Check for updates

# Be sensei, my friend: Aikido training with a remotely controlled proxy trainer

Takashige Suzuki[1]\*, Alarith Uhde[2], Takuto Nakamura[1],
Takuji Narumi[1], Tomohiro Amemiya[3] and Hideaki Kuzuoka[1]

[1]Kuzuoka Narumi Lab, Graduate School of Information Science and Technology, The University of Tokyo,
Tokyo, Japan, [2]Tokyo College, The University of Tokyo, Tokyo, Japan, [3]Information Technology Center,
The University of Tokyo, Tokyo, Japan

Contact sports such as Aikido are preferred to be trained in person with an experienced trainer, as the attempts of remote training before and during the COVID-19 pandemic failed to reproduce the quality of the in-person training benefiting from the trainer's physically present body. To address this issue of replicating in-person experience remotely, we proposed "Sensei possession," an xReality-based training method for contact sports in which the trainer remotely guides a person who performs the physical interaction with the trainee on behalf of the real trainer in real-time. In this study, to test the effectiveness of "Sensei possession" on training performance and examine the concerns about its possible side effects on motivation, we conducted a between-participants experiment with an Aikido training task [N = 10 pairs (20 people)]. We compared the groups with or without live feedback on the proxy trainers under our hypothesis that live feedback would enhance trainees' performance gain. As a result, the trainees in the group with live feedback on proxy trainers had more performance gain than those without live feedback, and no indication of negative effects on motivation was shown. We discussed our results concerning the previous research on micro-adaptive training and reported technical insights to improve the design of Sensei possession further.

KEYWORDS

sensei possession, Aikido, contact sports, remote training, xReality, in-person training, adaptive training, motivation

## 1 Introduction

Every year, hundreds of Aikido enthusiasts go on a journey to Aikikai Hombu Dojo, the headquarters of the world's largest Aikido organization (Kohn, 2011). They want to learn the art directly from the grandmasters to improve their skills, better understand the philosophy, and get a realistic assessment of their current level in Aikido, which helps them build their motivation for further practice (Stepanyuk et al., 2022).

In Aikido and other contact sports, such "in-person training" is crucial, and learning from the grandmasters themselves is a valuable experience (Figueiredo et al., 2020). In-person training allows the trainer to demonstrate how to establish contact, how to orient the body, and also to communicate the philosophy (Ikuta, 2000; Afana et al., 2023). Furthermore, by meeting in person, the trainer can adjust the level of difficulty for the trainee in real-time to ensure efficiency and safety (Martin-Niedecken et al., 2019; Lu et al., 2024).

However, such in-person training has an inherently limited capacity, especially if it exceeds short, 1-day visits. The 31 instructors at the Tokyo Hombu Dojo can only offer such classes to a few dozen participants per year. These classes currently require continuous in-person attendance, which implies financial and other challenges for people who cannot easily visit Tokyo, who relocate to other cities after some time, or who have an occupation that requires them to travel frequently (Burazerović, 2011; Aikikai Foundation, 2014; Aikikai Foundation, 2023). Moreover, the COVID-19 pandemic caused the dojo to cancel the in-person classes repeatedly. Other dojos faced similar problems (Aikido Yoshinkan, 2020). As a result, Aikido practitioners had to adopt online training using video calls during this period, reducing motivation and participation because the video calls could not provide the aforementioned valuable experience in the in-person training with grandmasters (Meyer et al., 2021).

To address this problem, we suggest a technical solution that tries to replicate such in-person experiences for Aikido and other contact sports in a highly immersive way. It is not supposed to fully replace in-person training with grandmasters, but it makes them more affordable and scalable. Specifically, we use an approach we call "Sensei possession," based on xReality [XR; which represents all "new" reality format including Virtual and Augmented Reality (VR and AR) (Rauschnabel et al., 2022)]. In "Sensei possession," a remote, experienced trainer in the dojo (the "sensei" or grandmaster) demonstrates how to establish contact and how to position the body, while being recorded. At the trainee's location, a remote "proxy trainer" (who can be a novice) with an XR headset views this recording and learns for a few minutes to imitate the movements of a virtual avatar along with audio commands. After the learning phase, the proxy trainer removes the XR headset and becomes "possessed," by receiving instructions from the sensei through wireless headphones. Finally, the trainee can practice together with the proxy trainer "as if" the grandmaster were seemingly present. An envisioned example use case for this setup is for some local club members to start learning a foreign contact sport, while receiving first-hand instructions from the grandmaster. Based on the popularity of "in-person training" with grandmasters in contact sports, we hypothesized that even if the performance of a possessed proxy trainer was imperfect compared to the sensei, trainees would be able to practice contact sports more effectively in in-person training with the possessed proxy trainer, than in video-based remote training with a sensei without the possession system.

To test our hypothesis, we first developed this Sensei possession system based on insights from the literature on remote training media and a pilot workshop with domain experts (described in Section 3.1). Then, we ran an experimental study with Aikido training and inexperienced participants to evaluate the effectiveness. This paper contributes details on the implementation of "Sensei possession", and also confirms its effectiveness on performance and motivation. Finally, our discussion covers our result's relation with previous research on micro-adaptive training and intrinsic motivation, as well as the design implications for more complex training situations.

# 2 Related work

## 2.1 XR sports training increases efficiency but May overwhelm users

Applications of XR technologies in remote training include not only the transmission of three-dimensional information from place to place but also the enhancement of training efficiency through information presentation methods that were not possible in reality. The most basic example is the placement of AR objects and signs in mid-air, which trainers can utilize to provide intuitive instructions of three-dimensional movements (Hou et al., 2013). Similarly, the visual presentation of instructed body movements from a first-person perspective can increase the intuitiveness. Yang and Kim (2002) introduced an interaction method called Just Follow Me (JFM) in which a trainee can observe the motion of a trainer using the ghost metaphor, a semi-transparent avatar displayed on top of the trainee's avatar in a VR environment, and found that JFM can improve the training efficiency compared to real-world teaching methods. Furthermore, Fribourg et al. (2021) introduced virtual co-embodiment where the movements of a trainer and a trainee are averaged into a single avatar, through which the trainer can intuitively express their intention of movement to the trainee, who can then acquire the motor skills more efficiently compared to JFM (Kodama et al., 2023; Takita et al., 2023).

However, the visual cues used in the methods above are known to be overwhelming for an inexperienced trainee, especially "when multiple body parts were involved in motion simultaneously" (Wang et al., 2024). As our goal is to realize the training of whole-body motor skills in contact sports, which is more complex than tasks implemented in the previous studies, we must investigate how much information should be presented when applying the above methods in the training. Meanwhile, we also assume that such investigations for adjustment will be necessary in each case of converting in-person training of specific skills to XR training, which will inevitably burden the trainers and the trainees.

## 2.2 Capability of "sensei possession" must be tested

A simple solution to lower the barriers to implementation is to minimize the experienced difference from conventional face-to-face training methods. Thus, we proposed "Sensei possession," a method to realize the remote learning of motor skills by replicating the face-to-face training environment over remote locations, with a proxy trainer whose movement can be controlled remotely by the real trainer (Suzuki et al., 2022). Similarly, Faridan et al. (2023) utilized ChameleonMask, a method introduced by Misawa and Rekimoto (2015) to display a person's motion and facial expression through another person, to let a teacher in a remote place teach a group of students through a proxy teacher, who wore an XR headset to observe and imitate the movements of the real teacher's hands.

Such proposals have the advantage of being able to reproduce remotely the same training that was done in person, which supports the preservation of tradition and lowers the implementation barriers. However, none have investigated how the training efficiency is influenced by the real-time control system, which

might not function when the proxy trainer has to follow the complex instructions of whole-body movements. Therefore, we tried to figure out a possible means to realize Sensei possession in this study according to the opinions collected from experienced trainers.

## 2.3 Possible side-effects on motivation may negatively affect sports training

The complex skills in the whole-body movements of contact sports contain "tacit knowledge" which cannot be expressed or even recognized directly (Polanyi, 1966). One of the means that allowed the transfer of such unrecognizable knowledge is "situated learning," in which a trainee's surroundings play an important role in the acquisition of skills (Lave and Wenger, 1991). Ikuta (2000) has found that in the traditional training community of whole body motor skills, a trainee must dwell in that community and its surrounding environment to be able to master the skills shown by the trainer because the techniques are to be controlled dynamically in association with the environment, based on the philosophy cultivated in the community. Moreover, the trainees' understanding of metaphors in the communication with their trainers, which helps the acquisition of tacit knowledge by inducing imagination (Petrie and Oshlag, 1993), is also heavily dependant on the contextual experience gained within the community (Ikuta, 2008). Thus, preserving the traditional training environment is important for the members to learn the skills at a high level, while it is not yet known which elements in the training environment are critical for skill acquisition.

In addition to the training environment, building motivation towards the training is also essential. Hayashibe and Amenomiya (2008) reported from the interviews on Japanese traditional artists, that "the strong will to learn and to endure years of training" is vital for the succession of complex skills. Despite that, the study on remote learning (Serhan, 2020) indicated that video meeting systems negatively affect the motivation to learn compared to face-to-face instructions, with interviewees' opinions that the remote interaction did not satisfy them in the same way that watching a sports game on TV is different from experiencing it on site. On the other hand, the report of Cheng et al. (2015), in which participants enjoyed the experience of being manipulated by remote instructions to enrich the experience of other participants, suggests the possibility of the "Sensei Possession" system overcoming the defect of motivation in remote learning. Consequently, we collected motivation-related data from the participants in our study to evaluate our proposed system as an assistive medium for preserving training activities.

# 3 Materials and methods

## 3.1 Pilot workshop

In preparation for developing Sensei Possession, we contacted 3 Aikido organizations and found that none of them had used XR communication systems for remote training before. Thus, we organized a first workshop with domain experts, to discuss with them how an XR-based system could be useful in remote training and to gather first insights about how to design it. From a technological point of view, we were focused on elements of in-person training the experts consider important, but problematic during conventional remote education as they had experienced during the COVID-19 pandemic. Thus, the workshop helped us narrow down the design space and clarify what factors to focus on in our design and evaluation. The workshop was approved by the ethics committee of the Graduate School of Information Science and Technology, the University of Tokyo (UT-IST-RE-221005).

We invited two expert trainers with different backgrounds. One trainer had 20 years of experience teaching Aikido, and the other had 4 years of experience teaching volleyball. We selected these two experts to collect Aikido-specific insights on the one hand, but also to develop a more generalizable understanding of how such a system could be useful in a broader spectrum of sports that involve dynamic body movements. We guided the invited trainers through two sessions, in which we first introduced 5 state-of-the-art remote training and XR technologies, such as XR sports training and wearable haptic displays. The trainers could explore these technologies with remote Aikido training simulations, in each of which the expert experienced a set of XR technologies for transferring movements to another participant as references. After the practice phase, we gave them an open-ended brief questionnaire about preferences, usability, other necessities, and possible concerns. In the end, we ran a semi-structured interview for around 20 min to provide more details about their answer to the questionnaire.

From the analysis of the opinions collected from 4 male participants aged from 22 to 39 (M = 27.0, SD = 8.0) in the 2-h workshop, we found the problems of conducting the whole-body motor skill training remotely in the conventional methods as follows: In the questionnaire, the difficulty of using XR headset during long-term training was mentioned, that two participants commented "it is difficult to look at my own body" due to the XR headset's limited vertical field of view. They also said, "By looking down, the neck started to ache." In the interview, the expressivity of the wearable haptic devices and the XR headset was referred to as insufficient, that the device "fails to reproduce the sense of physical contact during in-person training" and "visually transferring the whole-body movement is not appropriate." Finally, the concern about the influence of remote training on motivations was mentioned, that "the relationship with the remote person may not go well like in-person" although "trainees' motivation to learn the skill and the personality of their trainer is important".

Based on these findings, we revised our initial design of Sensei Possession. The opinions about the overwhelming complexity of understanding the instructed whole-body movement confirmed that developing a shared ground/context through the instructions is necessary to transfer movements efficiently. This is especially important if an inexperienced person serves as a proxy trainer, and a preliminary learning period would be necessary to be able to understand the commands of the real trainer. In addition, the trainer suggested in the interview that the roles of proxy trainer and trainee should be swapped during the training, "to replicate the traditional training environment with Senpai (trainee who started training earlier) and Kohai (trainee who started training later) because experiencing both teaching and being taught helps participants acknowledge their ability more precisely."
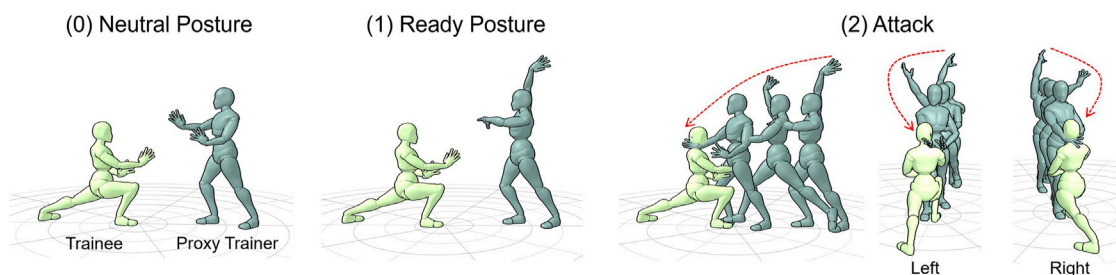
**FIGURE 1**
Proxy trainers start from the neutral posture, shift to the ready posture, perform one of the two attacks, and return to the neutral posture. Created with https://justsketch.me.
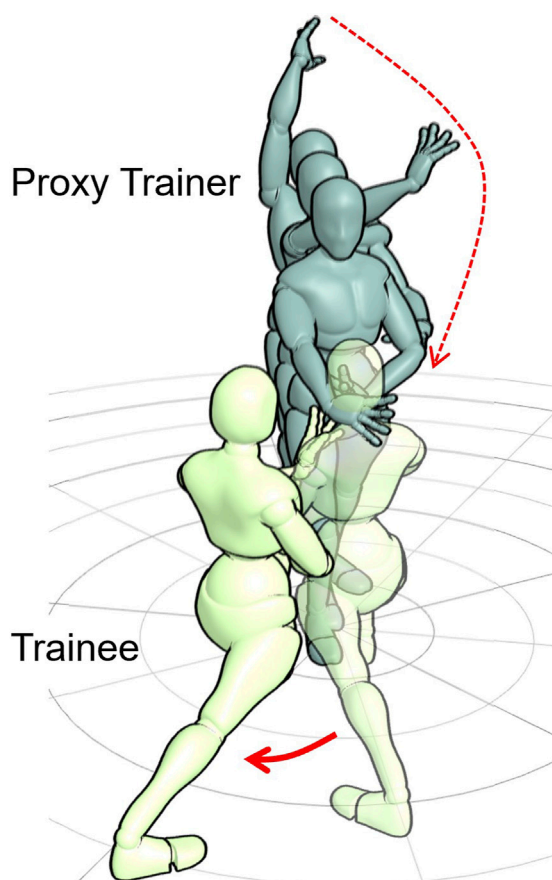


**FIGURE 2**
Trainees start from the lower neutral posture, evade the attack, and return to the neutral posture. Created with https://justsketch.me.

## 3.2 Main experiment

### 3.2.1 Objectives

In our experiment, we investigated how Sensei possession influences training efficiency and the subjective experiences of trainees and proxy trainers. We focused on the real-time-ness of Sensei possession because it is the most demanding element in the implementation and for the participants, and also because we expected it to be the key part of partner-centered contact sports

training. As we conducted the experiment under COVID-19 safety measures, we designed a partner-centered, non-contact training task under the supervision of an Aikido Club at Nihon University. Our goal was to determine whether the constellation with a proxy trainer, whose movements were controlled in real-time by a real trainer (*Possessed condition*), would improve the learning speed of a trainee in comparison with the partner-centered training with a proxy trainer without real-time control *(Unpossessed condition)*. Additionally, we aimed to examine the effect of real-time possession on participants' motivation, the overwhelmingness of which was raised as a concern in our opinion survey. This experiment was approved by the ethics committee of the Graduate School of Information Science and Technology, the University of Tokyo (UT-IST-RE-221201-1).

### 3.2.2 Method
#### 3.2.2.1 Participants

We recruited 20 participants [16 males and 4 females, aged 21 to 45 (M = 25.1, SD = 6.7)] who were healthy enough to exercise for about 15 min. All participants also agreed to practice Aikido with an unfamiliar training partner at a close distance. Two of the participants had experience with Aikido, but we assumed they were equivalent to the inexperienced trainees in our training tasks because they had only trained in different styles of Aikido less than 10 times. 9 other participants had experience in other martial arts, including Judo, Karate, Kendo, and Tai Chi Chuan.

We grouped participants into 10 pairs of the same gender and assigned 5 pairs to each condition. All assignments were made randomly while the distribution of different genders was leveled between 2 conditions [Possessed: aged from 21 to 45 (M = 24.8, SD = 6.8), Unpossessed: aged from 21 to 43 (M = 25.4, SD = 6.2)].

#### 3.2.2.2 Training task

The Aikido Club that supervised our task design belongs to a martial art style called Takeda Ryu Nakamura Ha, which is famous for its competitive style in which the student members can participate in national competitions held twice a year. The training task that we used in our experiment was a part of the fundamental training to prepare for such competitions, where a trainee must learn to evade an incoming attack at the right time and in the right position to later be able to catch the attack and fight back. We chose this task because it is one of the most important training tasks, and because it draws from interpersonal bodily interactions

that are hard to replicate via a video call. Thus, it is one of the tasks that could not be trained satisfactorily during quarantine, owing to its need for a skilled training partner.

On the side of the (proxy) trainer, the attack is performed as a sequence of 3 movements (Figure 1): (Movement 1) From the neutral posture, raise the right hand to be ready for attack and stick out the left hand as a defense from the incoming opponent. (Movement 2) Strike the partner's head diagonally from either the right or left temple while simultaneously stepping towards the partner. (Movement 3) Step back and make the neutral posture.

From the fighting perspective, the attack is traditionally performed by a trainer or a trainee with more experience, who can adjust the difficulty of evasion by changing the speed of the attack. During a competition, as the quality of attack is also subject to evaluation, all movements must be executed in the correct postures to give the opponent a fair chance to evade.

On the side of the trainees, the evasion is performed accordingly with the attack. They take 2 steps forward diagonally, away from the striking hand but still facing towards the center of the partner's body (Figure 2). From the fighting perspective, the key difficulty of this evasive motion is the time to start moving, which should neither be too early nor too late. When one starts evading before the striking motion starts, the opponent may decide to attack from the direction that one has escaped to, while one will be struck if evasion is not executed early enough. During the competition, the trainee must try to perform the evasion in the correct way because the postures during the evasion are also to be judged.

### 3.2.2.3 Procedures

In our experiment, we simulated the remote training of Aikido with Sensei possession. To do this, we artificially prepared the participants for Sensei possession and the remote training situation.

As for the participants, the first author played the role of the "real trainer," while two participants played the roles of "proxy trainer" and "trainee" in turn. We recruited the first author as the real trainer as he had trained in Aikido for more than 8 years and had experience teaching inexperienced trainees. Before the experiment, the first author trained additionally for a year with the members of the supervising Aikido club and learned how to evaluate the trainees' performance.

As for simulating the remote training situation, we used a simple experiment room and divided it into two zones. These two zones represented a zone for the proxy trainer and the trainee (participants' zone), and a remote zone for the real trainer. We experimented in a closed room and kept the participants' positions constant during the training using marks on the floor, because of safety protocols under COVID-19 and for reducing the risks of participants' injury during the training. We placed a partition between the 2 zones so that the real trainer (the first author) could hide from the participants and communicate with them via a monitor placed in the participants' zone. With this setup, the participants could only get feedback from the real trainer indirectly, as if they were in a place away from the real trainer.

With the above simulated remote training environment, our experiment consisted of 2 training phases with 3 phases of subjective data collection before, between, and after the training phases (See Figure 3).

In the first phase, participants regardless of their condition performed the 3-min warm-up exercise and answered the baseline questionnaire (the details are later described in Paragraph 3.2.2.5.2).

Next, in the 1st training phase, the real trainer asked each participant to play the role of the proxy trainer or the trainee. Then, the participant went under the preliminary learning step (described later in the next paragraph). After the participants learned what movement they should learn, the participants performed the attack and the evasion for 30 s by themselves, which we recorded as their baseline performance with a video camera per participant. For this recording, the real trainer asked the participants to perform the right and left types of attack and evasion alternately, while focusing on maximizing their performance. The next step was another preliminary learning step for the proxy trainer, followed by the interpersonal training step with the Sensei possession system comprising 5 sessions (also described later). At the end of this 1st training phase, participants again performed the attack and the evasion for 30 s by themselves, which we recorded as their final performance. For this final recording, the real trainer asked the participants again to perform right and left alternately with their best performance.

In the next phase, participants answered the phase 1 questionnaire. After that was the 2nd training phase, in which we asked the participants to play the other role and undergo the same procedures as in the 1st training phase. We recorded the baseline and final performance of the participants again in this training phase but with different roles.
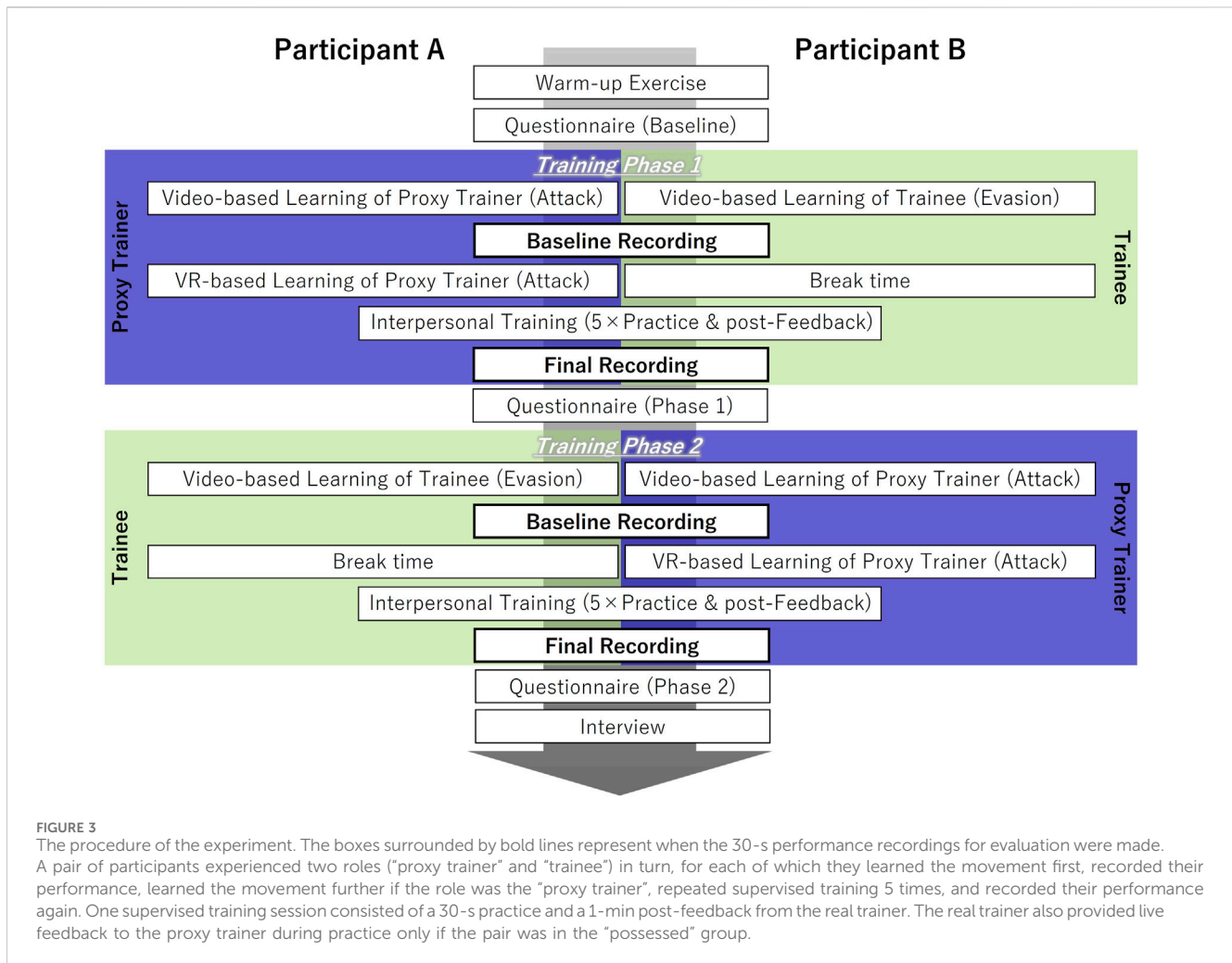
Finally, in the last phase, participants answered the phase 2 questionnaire and commented on the experiment in a semi-structured interview.

### 3.2.2.4 Technical setup of the sensei possession system

As described in the procedures above, the overall flow of a training phase consisted of 5 steps; a preliminary learning step for both roles, a baseline recording step, a preliminary learning step only for proxy trainers, an interpersonal training step consisting of 5 sessions, and finally a final recording step. We set the two preliminary steps based on the result of our pilot workshop.

In the first preliminary learning step, participants watched a 3-min video that we made for each role. In the video, the correct movements were explained in detail, covering all the evaluation criteria described in the next paragraph.

In the second preliminary learning step after the recording of baseline performance, participants playing the role of a proxy trainer practiced the attack movement along with auditory commands. The proxy trainers wore an XR headset (Meta Quest 2, which has a video-see-through function) for 2 min to observe an XR-based learning material that we developed in Unity. We decided to use the XR headset only in this step and not in the next interpersonal training step to prevent the neckache reported in our pilot workshop. In this XR learning material, the participant could observe an animated avatar of a trainer and a still avatar of a trainee. The trainer avatar repeatedly performed the attacking motion, against the trainee avatar which stood still in the neutral posture. We implemented the video-see-through function and also made the avatars semi-transparent so that the participants could compare their own posture with the avatar. We also placed an instruction panel next

**FIGURE 3**
The procedure of the experiment. The boxes surrounded by bold lines represent when the 30-s performance recordings for evaluation were made. A pair of participants experienced two roles ("proxy trainer" and "trainee") in turn, for each of which they learned the movement first, recorded their performance, learned the movement further if the role was the "proxy trainer", repeated supervised training 5 times, and recorded their performance again. One supervised training session consisted of a 30-s practice and a 1-min post-feedback from the real trainer. The real trainer also provided live feedback to the proxy trainer during practice only if the pair was in the "possessed" group.

to the avatars, which instructed the participant to perform attack motions according to the auditory commands: At the sound of a whistle, the trainer was supposed to get ready (Attack Step 1). With the swoosh sound of the strike, which was played from either the right or left ear, they were instructed to strike from the direction of the played sound (Attack Step 2).

In this XR-based learning material, the direction of attacks (right and left) was in turn for the first 1 min (level 1) and random for the next 1 min (level 2). The real trainer asked the proxy trainer to read the instruction panel and try to mimic the animated avatar in these 2 levels. This was for the participants to focus on the postures at level 1, and then learn the correspondence between the motions and the auditory commands in level 2. This correspondence was only to be used in the live-feedback system in the possessed condition. However, we decided to utilize this 2-min preliminary learning of proxy trainers for both conditions to level the amount of training before the next training step.

In the interpersonal training step, participants underwent 5 repeated sessions of practice, in each of which they performed the attack and evasion for 30 s followed by a 1-min break. During the 30 s of performance, the proxy trainers in the possessed condition received live feedback, while those in the unpossessed condition performed independently. For the live feedback in the possessed

condition, we decided to use a wireless headphone (Sony WH-1000XM3) instead of the XR headset to deliver the auditory commands to the proxy trainer. The reason was that the light and balanced weight of the headphones was suitable for long-term training. Also, the headphones allowed the real trainer and the proxy trainer to hide what movement was instructed next. We communicated the instructions from the real trainer to the proxy trainer through two sounds used in the preliminary XR-based learning. The real trainer could play the audio command of the strike (swoosh sound) from either the right or left side of the headphones to indicate the direction of the strike. On the other hand, the proxy trainers in the unpossessed condition wore the same headphones but they did not receive the audio command.

During the 1-min break in a repeated session, the real trainer talked with the proxy trainer and the trainee to point out what they should focus on to improve their performance based on the criteria (described later). This conversation took place via the aforementioned monitor in the participants' zone, which showed the real trainer behind the partition captured by an inner camera of a laptop PC (Panasonic CF-SZ6). We decided to provide this post-feedback in both conditions to provide participants with the instructions they would receive in conventional video-based remote training. In addition to the post-feedback, the real trainer asked the proxy trainer to change the difficulty

TABLE 1 Correlations between all measures. IM = intrinsic motivation. The values in brackets are the internal consistencies (Cronbach's $\alpha$; where applicable). Correlations marked with an asterisk are significant (*: $p < 05$, **: $p < 01$, ***: $p < 001$).

| Measure | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1. IM | (.93) | | | | | | | |
| 2. IM Excitement | .86*** | (.88) | | | | | | |
| 3. IM Challenge | .91*** | .56** | (.93) | | | | | |
| 4. Obsessive Passion | .48* | .37 | .49* | (.90) | | | | |
| 5. Harmonious Passion | .61** | .52* | .55* | .48* | (.81) | | | |
| 6. Speed Control | .32 | .34 | .23 | .07 | −.01 | | | |
| 7. Attack Quality | .26 | .28 | .16 | .16 | .48* | −.13 | | |
| 8. Evasion Success | .15 | .35 | −.06 | .25 | .26 | −.04 | .26 | |
| 9. Evasion Quality | .31 | .33 | .23 | .21 | .42 | .09 | .52* | .37 |

of the attack, to let the trainee eventually be able to evade the attack at the right time while keeping the correct posture: In the first three repetitions, the real trainer instructed the proxy trainer to strike alternately from left and right. At the same time, the real trainer advised the trainee to focus on the posture as the direction attacks are predictable. In the 4th repetition, the direction of attacks became random to let the trainee learn to evade as a reaction to the attack. Finally, in the 5th repetition, the real trainer asked to mix in some feint attacks to let the trainee learn to be patient and start evading at the latest time possible. Accordingly, the real trainer in the possessed condition played the audio commands during the 30-s performance to increase difficulty as above.

For both the baseline and final recording steps before and after the interpersonal training step, the proxy trainer took off the headphones and performed the attack independently. With this setup, we intended to measure the performance of the participants when they could not rely on the live feedback from the real trainer.

### 3.2.2.5 Measures
#### 3.2.2.5.1 Performance data (video analysis).
In preparation for evaluating the motor skill performance, we made a list of evaluation criteria under the supervision of the Aikido Club. In this list, the movements and postures of the participants were scored from 0 to 2 for each criterion (0: there was no effort shown to achieve the criterion or the criterion was misunderstood, 1: there was effort shown to achieve the criterion but it was not successful, 2: the criterion was successfully achieved). The example of the "Evasion Timing" criterion in the evasion is as follows:

2 *points*: The whole evasive motion [the head moved more than a distance of one head in the correct direction, and both legs have moved out from the attacker's direction of travel (the shoulder width zone in from of the attacker)] was executed while the attacker's hand has started moving towards the head and then reaches the height of the trainee's head before the attack.
1 *point*: Although the evasive motion was completed before the end of the attack, the trainee started evading before the attack had started.
0 *point*: The trainee failed to complete the evasion before the end of the attack, or the direction of evasion was the opposite.

We mentioned all these criteria in the videos for the preliminary learning, and the real trainer also referred to them when he provided feedback to the participants in the training phases.

Including the above criteria, the instructor at the Aikido club recommended 8 measures for our study, 4 for the proxy trainer and 4 for the trainees. Among those, speed control on the trainer side and evasion success on the trainee side were most important because we focused on real-time feedback. We summarized the other three scores for each role in a compound measure called "Quality". The detailed measures are as follows:

Attack Speed Control: We evaluated how much the real trainer controlled the proxy trainers by counting the repetitions in a recording. This is because the real trainer used live feedback to control the proxy trainer's timing and speed. We consulted with the Aikido trainer about the ideal number of repetitions in a 30-s recording, which resulted in 5. Thus, we subtracted 5 from the recorded number of repetitions and defined it as the score of speed control.

Attack Quality: We reviewed each of the proxy trainers' repeated movements using the 13 criteria for attack and calculated the mean of each in a recording. Then, we calculated the following by first dividing the 13 criteria into 3 groups:

Stability: The proxy trainers' ability to stabilize themselves during attack motion. The score comprises 4 criteria (ranging from 0 to 8).

Range: The proxy trainers' ability to distance themselves against their opponents during attack motion. The score comprises 4 criteria (ranging from 0 to 8).

Accuracy: The proxy trainers' accuracy of attack motion. The score comprises 5 criteria (ranging from 0 to 10).

We took the sum of scores in each group and z-standardized them. Then, we calculated the mean value of the 3 and used it as the representative value of quality.

Evasion Success: Learning to evade successfully is the most important factor in this training of evasion. To evaluate whether the trainees learned to evade, we reviewed each trainee's repeated movements in a recording using the "Evasion Timing" criterion. Then, we calculated the ratio of success (2 points) in a recording and used the number (ranging from 0 to 1) as the score of evasion success.

Evasion Quality: We reviewed each of the trainees' repeated movements using the 11 criteria for evasion (not including "Evasion Timing") and calculated the mean of each in a recording. Then, we calculated the following by first dividing the 11 criteria into 3 groups:

Stability: The trainees' stability of their position and movement. The score comprises 6 criteria (ranging from 0 to 12).

Range: The trainees' ability to distance themselves against the attacking opponent. The score comprises 2 criteria (ranging from 0 to 4).

Accuracy: The trainees' accuracy of the evasion technique. The score comprises 3 criteria (ranging from 0 to 6).

Similarly to the Attack Quality, we took the sum of scores in each group and z-standardized them. Then, we calculated the mean value of the 3 and used it as the representative value of quality.

To evaluate the above 4 measures, we first shuffled the recordings of all participants. Then, after at least 1 week of memory washout period, the first author observed the video at half speed and coded the performance of participants based on the list of criteria.

**3.2.2.5.2 Subjective data.**       As mentioned in the procedures, we collected subjective data by asking participants to complete 3 questionnaires at the beginning, the middle, and the end of the experiment. Each questionnaire comprised items for the following factors:

Intrinsic Motivation: We used the SMS (Sports Motivation Scale) to measure intrinsic motivation. The SMS is a questionnaire established by Pelletier et al. (1995) to measure the motivation of sports athletes. We used the Japanese translation (Matsui, 2014). It had a high internal consistency (See Table 1).

An exploratory Principal Component Analysis (PCA) using oblimin rotation and parallel analysis revealed an interesting substructure of the IM scale in our study context, with two different facets (See Supplementary Appendix A for an exact item overview). Six items formed a motivation facet we labeled "challenge." An example item of this facet is "Because I feel a lot of personal satisfaction while mastering certain difficult training techniques." Items of this facet describe how motivation relates to being challenged and achieving difficult goals. The other facet contained five items and we labeled it as "excitement." An example item is "For the pleasure I feel in living exciting experiences." One item loaded on both facets ("For the pleasure that I feel while executing certain difficult movements"). This seems plausible because it describes both the pleasurable feeling (as in the excitement facet) and the cause in difficult movements (as in the challenge facet). We excluded it for the facet scores. For the IM scale, we calculated the overall individual scores for each participant and the facet scores as the means of the associated items. In our analysis below, we include an exploratory analysis of the two subscales.

Passion: We used the PSS (Passion Scale for Sports) established by Vallerand et al. (2003) to measure passion. The PSS comprises 10 items and was translated to Japanese by Fujita (2011). We included it in our study because we assumed that our proposed method might enhance proxy trainers' obsessive passion for intensively realizing the commanded movement, which could cause them to be narrow-minded and increase the risk of injury (Akehurst and Oliver, 2014). It contains two subscales: Harmonious Passion and Obsessive Passion (Vallerand et al., 2003). Both

subscales had good internal consistencies (See Table 1). A PCA using oblimin rotation and parallel analysis confirmed the questionnaire's component structure. Individual scores are the means of all subscale items.

Mindfulness: We included the AMQ (Athlete Mindfulness Questionnaire) established by Amemiya et al. (2015) based on the Five Facet Mindfulness Questionnaire (Baer et al., 2006) to measure participants' mindfulness with 21 items. The AMQ has been used in clinical sports psychology to evaluate the mindfulness of athletes concerning burnout and performance (Amemiya and Sakairi, 2019; 2021). It contains four subscales: Awareness, NonJudging, Observing NonReactivity, and Describing.

However, we found some problems with the AMQ's psychometric quality in the context of our study. Internal consistency was good for Awareness ($\alpha$ = .87) and Describing ($\alpha$ = .88). However, it was insufficient for Nonjudging ($\alpha$ = .57) and borderline for Observing NonReactivity ($\alpha$ = .60). In both cases, the internal consistency could not be increased to an acceptable level via item exclusion.

In addition, we ran a PCA to confirm the questionnaire's internal structure. We explored different PCA models, using oblimin and varimax rotation and the number of components either set to 4 (based on the model) or determined by parallel analysis. However, the items did not follow the suggested component structure of the AMQ in any of these variations. Several items had a factor loading > .4 on two or three components. Moreover, items from the same subscale were spread across multiple components for Awareness, NonJudging, and Observing NonReactivity. Given these problems with internal consistency and internal structure in our study context, we excluded the AMQ from further analysis.

Interest: To measure the interest in the training, we included two items from the motivation questionnaire originally introduced by Yogi (2012) for the class of Judo, a traditional martial art similar to Aikido. We changed the terms in the items to apply to our Aikido tasks, and the final items are translated as follows:
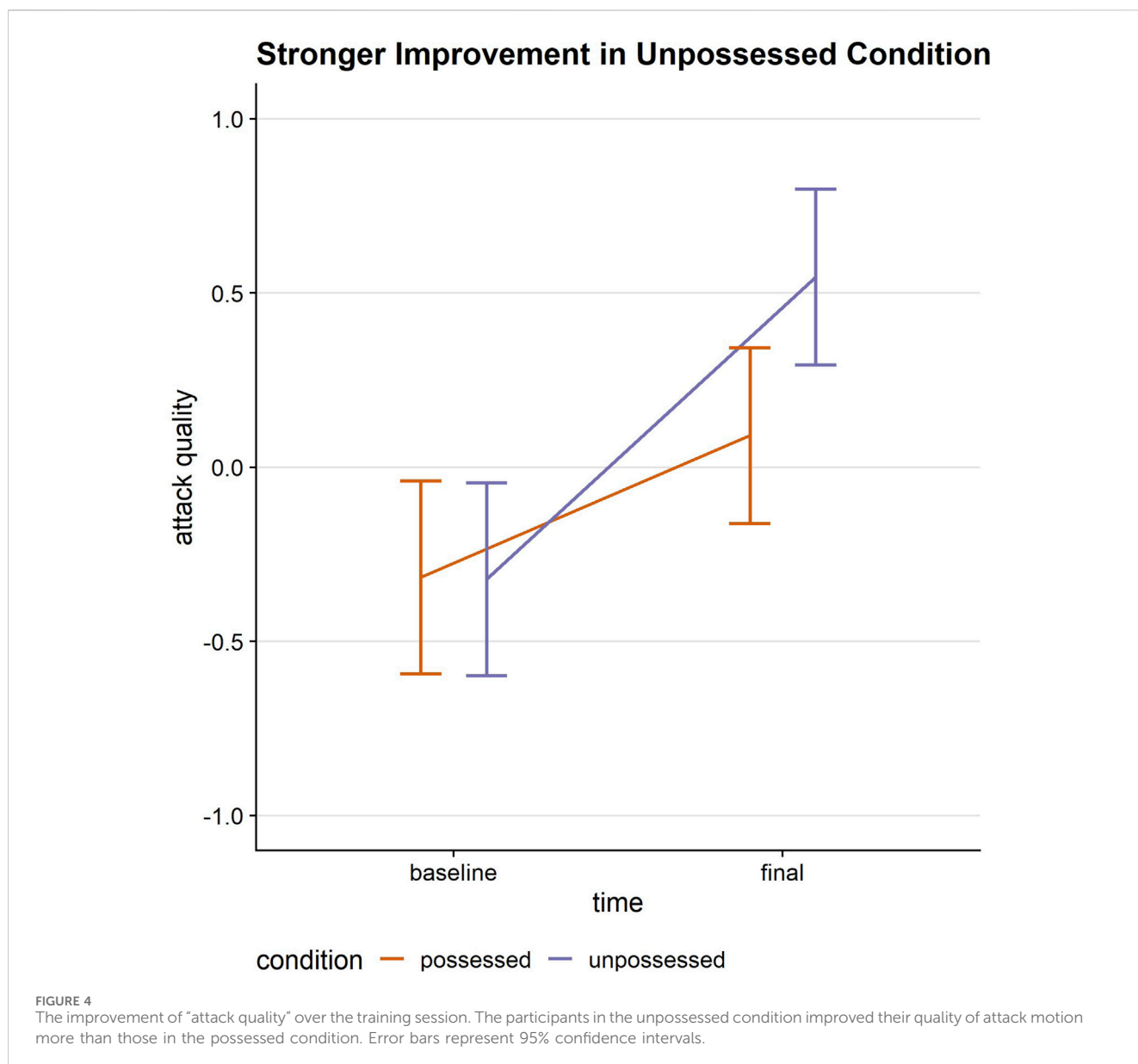
- What do you think about the training of Aikido?
- What do you think about exercises with physical contact like Aikido?

Yogi collected responses using a 4-point Likert-type scale, but we used a 6-point Likert-type scale (1-I am not interested at all, 6-I am very interested) instead to analyze the measure in more detail. We used the responses to the above two items respectively as the scores of Aikido-interest and exercise-interest.

In addition to the items above, we added 4 items below to the second and third questionnaires (phase 1 and phase 2) for measuring the impression of the proxy trainer toward the trainee:

- The partner was considerate of my skill level.
- The speed of my improvement increased thanks to the partner's consideration.
- The proxy trainer moved exactly as the real trainer had commanded.
- The proxy trainer was like the real trainer during the training.

For the third item, we asked the proxy trainers to respond with their self-evaluation, while we asked trainees to estimate how well

**FIGURE 4**
The improvement of "attack quality" over the training session. The participants in the unpossessed condition improved their quality of attack motion more than those in the possessed condition. Error bars represent 95% confidence intervals.

another participant performed. The responses were given using a 7-point Likert-type scale (1-strongly disagree, 7-strongly agree).

For the interview at the end of the experiment, we assumed a semi-structured interview would be appropriate after conducting the preliminary experiment with the same training task and the unstructured interview 2 times (N = 4). Thus, we asked the following questions in the main experiment:

- Would you like to keep using the training system that you have used, and for what reason?
- Did your motivation to learn Aikido such as "I want to be better" change, and why?
- Do you think you had understood or acquired the "Timing" in the last 2 repetitions of the training, and why?

Additionally, participants were asked to comment freely if they had noticed anything during the experiment. We also recorded the videos of the entire training and the interview for later analysis.
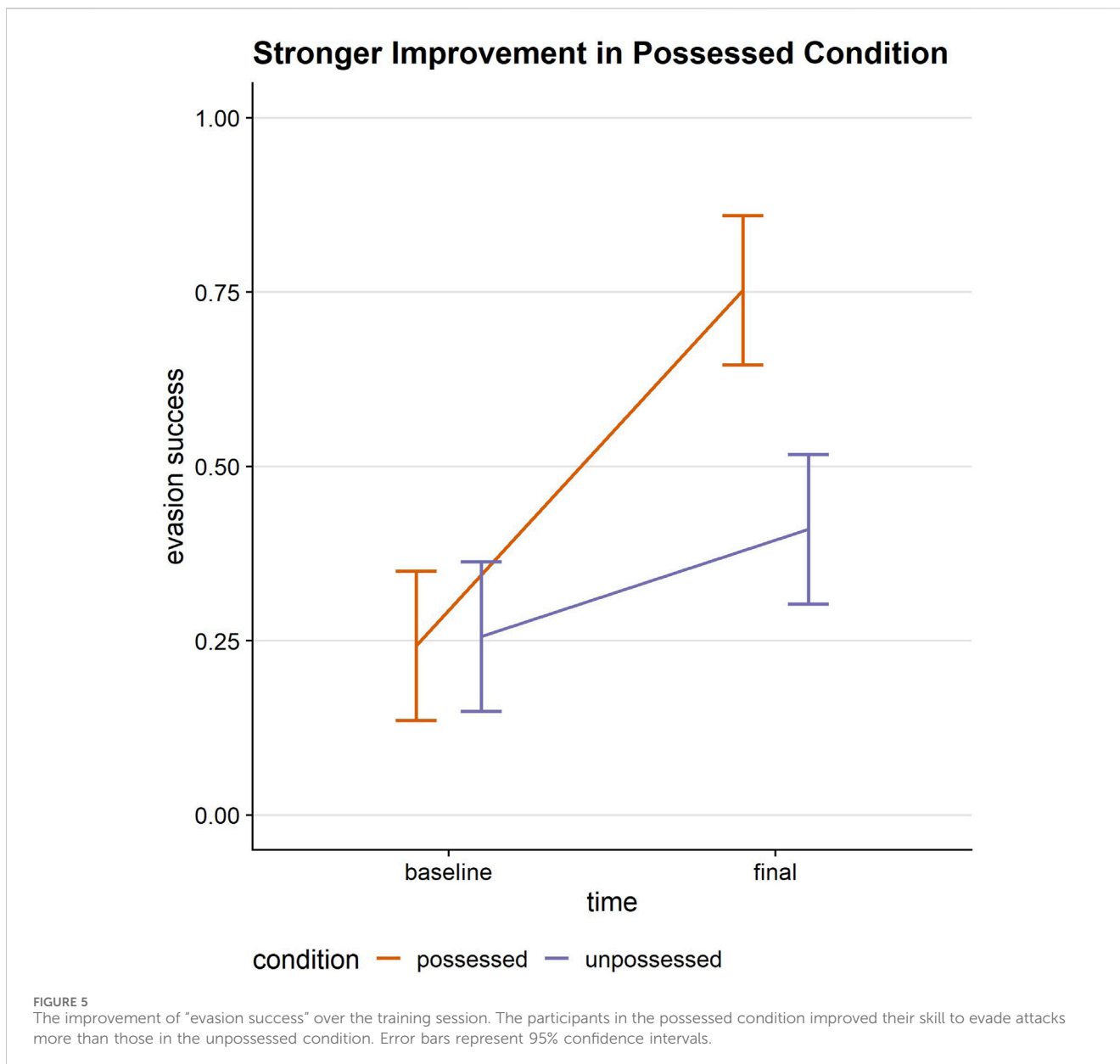
# 4 Results

## 4.1 Performance data

### 4.1.1 Attack speed control: worsening control over time

We hypothesized that the "possessed" proxy trainers with live feedback during training would have a stronger training effect toward optimal attack speed control than the "unpossessed" proxy trainers without live feedback. Statistically, the "attack speed control" scores (the number of attack repetitions subtracted by the ideal number) of the "possessed" group should be closer to zero after training than the "unpossessed" group.

To test this, we ran a $2 \times 2 \times 2$ mixed ANOVA with the repeated factor "time" (baseline, final), and the between-participants factors "condition" (real-time possessed vs. unpossessed) and "first role" (proxy trainer vs. trainee). As the dependent variable, we used the attack speed control score.

**FIGURE 5**
The improvement of "evasion success" over the training session. The participants in the possessed condition improved their skill to evade attacks more than those in the unpossessed condition. Error bars represent 95% confidence intervals.

We found a main effect of "time" ($F$ (1, 16) = 10.32, $p < .01$ $\eta_p^2 = .39$), which showed a significant increase of the score over time. This indicated that the proxy trainers' control of attack speed worsened over the training, regardless of the conditions.

No other effect was significant.

### 4.1.2 Attack quality: improvement over time

We checked if all the proxy trainers in both "possessed" and "unpossessed" groups learned to perform the correct attacking motion. As a result of a 2 × 2 × 2 mixed ANOVA, we found that the main effect of time was significant ($F$ (1, 16) = 44.40, $p < .001$ $\eta_p^2 = .74$). However, in addition to this main effect, we also found a significant interaction effect of time and condition ($F$ (1, 16) = 5.83, $p < .05$ $\eta_p^2 = .27$). The attack quality improved for participants in both the possessed condition ($t$ (16) = 3.01, $p < .05$) and the unpossessed condition ($t$ (16) = 6.42, $p < .001$) over time, which confirms our assumption of proxy trainers' learning.

Meanwhile, the interaction plot seemed to indicate a stronger effect in the unpossessed condition than in the possessed condition (See Figure 4). However, the *post hoc* comparison of the ANOVA showed no significant effect between the final attack quality in the possessed and unpossessed condition ($t$ (16) = 2.10, $p = .16$). In sum, the interaction effect seems to indicate that attack quality was slightly better in the unpossessed condition, although the findings are not conclusive.

### 4.1.3 Evasion success: stronger improvement in possessed condition

We hypothesized that the "possessed" group with live feedback would have stronger training effects in evasion success than the "unpossessed" group without live feedback. To test this, we analyzed the "evasion success" data from the video recordings. We ran a 2 × 2 × 2 mixed ANOVA with the repeated factor "time" (baseline, final), and the between-participants factors "condition" (real-time

possessed vs. unpossessed) and "first role" (proxy trainer vs. trainee). As the dependent variable, we used the ratio of successful evasion during the recording.

We found a main effect of "time" ($F$ (1, 16) = 52.92, $p < .001$ $\eta_p^2 = .77$), which indicated that training had a positive overall effect on the participants' evasion skill. However, in addition to this main effect, we also found a significant interaction effect of time and condition ($F$ (1, 16) = 15.21, $p < .01$ $\eta_p^2 = .49$), which provides further insights. Although the success rate improved for participants in both the possessed condition ($t$ (16) = 7.90, $p < .001$) and the unpossessed condition ($t$ (16) = 2.39, $p < .05$), the effect was stronger in the possessed than the unpossessed condition ($t$ (21.60) = 2.94, $p < .01$). This confirms our hypothesis that the real-time feedback on the proxy trainers enhances the training efficiency of the opposing trainees (See Figure 5).

In addition, we found a two-way interaction effect of time and "first role" ($F$ (1, 16) = 5.81, $p < .05$ $\eta_p^2 = .27$) and a three-way interaction effect of time, first role, and condition ($F$ (1, 16) = 6.69, $p < .05$ $\eta_p^2 = .30$). A *post hoc* exploratory analysis of the two-way interaction revealed that both participants who first took the role of proxy trainers ($t$ (16) = 3.44, $p_{Holm} < .05$) and participants first starting as trainees ($t$ (16) = 6.85, $p_{Holm} < .001$) overall improved their performance through the training. On this level of analysis, improvement was similarly strong, independent of which role they took first ($t$ (16) = 0.92, $p_{Holm} = .69$). However, the three-way interaction provided further insights. In the possessed conditions, participants improved both when starting as proxy trainers ($t$ (16) = 5.68, $p_{Holm} < .001$) and as trainees ($t$ (16) = 5.50, $p_{Holm} < .01$). But in the unpossessed condition, participants only improved when starting as trainees ($t$ (16) = 4.19, $p_{Holm} < .05$). There was no improvement in evasion success for participants who started as proxy trainers in the unpossessed condition ($t$ (16) = −0.81, $p = 1.00$). In other words, the training improvement for participants who started as proxy trainers was only based on the possessed condition.

In sum, the possessed XR intervention had a stronger positive effect on evasion success, compared with the unpossessed condition. In addition, the improvement was independent of whether participants started as proxy trainers or as trainees. In contrast, the unpossessed condition also had positive effects on evasion success, but only for participants who started as proxy trainers, not for participants who started as trainees.

### 4.1.4 Evasion quality: overall improvement over time

We hypothesized that the trainees of the "possessed" group with live feedback would have better performance gains of evasion quality than the "unpossessed" group without live feedback because of the higher quality of opposing proxy trainers. However, a 2 × 2 × 2 mixed ANOVA showed no significant effects other than the main effect of time ($F$ (1, 16) = 42.86, $p < .001$ $\eta_p^2 = .73$).

## 4.2 Subjective data

### 4.2.1 Intrinsic motivation

The real trainer guided down to the small details of the speed of the proxy trainer in the possessed condition. Thus, we hypothesized

that this real-time feedback would reduce the proxy trainers' intrinsic motivation toward the training. To test this, we ran a 3 × 2 × 2 mixed ANOVA with the repeated factor "time" (baseline, session 1, session 2) and the between-participants factors "condition" (real-time possessed vs. unpossessed) and "first role" (proxy trainer vs. trainee). The dependent variable was "intrinsic motivation". For an overview of descriptive data, see Table 2.

#### 4.2.1.1 Growing motivation over time

We found a main effect of "time" ($F$ (2, 32) = 5.54, $p < .01$, $\eta_p^2 = .26$). Intrinsic motivation was higher after session 1, compared with the baseline ($t$ (16) = 3.03, $p_{Holm} < .05$). It was also higher after session 2, compared with the baseline ($t$ (16) = 2.53, $p_{Holm} < .05$). The increase from session 1 to session 2 was not significant ($t$ (16) = 1.39, $p_{Holm} = .18$).

#### 4.2.1.2 Higher motivation when starting as trainee

We also found a main effect of "first role" ($F$ (1, 16) = 4.77, $p < .05$, $\eta_p^2 = .23$). Participants who started as trainees had a higher overall intrinsic motivation than participants who first started as proxy trainers.

No other effects were significant.

#### 4.2.1.3 Exploration: time effect based on "challenge," first role effect based on "excitement"

We ran a further exploratory analysis with the intrinsic motivation facets "challenge" and "excitement" described above. To that end, we ran two further 3 × 2 × 2 mixed ANOVAS with the same independent factors and the two facets as dependent variables.

In the first ANOVA with the "challenge" facet, we found the same main effect of "time." Challenge-based intrinsic motivation was higher after the first session than after the baseline ($F$ (2, 32) = 7.46, $p < .01$, $\eta_p^2 = .32$). In addition, the increases from baseline to session 1 ($t$ (16) = 3.36, $p_{Holm} < .05$) and from baseline to session 2 ($t$ (16) = 2.85, $p_{Holm} < .05$) were significant. However, the increase from session 1 to session 2 was not ($t$ (16) = 1.35, $p_{Holm} = .20$). In this analysis, the "first role" effect disappeared ($F$ (1, 16) = 1.86, $p = .19$).

We found the opposite pattern in the second ANOVA with the "excitement" facet. Here, the "time" effect disappeared ($F$ (2, 32) = 0.57, $p = .57$), but the "first role" effect remained significant ($F$ (1, 16) = 6.13, $p < .05$, $\eta_p^2 = .28$).

In sum, our exploratory analysis based on the two facets, "challenge" and "excitement" of intrinsic motivation, indicates a distinctive effect (See Figure 6). The increase in intrinsic motivation seems to be based on a challenging experience. The higher motivation in the group that started as trainees seems to be based on an experience of excitement.

### 4.2.2 Obsessive and harmonious passion: no effects

We developed the XR intervention with the intention of reproducing the physical body of a real trainer over a proxy trainer, which means the proxy trainer had to focus intensively on obeying the real-time commands from the real trainer. Thus, we assumed a possible risk of proxy trainers' obsessive passion being higher in the possessed condition, which could eventually lead to employing maladaptive behaviors in the future.

TABLE 2 The overview of intrinsic motivation score. "Condition" and "first role" are the between-participants factors. 5 participants belonged to each combination of the two factors. "S.D." stands for standard deviation. "Time" is the repeated factor, comprising "baseline" measured before two sessions of training, "phase 1" measured right after the first training phase, and "phase 2" measured after the second training phase.

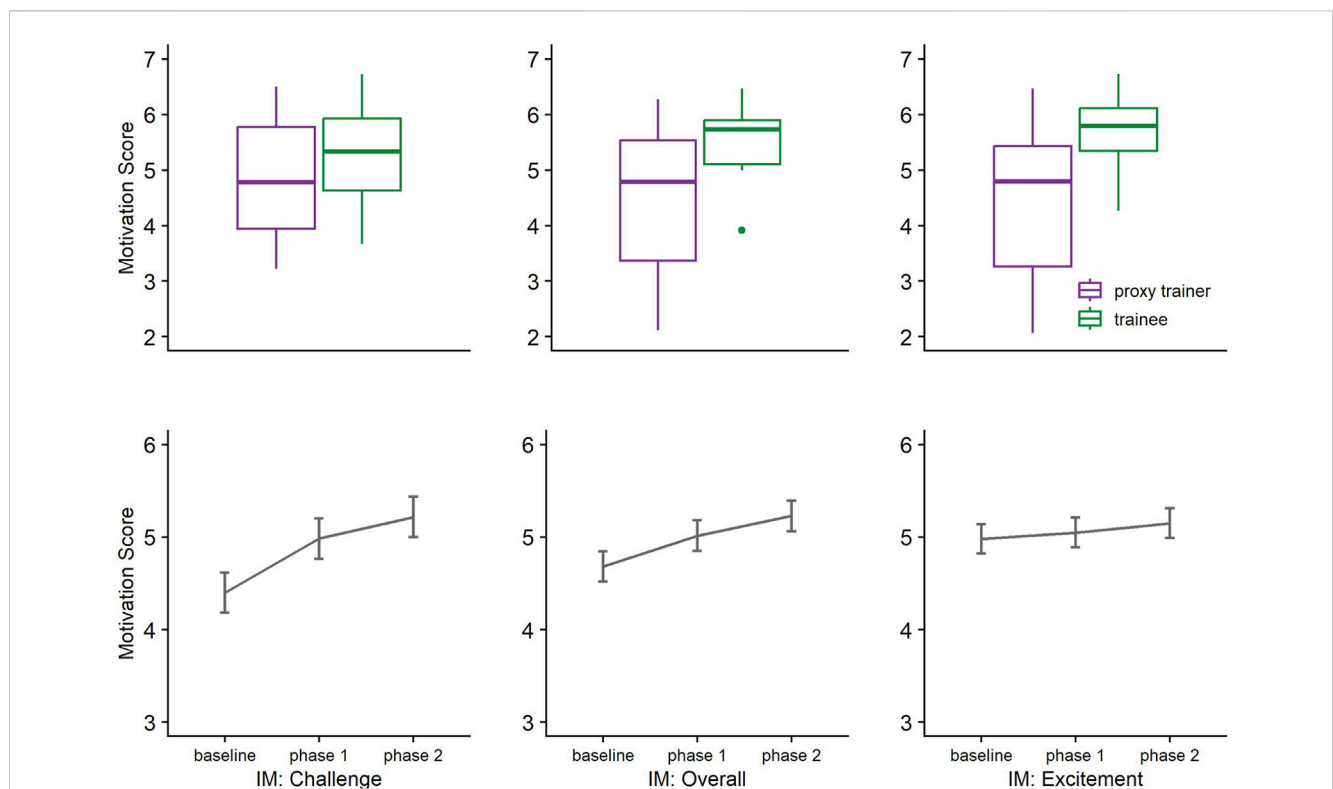| Condition | First role | Time | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | Baseline | | Phase 1 | | Phase 2 | |
| | | Mean | S.D. | Mean | S.D. | Mean | S.D. |
| Possessed | Trainee | 5.35 | 0.99 | 5.72 | 0.95 | 5.80 | 1.06 |
| | Proxy Trainer | 4.90 | 1.29 | 5.07 | 1.28 | 5.30 | 1.27 |
| Unpossessed | Trainee | 5.05 | 0.35 | 5.37 | 0.29 | 5.70 | 0.65 |
| | Proxy Trainer | 3.43 | 1.87 | 3.92 | 1.68 | 4.12 | 1.00 |



FIGURE 6
The two intrinsic motivation effects. The top row shows that the effect of "first role" on overall motivation (middle) is driven by a difference in the "excitement" facet (right). In contrast, the bottom row shows how the increase in intrinsic motivation over time is driven by an increase in the "challenge" facet (left). IM = "intrinsic motivation". Error bars represent standard errors.

To check for this, we ran two $3 \times 2 \times 2$ mixed ANOVAs with the repeated factor "time" (baseline, session 1, session 2) and the between-participants factors "condition" (real-time possessed vs. unpossessed) and "first role" (proxy trainer vs. trainee). The dependent variables were "obsessive passion" and "harmonious passion," respectively. For "obsessive passion," we only found a significant three-way interaction effect ($F_{(2, 32)} = 3.63$, $p < .05$, $\eta_p^2 = .19$). However, it did not relate to our assumptions and a *post hoc* analysis revealed no significant pairwise differences. No other effect was significant (all $p > .21$). In the analysis of "harmonious passion," no effects were significant (all $p > .18$).

In sum, although our intervention design may make the risk of being obsessively passionate about performance plausible, we found no indication of an increased risk regardless of real-time possession.

### 4.2.3 Qualitative data/interviews

To complement the quantitative analysis, we looked into the interview data. Specifically, we were interested in whether these could give us further insight into three effects, that we did not anticipate before the experiment: 1) slightly stronger attack quality gains in the unpossessed condition, 2) the distinction between intrinsic motivation facets and their distinct effects, and 3) first role/condition 3-way interaction.

#### 4.2.3.1 Attack quality: live feedback can be disturbing

Participants in the possessed group commented on the difficulty they had while following the commands they heard from the headphones. One participant said: "It took me some time to decide my movements based on the audio commands," and another one "The audio commands of direction was not intuitive." These comments indicate the excessive use of proxy trainers' cognitive resources during the training in possessed conditions, which could have reduced their attack quality gains.

#### 4.2.3.2 IM facets: challenges found through the training increased motivation

In the interview, participants commented on their increased motivation by narrating how much they had improved and becoming interested in the next challenges they wanted to overcome. One participant said: "Through my experience of acquirement, I thought I wanted to acquire more," and another one "I became motivated to be better. Although I thought the movements would be easy once I memorized them at first, I eventually became aware of the need to look at the opponent and realized the deepness of the skill." These comments support the growth of their motivation over time, which matches the time effect on "challenge" found in the quantitative analysis of intrinsic motivation.

#### 4.2.3.3 First role: poor quality of preliminary learning impacted first proxy trainers

Participants commented on the difficulty of movement in both roles. Although one of the participants whose first role was the trainee commented, "the partner's attack movement in the former session helped me perform in my turn" and no such comments were given by the participants whose first role was the proxy trainer. In addition, many whose first role was the proxy trainer mentioned the difference between the attack motion shown in the XR preliminary learning material and the reality, and that "the motion of a real human was smoother and more helpful than that of the virtual avatar." These opinions indicated the lack of quality in the preliminary learning material for proxy trainers and that these were more recognized by the group who initially performed proxy trainers because the other half could refer to their partners in the former session, which may have influenced their different excitement levels.

## 5 Discussion

In this paper, we tested how XR-based preliminary learning material and the sensei possession (i.e., remote control of a proxy trainer via live feedback from a real trainer) can be used for contact sports training. Specifically, we tested the effect of live feedback on a paired training of Aikido and assumed that participants in the possessed group with live feedback would show greater performance gains and no inferior motivation compared to those in the unpossessed group without live feedback. We found that the ratio of the successful Aikido-style evasive motion improved more in the possessed condition than in the unpossessed condition, while we found no sign of a negative influence on participants' motivation in both conditions.

Our main contributions are the implementation of XR-based preliminary learning materials to the Sensei possession system for contact sports training, and the confirmation of the increased effectiveness of Aikido training when proxy trainers had live feedback from the real trainer. Also, our analysis of the subjective intrinsic motivation revealed two facets "challenge" and "excitement", on each of which the training and the quality of preliminary learning materials might have influenced.

## 5.1 Broader impact

Marraffino et al. (2021) reported that micro-adaptive training resembling in-person tutoring improved trainee's skill more than less-frequently scheduled adaptive training. However, we extended their finding that inexperienced proxy trainers could offer the micro-adaptation more effectively with live intervention by a real experienced trainer, compared to the unpossessed proxy trainers. We assume that the attempt of the unpossessed proxy trainers to conduct adaptive training for the trainees did not succeed because they lacked the necessary skills to recognize the trainees' needs due to lack of experience compared to the real trainer (Sharma et al., 2017).

Matsui (2014) performed principal component analysis on the Sports Motivation Scale with the responses he collected from multiple sports clubs at once, from which he found one factor "intrinsic motivation". Here, we found that two components "challenge" and "excitement" appeared when this scale was used repeatedly during the training with the Sensei possession system. Our exploratory analysis indicated that the growing intrinsic motivation over time was based on the "challenge," and the higher motivation experienced by the group who started as trainees was based on the "excitement". Participants' comments in the interview indicated that they increased their motivation over the training by finding "challenges" on the way to improving their performance. Also, we found the factor that may have reduced the "excitement" of participants who became proxy trainers first: Participants' comments indicated that the XR-based preliminary learning materials lacked quality, and the participants who became proxy trainers first had more confusion than those who could observe the opponent's performance as a proxy trainer before playing the part themselves. We assume that we were able to find these two components because the above factors that may have affected the two components were absent in the experimental setting of previous work.

## 5.2 Live feedback assures evasion training effectiveness

In our analysis, we found two interaction effects in relation to participants' first role in the training. The first, two-way interaction suggested that participants improved their evasion performance, independent from the first role. However, the second, three-way interaction revealed that participants in the unpossessed group only improved their performance if their first role was the proxy trainer.

We attempted to further investigate the cause of this complex interaction. Specifically, we hypothesized that the particular training order helped participants who started as proxy trainers and then became trainees to compensate for the lack of live feedback. For example, the participants who started as the proxy trainers in the

unpossessed conditions may have learned by observing their partners before switching roles (Wulf et al., 2010). This is in line with one participant's comment in the interview about observing the partner's attacking movement in phase 1 and utilizing the experience in phase 2. This did not refer to evasion, but we could assume some similarities. However, observation effects were not part of our study design, so we did not measure them and only had a single qualitative comment about this. Thus, the cause of this complex interaction remains uncertain, and we may need to conduct further studies to clarify it.

Despite this, our overall findings indicate that Sensei possession with live feedback was more effective when trainees had no experience with the Aikido training task. Although some trainees in the unpossessed condition could improve the evasion technique, this effect was more robust in the possessed condition, and independent of role switching, allowing for more time-efficient training.

## 5.3 Design implication

In our study, we found that improvement of the Sensei possession system is possible from three perspectives; the video-based preliminary learning material, the XR-based preliminary learning material, and the live feedback for proxy trainers.

First, the participants commented that the 3D animation of XR avatars was more helpful compared to the video of actual humans in the video-based learning material. This indicates that 3D video, which can be recorded with smartphones for the general public, would work more effectively as the preliminary learning material. However, when we play the 3D video with an XR headset, we should scale and place the video player to the level that observers need not move their necks.

Second, as reported in our result, some participants commented negatively on the lack of XR animation quality compared to actual human movement. We assumed we could not achieve high quality because we developed the animation by manipulating avatars inside Unity software, which is not built for animation design. Therefore, we suggest that future designers either use a motion-tracking system to record the movement of experienced trainers or employ appropriate software and developers.

Third, the participants' opinions in the interview indicated that our live feedback design requires improvement. From the opinion that participants required some time to understand the abstract sound of audio commands, we assume that the sound effects (beeping whistle and swooshing attack) should be replaced with language-based audio commands, such as "ready," "right," and "left." This use of language-based commands would be beneficial when more actions are to be commanded in more complex training.

Also, we learned from the participants' opinions that the cognitive load of focusing on the whole-body movement was overwhelming. One said, "When I focused on either arm or foot, I became unable to control the other, regardless of the role I was playing." Regarding this opinion and the other comments of confusion about the audio commands, we suggest implementing more intuitive commands in addition to audio. For example, we could use wearable vibrotactile devices on legs for the commands of leg movement, which Tanaka et al. (2022) reported to be an effective method for presenting feedback on whole-body movement.

## 5.4 Limitations

The first limitation of our study is the small sample size. Our power analysis using MorePower 6.0 (Campbell and Thompson, 2012) with $\alpha = .05$ and $\eta^2 = .06$ showed that the test powers were .31 for $2 \times 2 \times 2$ mixed ANOVA and .65 for $3 \times 2 \times 2$ mixed ANOVA. This means that the confidence level in the non-significant effects was relatively low. This small sample size was attributed to our aim to test the contact sports training by inexperienced participants, for which we had to test the recruited participants in pairs and also minimize participants' risk of infection under the COVID-19 safety measures. Although the widespread of COVID-19 was our motivation to develop a remote training method, future works should choose when the risk of disease is relatively few and try to recruit more participants for experiments. In such an experiment with more participants, the automation of video analysis would be useful to code the performance data, as suggested by Bridgeman et al. (2019).

Secondly, our study lacked comparisons with conventional training methods such as fully-video-based training or face-to-face training. In our study, all conditions included the XR-based training method to focus on the effect of live feedback on proxy trainers. Hence, although we partially revealed the advantages of our proposed method over the training without live feedback, whether the conventional training can be replaced by our remote method is not yet evident. To clarify how much our proposed method can achieve in reproducing in-person training, future works should add a fully-video-based condition and a face-to-face condition. In addition, to robustly test the superiority of the Sensei possession over conventional training methods, future works should experiment with more variety of tasks.

The third limitation is the lack of long-term experiments. Although our study's focus on short-term training outcomes provides valuable aspects about the use of the Sensei possession system, the long-term effects on skill development, motivation, and mindfulness will also be crucial when applying this system to habitual training activity. Therefore, future works should monitor the long-term effects by employing a more complex task that requires a longer period of training.

Lastly, our study failed to examine the mindfulness of the participants. Since the mindfulness of the trainees relates to the long-term quality of training such as burnout risk, we aimed to test whether the use of Sensei possession has any influence on it. However, we found in our analysis that the AMQ of Amemiya et al. (2015) was not suitable for our experiment. We assumed that participants might have been confused by some items of AMQ, which was originally designed for team sports or outdoor sports (e.g., "I get distracted even when my teammates and coaches are talking." "I draw attention to the natural sensations such as wind and sunlight that blow while playing."). Thus, to test the mindfulness of participants in this type of experimental setting, future works should arrange AMQ and establish a novel mindfulness scale for indoor sports training without teammates.

## 5.5 Future work

During the development of the Sensei possession system, we underwent three challenges that we expect future work would also

encounter: development of live feedback, reduction of headset usage time, and design of safety.

The development of live feedback required an effort because the commands of movements had to be both short and intuitive. We resorted to reducing the variety of commands and utilizing simple sound effects (only three types, two of which were the same sound but directionally presented), which Lei et al. (2022) reported to have less stress on working memory. Nevertheless, we had a negative comment from a participant specifically about the directional command as mentioned in the result. This phenomenon that commands with directional audio is not always precise is in line with the previous studies of navigation (Walker and Lindsay, 2006; Marston et al., 2007), which also reported that haptic feedback was equally precise as audio and offering "a choice of audio or vibrotactile signals" was desirable. Hence, the adaptation of haptic displays as we suggested in the subsection 5.3 would contribute to improving the proxy trainer's user experience, while the precision may not improve. Therefore, in case the precision of command delivery is more crucial for the training tasks, future works should also consider the use of a stronger intervention such as electrical muscle stimulation (Tamaki et al., 2010; Pfeiffer et al., 2015), which can realize or navigate a variety of body movements without depending on the user's attention.

The reduction of headset usage time became our problem as the contents of preliminary learning materials were all originally taught in a face-to-face setting in the Aikido club, which we found easier to deliver in 3D animation than 2D video. With the requirement of keeping the XR head usage time at a minimum to prevent neck aches, we had to prioritize the contents for the role of the proxy trainer by discussing repeatedly with the Aikido club instructor and chose to deliver only the most important part of attacking movement via the XR headset, leaving the rest of knowledge to deliver via 2D video. However, future works could solve this problem by waiting for the development of lighter XR headsets, or relieve this problem by utilizing a stationary XR display such as Cave automatic virtual environment (Cruz-Neira et al., 1992) that Chen et al. (2019) reported to provide a better learning experience than 2D display while keeping the discomfort less than a headset.

Lastly, the safety of participants was also our concern because we expected inexperienced participants could accidentally hurt themselves from their lack of control when performing unfamiliar movements. Although we carefully chose the least dangerous training tasks from the paired Aikido training, we consulted many times with the Aikido club instructor, who had rich experience in teaching beginners and supported us in listing the possible ways injuries could occur. Based on this list, we designed the experiment environment to be as safe as possible and also included words of caution in the instructions for the participants that helped participants not to unnecessarily perform risky movements. However, if future works are to adopt more complex training tasks with higher risk, the experimenters should consider the use of tapings, protective gears, and an appropriate mat that can reduce the damage to the participants when they collide with each other or fall on the floor (Rainey, 2009; Pocecco et al., 2013).

## 6 Conclusion

In our study, we tested the use of a remotely controlled proxy trainer for the first time in training for a contact sport, Aikido. We found that the Sensei possession system with live feedback on proxy trainers can positively influence trainees to improve the success rate of performing evasive movements at the right time. Additionally, we found that the participants' intrinsic motivations for the training increased along with the perceived challenge and excitement. In the future, we plan to develop our live feedback interface for the proxy trainer further and test the effectiveness of Sensei possession with more complex training tasks over a longer period.

## Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

## Ethics statement

The studies involving humans were approved by the ethics committee of the Graduate School of Information Science and Technology, the University of Tokyo. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## Author contributions

TS: Writing–review and editing, Writing–original draft, Visualization, Validation, Project administration, Methodology, Investigation, Formal Analysis, Data curation, Conceptualization. AU: Writing–review and editing, Validation, Supervision, Formal Analysis. TNk: Writing–review and editing, Supervision, Methodology, Formal Analysis, Conceptualization. TNr: Writing–review and editing, Supervision, Resources, Funding acquisition. TA: Writing–review and editing, Supervision, Software, Resources, Methodology, Conceptualization. HK: Writing–review and editing, Supervision, Resources, Project administration, Methodology, Funding acquisition.

## Funding

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frvir.2024.1392635/full#supplementary-material

## References

Afana, J., Marshall, J., and Tennent, P. (2023). *Designing bodily materialistic representation for effective communication within virtual reality multimedia.* CHI' 23 Workshop - Body x Materials.

Aikido Yoshinkan (2020). *Dojocho noriki greeting all ayf members.* Available at: https://www.yoshinkan.net/cgi-bin/osirase_e/cgi/msgs.cgi?do=jwd&wn=1587712188 (Accessed January, 2024).

Aikikai Foundation (2014). *Instructors.* Available at: http://www.aikikai.or.jp/eng/information/leader.html (Accessed December 20, 2023).

Aikikai Foundation (2023). Aikido academy. Available at: http://www.aikikai.or.jp/eng/information/school.html (Accessed December 20, 2023).

Akehurst, S., and Oliver, E. J. (2014). Obsessive passion: a dependency associated with injury-related risky behaviour in dancers. *J. sports Sci.* 32, 259–267. doi:10.1080/02640414.2013.823223

Amemiya, R., and Sakairi, Y. (2019). The role of mindfulness in performance and mental health among Japanese athletes: an examination of the relationship between alexithymic tendencies, burnout, and performance. *J. Hum. Sport Exerc.* 14, 456–468. doi:10.14198/jhse.2019.142.17

Amemiya, R., and Sakairi, Y. (2021). Relationship between mindfulness and cognitive anxiety-impaired performance: based on performance evaluation discrepancies. *Asian J. Sport Exerc. Psychol.* 1, 67–74. doi:10.1016/j.ajsep.2021.06.001

Amemiya, R., Yusa, Y., and Sakairi, Y. (2015). Development of a mindfulness questionnaire for athletes. *Jpn. J. cognitive Ther.* 8, 106–115.

Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., and Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment* 13, 27–45. PMID: 16443717. doi:10.1177/1073191105283504

Bridgeman, L., Volino, M., Guillemaut, J.-Y., and Hilton, A. (2019). "Multi-person 3d pose estimation and tracking in sports," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 0–0.

Burazerović, M. (2011). Similarities and differences between various aikido and real aikido techniques. *Sci. Pract.* 115.

Campbell, J. I., and Thompson, V. A. (2012). Morepower 6.0 for anova with relational confidence intervals and bayesian analysis. *Behav. Res. methods* 44, 1255–1265. doi:10.3758/s13428-012-0186-0

Chen, X., Chen, Z., Li, Y., He, T., Hou, J., Liu, S., et al. (2019). Immertai: immersive motion learning in vr environments. *J. Vis. Commun. Image Represent.* 58, 416–427. doi:10.1016/j.jvcir.2018.11.039

Cheng, L-P., Roumen, T., Rantzsch, H., Köhler, S., Schmidt, P., Kovacs, R., et al. (2015). "Turkdeck: physical virtual reality based on people," in *Proceedings of the 28th annual ACM Symposium on user interface software and Technology* (New York, NY, United States: Association for Computing Machinery), 417–426. doi:10.1145/2807442.2807463

Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., and Hart, J. C. (1992). The cave: audio visual experience automatic virtual environment. *Commun. ACM* 35, 64–72. doi:10.1145/129888.129892

Faridan, M., Kumari, B., and Suzuki, R. (2023). "Chameleoncontrol: teleoperating real human surrogates through mixed reality gestural guidance for remote hands-on classrooms," in *Proceedings of the 2023 CHI conference on human factors in computing systems*, 1–13.

Figueiredo, A. A., Błach, W., Bujak, Z., Maroteaux, R. J., and Cynarski, W. J. (2020). Martial arts tourism of the "europe—far east" direction, in the opinion of grand masters. *Sustainability* 12, 8857. doi:10.3390/su12218857

Fribourg, R., Ogawa, N., Hoyet, L., Argelaguet, F., Narumi, T., Hirose, M., et al. (2021). Virtual co-embodiment: evaluation of the sense of agency while sharing the control of a virtual body among two individuals. *IEEE Trans. Vis. Comput. Graph.* 27, 4023–4038. doi:10.1109/tvcg.2020.2999197

Fujita, T. (2011). Preliminary study of passion in sport (in Japanese) [supotsu ni okeru jounetsu no yobiteki kentou]. *Bull. Fac. Educ. Kagoshima Univ. Cult. Soc. Sci.* 63, 81–87.

Hayashibe, K., and Amenomiya, M. (2008). "Recognition and transmission process of skill acquirement in the Japanese traditional crafts and modern factories," in *International journal of psychology* (27 CHURCH RD, HOVE BN3 2FA, EAST SUSSEX, ENGLAND: PSYCHOLOGY PRESS), 43, 314.

Hou, L., Wang, X., Bernold, L. E., and Love, P. E. D. (2013). Using animated augmented reality to cognitively guide assembly. *J. Comput. Civ. Eng.* 27, 439–451. doi:10.1061/(ASCE)CP.1943-5487.0000184

Ikuta, K. (2000). What are the implications of the teaching and learning method of traditional Japanese artistic performances? *Bild. Und Erzieh.* 53, 429–440. doi:10.7788/bue.2000.53.4.429

Ikuta, K. (2008). "The role of "craft language" in learning "waza"," in *Cognition, communication and interaction: transdisciplinary perspectives on interactive Technology*, 405–414.

Kodama, D., Mizuho, T., Hatada, Y., Narumi, T., and Hirose, M. (2023). Effects of collaborative training using virtual co-embodiment on motor skill learning. *IEEE Trans. Vis. Comput. Graph.* 29, 2304–2314. doi:10.1109/tvcg.2023.3247112

Kohn, T. (2011). "Ownership and appropriation," in *chap. Appropriating an Authentic Bodily Practice from Japan: on 'being there', 'having been there' and 'virtually being there'.* 1st ed. (London, UK: Routledge), 65–85.

Lave, J., and Wenger, E. (1991). *Situated learning: legitimate peripheral participation.* Cambridge, UK: Cambridge University Press.

Lei, Z., Ma, S., Li, H., and Yang, Z. (2022). The impact of different types of auditory warnings on working memory. *Front. Psychol.* 13, 780657. doi:10.3389/fpsyg.2022.780657

Lu, Y., Leng, X., Yuan, H., Jin, C., Wang, Q., and Song, Z. (2024). Comparing the impact of personal trainer guidance to exercising with others: determining the optimal approach. *Heliyon* 10, e24625. doi:10.1016/j.heliyon.2024.e24625

Marraffino, M. D., Schroeder, B. L., Fraulini, N. W., Van Buskirk, W. L., and Johnson, C. I. (2021). Adapting training in real time: an empirical test of adaptive difficulty schedules. *Mil. Psychol.* 33, 136–151. doi:10.1080/08995605.2021.1897451

Marston, J. R., Loomis, J. M., Klatzky, R. L., and Golledge, R. G. (2007). Nonvisual route following with guidance from a simple haptic or auditory display. *J. Vis. Impair. Blind.* 101, 203–211. doi:10.1177/0145482x0710100403

Martin-Niedecken, A. L., Rogers, K., Turmo Vidal, L., Mekler, E. D., and Márquez Segura, E. (2019). "Exercube vs. personal trainer: evaluating a holistic, immersive, and adaptive fitness game setup," in *Proceedings of the 2019 CHI conference on human factors in computing systems*, 1–15.

Matsui, K. (2014). Students intrinsic motivation (in Japanese) koukou undoubukatsudou ni okeru seito no naihatsuteki doukiduke. *Jpn. J. sport Psychol. advpub*, 2013–1301. doi:10.4146/jjsposy.2013-1301

Meyer, M., Molle, A., Bowman, P., and Judkins, B. (2021). Martial arts in the pandemic. *Martial Arts Stud.* 0, 7–31doi. doi:10.18573/mas.134

Misawa, K., and Rekimoto, J. (2015). "Chameleonmask: embodied physical and social telepresence using human surrogates," in *Proceedings of the 33rd annual ACM conference extended abstracts on human factors in computing systems*, 401–411.

Pelletier, L. G., Tuson, K. M., Fortier, M. S., Vallerand, R. J., Briere, N. M., and Blais, M. R. (1995). Toward a new measure of intrinsic motivation, extrinsic motivation, and amotivation in sports: the sport motivation scale (sms). *J. sport Exerc. Psychol.* 17, 35–53. doi:10.1123/jsep.17.1.35

Petrie, H. G., and Oshlag, R. S. (1993). *Metaphor and thought.* Cambridge University Press, 579–609. chap. Metaphor and learning.

Pfeiffer, M., Dünte, T., Schneegass, S., Alt, F., and Rohs, M. (2015). "Cruise control for pedestrians: controlling walking direction using electrical muscle stimulation," in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 2505–2514.

Pocecco, E., Ruedl, G., Stankovic, N., Sterkowicz, S., Del Vecchio, F. B., Gutiérrez-García, C., et al. (2013). Injuries in judo: a systematic literature review including suggestions for prevention. *Br. J. sports Med.* 47, 1139–1143. doi:10.1136/bjsports-2013-092886

Polanyi, M. (1966). *The tacit dimension*. Garden City, NY: Doubleday.

Rainey, L. C. E. (2009). Determining the prevalence and assessing the severity of injuries in mixed martial arts athletes. *North Am. J. sports Phys. Ther. NAJSPT* 4, 190–199.

Rauschnabel, P. A., Felix, R., Hinsch, C., Shahab, H., and Alt, F. (2022). What is xr? towards a framework for augmented and virtual reality. *Comput. Hum. Behav.* 133, 107289. doi:10.1016/j.chb.2022.107289

Revelle, W. (2019). *psych: procedures for psychological, psychometric, and personality research r package*. Evanston, IL: Northwestern University.

Serhan, D. (2020). Transitioning from face-to-face to remote learning: students' attitudes and perceptions of using zoom during covid-19 pandemic. *Int. J. Technol. Educ. Sci.* 4, 335–342. doi:10.46328/ijtes.v4i4.148

Sharma, N., Doherty, I., and Dong, C. (2017). Adaptive learning in medical education: the final piece of technology enhanced learning? *Ulster Med. J.* 86, 198–200.

Stepanyuk, S., Serputko, O., Veselovskii, A., Horodinska, I., and Kharchenko-Baranetska, L. (2022). *Aikido seminar as a universal form of training and motivation for athletes of all levels of mastery*. Scientific journal National Pedagogical Dragomanov University, 24–27.

Suzuki, T., Kodama, D., Nakamura, T., Narumi, T., Amemiya, T., and Kuzuoka, H. (2022). "The proposal of sensei possession and practice of workshop (in Japanese) [kyoushi no hyoshintai wo mochiita enkaku ginou denshou no teian to wakushoppu no jissen]," in Proceedings of the 27th virtual reality Society of Japan conference *(the virtual reality society of Japan)*, 1–4.

Takita, H., Hatada, Y., Narumi, T., and Hirose, M. (2023). "Effects of virtual co-embodiment on declarative memory-based motor skill learning," in *ACM symposium on applied perception 2023* (New York, NY, USA: Association for Computing Machinery). SAP '23. doi:10.1145/3605495.3605797

Tamaki, E., Miyaki, T., and Rekimoto, J. (2010). "Possessedhand: a hand gesture manipulation system using electrical stimuli," in *Proceedings of the 1st augmented human international conference*, 1–5.

Tanaka, Y., Nakamura, T., and Koike, H. (2022). Posture-based golf swing instruction using multi-modal feedback. *J. Inf. Process.* 30, 107–117. doi:10.2197/ipsjjip.30.107

Vallerand, R. J., Blanchard, C., Mageau, G. A., Koestner, R., Ratelle, C., Léonard, M., et al. (2003). Les passions de l'ame: on obsessive and harmonious passion. *J. personality Soc. Psychol.* 85, 756–767. doi:10.1037/0022-3514.85.4.756

Walker, B. N., and Lindsay, J. (2006). Navigation performance with a virtual auditory display: effects of beacon sound, capture radius, and practice. *Hum. factors* 48, 265–278. doi:10.1518/001872006777724507

Wang, L., Gai, W., Jiang, N., Chen, G., Bian, Y., Luan, H., et al. (2024). Effective motion self-learning genre using 360° virtual reality content on mobile device: a study based on taichi training platform. *IEEE J. Biomed. Health Inf.*, 1–12. doi:10.1109/jbhi.2024.3371470

Wulf, G., Shea, C., and Lewthwaite, R. (2010). Motor skill learning and performance: a review of influential factors. *Med. Educ.* 44, 75–84. doi:10.1111/j.1365-2923.2009.03421.x

Yang, U., and Kim, G. J. (2002). Implementation and evaluation of "just follow me": an immersive, vr-based, motion-training system. *Presence* 11, 304–323. doi:10.1162/105474602317473240

Yogi, Y. (2012). Effects of physical education classes using a judo lesson planning manual on motivation to learn (in Japanese) [judo jugyo dukuri kyouhon wo mochiita taiiku jugyou ga gakushuu iyoku ni oyobosu eikyou]. *Res. J. Budo* 45, 135–142. doi:10.11214/budo.45.135