorly, I would notice. |
| Dans cette vidéo, si le musicien commettait une erreur, cela passerait inaperçu. | In this video, if the musician made a mistake, it would go unnoticed. |
| Dans cette vidéo, le musicien et l'ordinateur contrôlent le son autant l'un que l'autre. | In this video, the musician and the computer control the sound equally. |
| Dans cette vidéo, seul le musicien était en contrôle du son. | In this video, only the musician was in control of the sound. |
| Dans cette vidéo, je perçois bien ce qui, dans le son, est contrôlé par le musicien et par l'ordinateur. | In this video, I can clearly perceive what aspects of the sound are controlled by the musician and what aspects are controlled by the computer. |
| Dans cette vidéo, je distingue très bien ce qui, dans le son, est contrôlé automatiquement de ce qui est contrôlé par le musicien. | In this video, I can very clearly distinguish what in the sound is controlled automatically from what is controlled by the musician. |
| Dans cette vidéo, je peux voir quand l'ordinateur modifie le son. | In this video, I can see when the computer modifies the sound. |
| Selon vous, qui était le plus impliqué dans le contrôle du son dans la dernière vidéo ? | In your opinion, who was most involved in controlling the sound in the last video? |

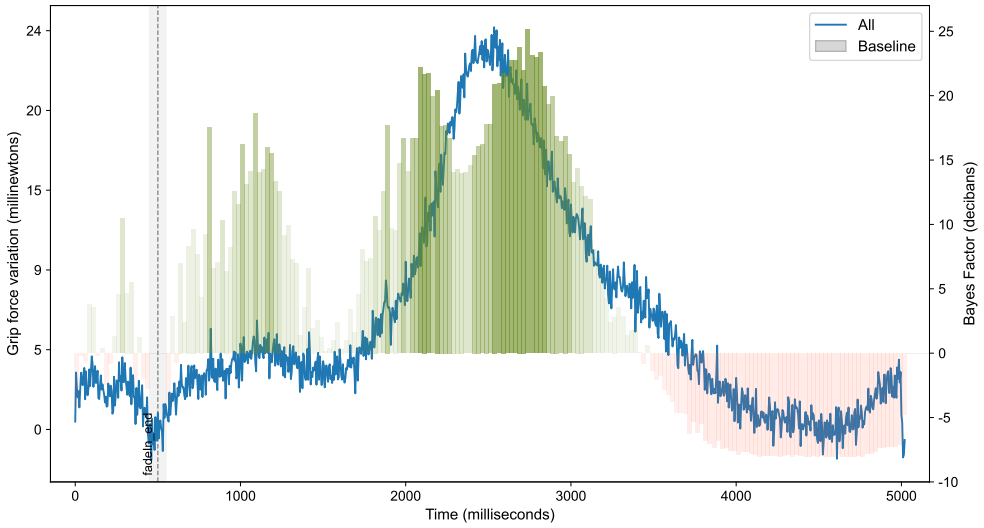



**Fig. 7**: Grand mean grip force signal across all conditions and repetitions in the full initial sample (n = 49), displayed as a raw signal without filtering or temporal alignment, prior to artifact rejection, and baseline-corrected at *fadeIn_end* (dashed vertical line). The gray shaded area indicates the 100 ms baseline window. The TBED trace on the right axis shows pointwise Bayesian evidence for the baseline contrast in decibans; green bars indicate evidence for $H_1$ (difference from baseline), red bars indicate evidence for $H_0$ (no difference from baseline), and opacity scales with evidential strength. To reduce point-to-point noise in the unfiltered display, $dBF_{10}$ values are averaged within 25 ms bins

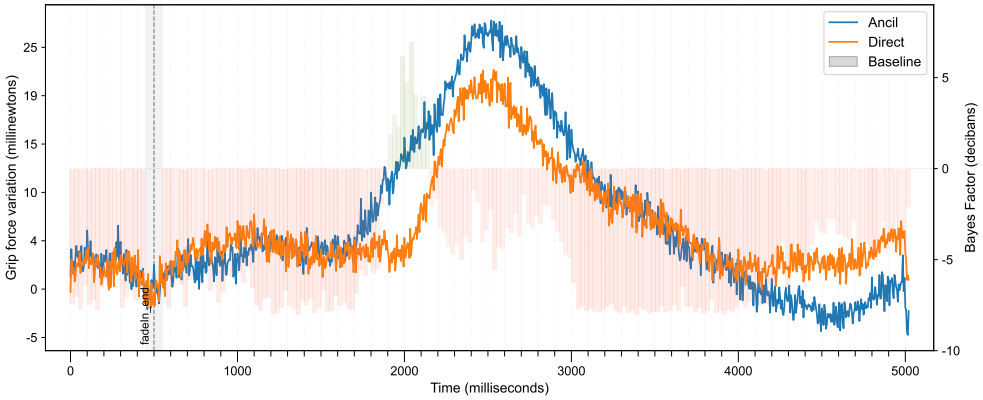



**Fig. 8**: Mean grip force signal for the *Ancil* (blue) and *Direct* (orange) conditions in the full initial sample (n = 49), displayed as a raw signal without filtering or temporal alignment, prior to artifact rejection, and baseline-corrected at *fadeIn_end* (dashed vertical line). The gray shaded area indicates the 100 ms baseline window. The TBED trace on the right axis shows pointwise Bayesian evidence for the condition contrast in decibans; green bars indicate evidence for $H_1$ (difference between conditions), red bars indicate evidence for $H_0$ (no difference between conditions), and opacity scales with evidential strength. To reduce point-to-point noise in the unfiltered display, $dBF_{10}$ values are averaged within 25 ms bins

**Table 3**: Durations (in milliseconds) between key timepoints for the *Ancil* and *Direct* base video recordings. The Mean column corresponds to the target duration used to time-normalize each phase. This procedure aligned homologous phase transitions across conditions while preserving within-phase signal dynamics. Raw, non-time-aligned signal plots are provided in Appendix Figs. 6, 7, and 8 as robustness checks

| **Interval** | **Ancil** | **Direct** | **Mean** |
|---|---|---|---|
| stim_start → fadeIn_end | 500 | 500 | 500 |
| fadeIn_end → mov_start | 850 | 1167 | 1008.5 |
| mov_start → trigger_start | 633 | 233 | 433 |
| trigger_start → trigger_end | 650 | 817 | 733.5 |
| trigger_end → mov_end | 250 | 300 | 275 |
| mov_end → fadeOut_start | 1634 | 1500 | 1567 |
| fadeOut_start → stim_end | 500 | 500 | 500 |

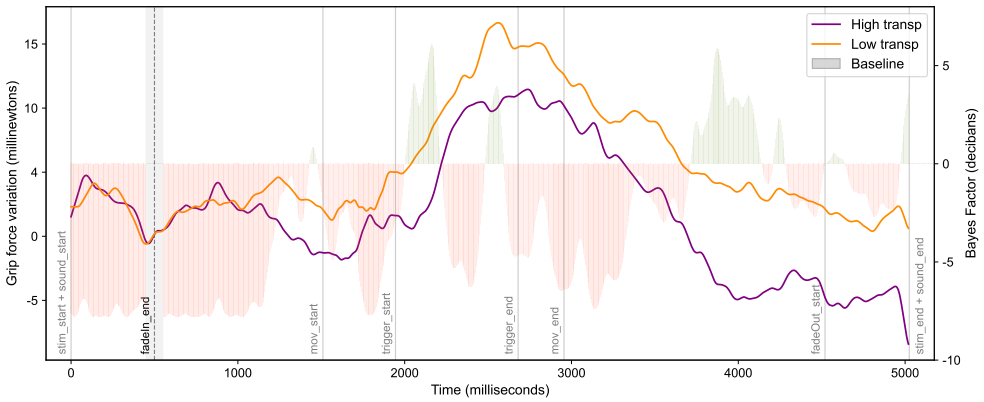



**Fig. 9**: Mean grip force signal for the high-transparency (purple) and low-transparency (orange) conditions (n = 42), low-pass filtered and smoothed, temporally aligned, and baseline-corrected at *fadeIn_end* (dashed vertical line). The high-transparency condition featured a salient and synchronous sound modification, whereas the low-transparency condition produced a barely perceptible and temporally offset change. The gray shaded area indicates the 100 ms baseline window. The TBED trace on the right axis shows pointwise Bayesian evidence for the transparency contrast in decibans; green bars indicate evidence for $H_1$ (difference between transparency conditions), red bars indicate evidence for $H_0$ (no difference between transparency conditions), and opacity scales with evidential strength. This comparison provides an indirect control for the contribution of auditory salience to the grip force profile. Notably, the lower-salience condition produced comparable or higher GF amplitudes than the higher-salience condition, inconsistent with auditory salience as a primary driver of the observed modulations

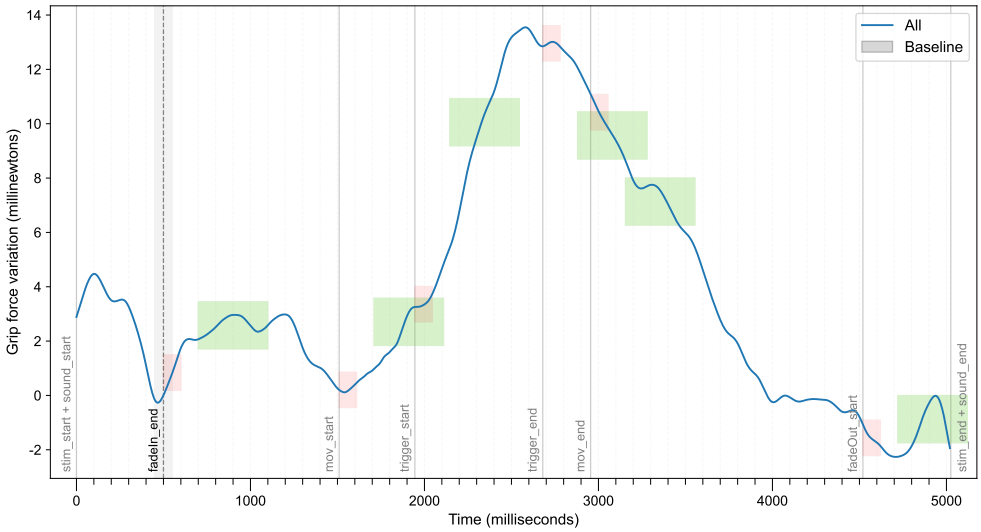



**Fig. 10**: Grand mean grip force signal (n = 42), low-pass filtered and smoothed, temporally aligned, and baseline-corrected at *fadeIn_end* (dashed vertical line). Red rectangles correspond to 0–100 ms windows after key phase transitions; green rectangles highlight 200–400 ms windows. These windows are overlaid to explore temporal patterns described in the motor resonance literature on action observation: the early non-specific phase (red) and the later context-specific phase (green)

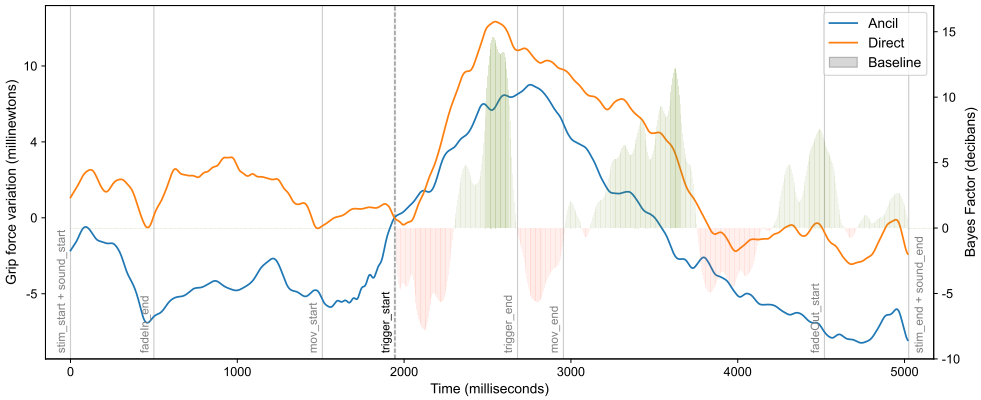



**Fig. 11**: Full time course of the mean grip force signal for the *Ancil* (blue) and *Direct* (orange) conditions (n = 42), low-pass filtered and smoothed, temporally aligned, and rebaselined at *trigger_start*. Vertical lines indicate key phase transitions. The gray shaded area indicates the 100 ms baseline window. The TBED trace on the right axis shows pointwise Bayesian evidence for the condition contrast in decibans; green bars indicate evidence for $H_1$ (difference between conditions), red bars indicate evidence for $H_0$ (no difference between conditions), and opacity scales with evidential strength. The zoomed view of the trigger window is shown in Fig. 4

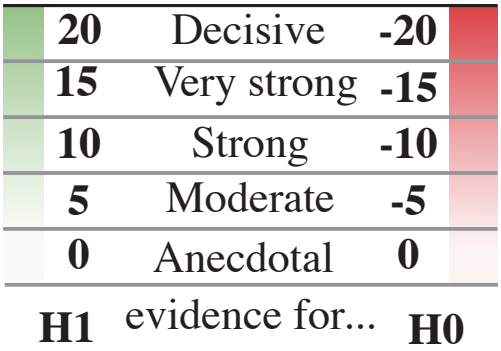



**Fig. 12**: Interpretation scale for deciban values ($dBF_{10}$) used in the TBED, following Jeffreys (1961). Bar color (green: evidence for $H_1$; red: evidence for $H_0$) and opacity reflect evidential strength, consistent with the display convention used in all TBED figures

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1.pdf
- ESM2.pdf
- ESM3.pdf