2025 IEEE World Haptics Conference (WHC)
Suwon Convention Center, Suwon, Korea
July 8 ~ 11, 2025

Watch, Feel, Do: Investigating the Impact of Naturalistic Vibrotactile Feedback and Active Imitation on Manual Skill Learning

Ravali Gourishetti, Autumn G. Hughes, Bernard Javot, and Katherine J. Kuchenbecker

I. Motivation

This study aims to investigate how well novice participants can learn manual skills by experiencing demonstrations recorded from expert teachers; each recording includes video, audio (interaction sound and spoken instructions), and naturalistic vibrotactile feedback [1]. Our specific goal is to understand how feeling this type of haptic feedback and imitating the demonstrated motion each affect manual skill learning. The results from this study will help us improve the tested technology and advance scientific understanding of human sensory-motor learning.

II. Methods

We conducted a user study to investigate learning of manual skills from recorded expert demonstrations. It was structured into three overall stages: an initial briefing, four experimental sessions, and a conclusion. The experimental sessions each comprised a demonstration and an evaluation phase from one of four fields: hair cutting, block printing, stone and wood sculpting, and violin playing.

At the start of the study, the experimenter gave a brief overview and instructed the participant to turn off all electronic devices to minimize distractions. Participants then measured their height and weight using provided tools, and these data, along with age and gender, were entered into a wristband device to monitor heart rate throughout the study. Participants put on the wristband (Polar Vantage V2) and an accompanying chest strap (Polar H10), followed by a calibration procedure. They then did the calibration steps for a screen-based eye-tracker (Tobii Pro) attached to the instrumented computer monitor used to display the study videos. A short questionnaire captured their demographics, handedness, and experience with manual tasks.

The study's four experimental conditions are IH: Imitation with haptic feedback, IN: Imitation with no haptic feedback, PH: Passive watching with haptic feedback, and PN: Passive watching with no haptic feedback. A preliminary demonstration from the field of metalworking was shown to expose participants to all four experimental conditions. Each participant learned all four manual skills, one in each experimental condition, with condition assignment randomized.

During the demonstration phase of each session, the participant experienced a video demonstration of an expert performing manual tasks; the supplementary video shows part of the violin demonstration. In sessions involving vibrotactile feedback, they held one or two stylus-shaped tools equipped with voice-coil actuators to deliver the recorded vibrotactile feedback, along with accelerometers to verify that delivery. Standard video recordings captured the participant's hand movements to check whether they followed the instructed condition (imitation of the expert's movements vs. passive watching). After each demonstration, participants completed a short questionnaire evaluating their subjective experience.

In the evaluation phase that followed each demonstration phase, participants performed a task similar to the task they had just learned about. These performances were recorded using a head-mounted camera (first-person perspective), an external camera (third-person view), a directional microphone for audio capture, and tool-mounted accelerometers to record task vibrations. Upon completion of each task, the participant filled out a questionnaire reflecting on their performance. Short breaks were offered between sessions to reduce fatigue. After completing all four sessions, participants filled out a final questionnaire and had a debriefing conversation with the experimenter.

III. Preliminary Results

We collected a range of quantitative and qualitative data, including questionnaires, heart rate, eye tracking, and expert evaluations of task performance. Below, we present a few observations from this dataset, only for the violin task.

Participant Background and Demographics: The study included 32 participants, comprising 20 female and 12 male, with a mean age of 35.5 years (SD = 7.8). Participants had diverse academic backgrounds, with the majority being students or early-career professionals. 28 were right-handed, two were left-handed, and the other two were ambidextrous. Five participants reported prior experience playing the violin, while others had little to no experience.

Post-Demonstration Questionnaire Responses: Figure 1 summarizes participant performance on the three evaluative questions presented after the violin demonstration; these questions were designed to assess the participant's understanding of the presented techniques. The IH group showed the highest accuracy (88%) on the first question, with the other three groups performing somewhat worse ($\leq$62.5%). For the second and third questions, all groups performed relatively well ($\geq$70%). These results show that the majority of participants understood the demonstration, and they hint that the IH group may have had an advantage for learning certain aspects of skills, though this apparent trend needs to be tested through statistical analysis across all four skills.

All authors are with the Haptic Intelligence Department, Max Planck Institute for Intelligent Systems, Stuttgart, Germany. E-mail: gourishetti@is.mpg.de, kjk@is.mpg.de

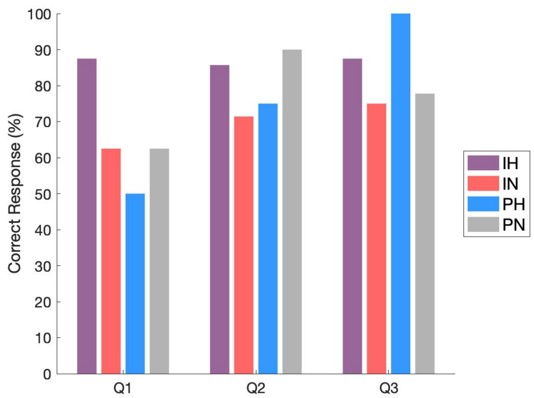

Fig. 1. Percentage of correct responses to the three post-demonstration questions for violin playing across the four experimental conditions. Image-based Q1 assessed the recognition of correct bow-holding technique. Video-based Q2 evaluated whether participants identified the appropriate pressure applied to the bow. Video-based Q3 asked participants to select the proper interpretation of the video's content, with options such as the incorrect amount of pressure applied on the bow. Bars represent the proportion of correct responses per condition for each question, and the color of each bar corresponds to the participant group.

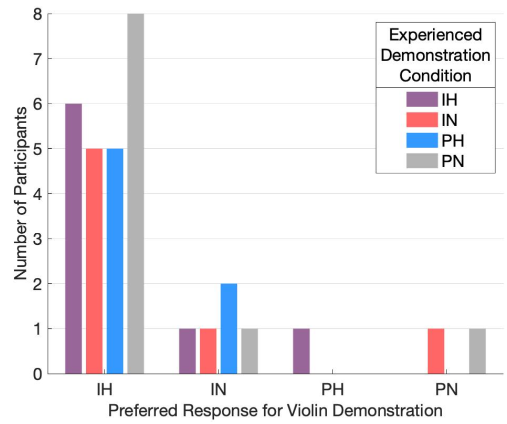

Fig. 2. Participants' preferred condition for experiencing the violin demonstration. The x-axis indicates the condition each participant preferred to experience, while the color of each bar shows the condition in which the participant experienced the violin demonstration.

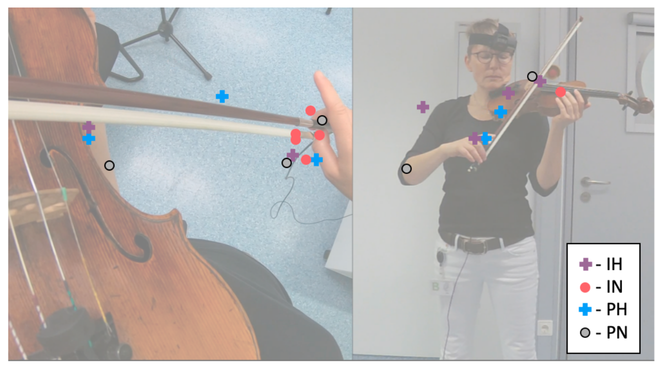

Fig. 3. A representative gaze overlay frame of all the participants during the demonstration phase. Participants predominantly fixated on the expert's hand and bow, especially in the first-person view.

Post-Task Reflections and Final Questionnaire Responses: Self-assessment responses indicated that most participants had moderate to high confidence that they achieved the task goals, held the tools correctly, and replicated the expert's movements properly. In the final questionnaire, participants were asked to indicate their preferred condition for experiencing the violin demonstration, as shown in Figure 1. Of 32 participants, 29 (90.6%) reported that the most effective approach involves motion imitation (IH or IN), with 24 (75%) specifically finding imitation with vibrotactile feedback (IH) to be most helpful (see Figure 2). These findings align with prior research highlighting the benefits of imitative learning over verbal instruction for motor-skill acquisition [2], as well as appreciation for naturalistic vibrotactile feedback when observing physical interactions [1].

Eye-Tracking Data: Eye-tracking data reveals participants' visual attention during the demonstration phase. A representative frame from the gaze recordings (Figure 3) shows that almost all participants focused on the expert's hands and the violin across all conditions. This pattern indicates a high level of engagement with the task. Additionally, participants generally looked more at the first-person perspective, suggesting that this view may facilitate a more intuitive and effective learning experience.

IV. Conclusion and Future Work

We presented an investigation into how different conditions of recorded expert demonstrations impact manual skill acquisition. Our preliminary findings on learning to play violin indicate that participants tended to focus visually on hand positioning and tool manipulation, particularly in the first-person videos. This finding suggests that the spatial alignment offered by first-person views may enhance the understanding and imitation of fine-motor skills. Additionally, most participants reported that they would prefer to experience the violin demonstration with motion imitation and specifically imitation with haptic feedback.

In future work, we aim to perform a comprehensive statistical analysis (self-assessment and expert evaluations) of task performance across the four experimental conditions to investigate the correlations between the conditions and the learning outcomes. Also, we plan to explore physiological engagement and stress level through heart rate variability. Ultimately, this research aims to contribute to the design of effective, scalable learning systems that leverage rich, multi-sensory demonstrations to enhance manual skill acquisition in both educational and professional contexts.

References

[1] Y. Gong, B. Javot, A. P. R. Lauer, O. Sawodny, and K. J. Kuchenbecker, "Naturalistic vibrotactile feedback could facilitate telerobotic assembly on construction sites," in *Proceedings of the IEEE World Haptics Conference (WHC)*, Delft, The Netherlands, Jul. 2023, pp. 169–175.

[2] E. De Stefani, F. Rodà, E. Volta, V. Pincolini, A. Farnese, S. Rossetti, F. Pedretti, and P. F. Ferrari, "Learning new sport actions: Pilot study to investigate the imitative and the verbal instructive teaching methods in motor education," *PLOS One*, vol. 15, no. 8, p. e0237697, 2020.