# Using virtual humans and computer animations to learn complex motor skills: a case study in karate

Anne-Marie Burns[*†]   Richard Kulpa[*†]   Annick Durny[*]   Bernhard Spanlang[‡]   Mel Slater[‡§]

Franck Multon[*†]

*(\*) M2S Lab., UFR APS, U. of Rennes 2 - ENS Cachan, Av. Charles Tillon, 35044 Rennes, France*
*(†) INRIA MimeTIC project, Campus de Beaulieu, 35042 Rennes, France*
*(‡) Event Lab, Facultat de Psicologia, Universitat de Barcelona, Campus de Mundet - Edifici Teatre, Passeig de la Vall d'Hebron 171, 08035 Barcelona, Spain*
*(§) Institucio Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys, 23,08010 Barcelona, Spain*
*E-mail: annema.burns@gmail.com, rkulpa@irisa.fr, annick.durny@uhb.fr, bspanlang@gmail.com, melslater@ub.edu, fmulton@irisa.fr*

## Abstract

*Learning motor skills is a complex task involving a lot of cognitive issues. One of the main issues consists in retrieving the relevant information from the learning environment. In a traditional learning situation, a teacher gives oral explanations and performs actions to provide the learner with visual examples. Using virtual reality (VR) as a tool for learning motor tasks is promising. However, it raises questions about the type of information this kind of environments can offer. In this paper, we propose to analyze the impact of virtual humans on the perception of the learners. As a case study, we propose to apply this research problem to karate gestures. The results of this study show no significant difference on the after training performance of learners confronted to three different learning environments (traditional group, video and VR).*

## 1. Introduction

Learning a morphokinetic gesture is a complex task requiring to put in place motor, cognitive, affective and sensorimotor processes. In order to learn a new motor skill, one should first understand it, visualize and build its internal representation, and then, repeat it throughout many hours of practice. A typical manner to learn motor skills is to follow the examples and advices of an expert of that skill.

In order to determine if learners can acquire proper motor skills in a given environment, one should first determine if they can pick-up all the necessary information from it. This is particularly important in the case of virtual reality where the teacher is a model-based animation. In fact, even if animated with motion capture data of a real teacher, the virtual teacher might exhibit some artifacts that will alter the quality of the information provided to the learner.

In this paper, we pose the question of investigating if virtual teacher animated with motion capture data can actually provide correct visual information to learners. We apply this analysis to karate since it is based on standardized and complex motions. Hence, we compare how learners improve their performance when participating in traditional group lessons, lessons based on videos of a teacher, and lessons based on watching a virtual humans create from motion capture of that same teacher.

## 2. Related Works

The question of learning complex sportive motor skill in VR has already been explored. In 2003, Chua et al. [**?**] tried to determine which type of visual feedbacks could positively contribute to the learning process. Their work was applied to Tai Chi forms. Subjects were immersed throughout the use of a head mounted display (HMD) and were asked to watch a virtual teacher in order to mimic Tai Chi gestures. Five

different types of visual feedback were experimented, but the project failed to demonstrate any advantages of one over another.

In a later works, Patel et al. [?] claimed that VR is undoubtedly a superior learning tool compared to classical monoscopic videos. However, as it is the case in Chua et al. [?], the task of the participants was to mimic a gesture instead of learning it. In either cases, no lesson was proposed to the participants.

VR systems are generally based on virtual humans. However, several studies have shown that technical choices, such as graphical quality [?], may have an impact on how people perceive the virtual humans. Furthermore, these studies have shown that virtual humans animated using motion capture data convey relevant information for an already trained goalkeeper to react as in real duel with an opponent.

However, to our knowledge, there is no study aiming at analyzing if virtual humans can be efficiently used to learn complex motor skills. In this paper, we propose to address this point by designing three different teaching conditions and evaluating if using virtual teachers lead to different levels of improvements when compared to traditional lessons with a physical teacher or videos.

## 3. Methodology

### 3.1. Participants

For the purpose of our study, 30 sports sciences students aged between 18 and 25 years old were randomly divided into three groups, namely a traditional, a video-based and a VR-based group. All learners were practicing more than two hours of sports per week, but none of them had ever practice karate or a similar martial art.

### 3.2. Gestures

The subjects were asked to train on three basic karate gestures, namely, one frontal punch at the torso level (oi tsuki chunan), one frontal kick (mae geri) and one defense using the forearm in an external to internal motion that could be used to stop a torso level attack (soto uke). All three gestures were performed in a natural static standing position (hachiji dachi). Each of the gestures was choosen for its specific difficulties or skills.

Tsuki requires synchronization of the two arms. While one retires in armed position (hikite) the other one punches synchronously. Both wrists should rotate at the same time at the very end of the gesture. The trajectory of the punch should follow a perfectly linear path and aim at the plexus.

Mae geri is a frontal kick requiring balance in order to alternate both legs. The motion is in three phases (figure **??**). First, the knee should be elevated over the belt. Then, the leg should be extended in a rapid whipped
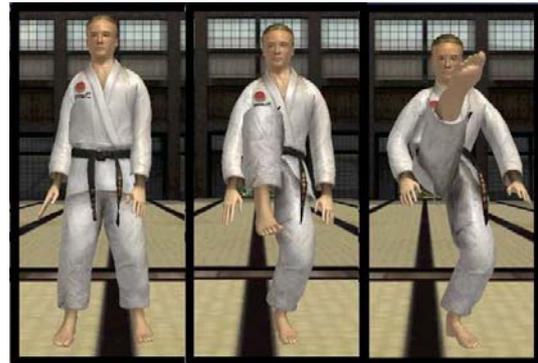


Figure 1: Three phases of the mae geri kick: (a) Standing position (b) Elevate the knee (c) Kick and go back to (b)

motion. Finally, the leg should be retracted quickly to the first position before returning to the standing position.

Soto uke is certainly the most difficult of the three gestures in that it requires a lot of coordination. Most subjects failed to understand it at the initial evaluation and took more than one lesson to get a gross idea of it. One arm starts at the ear level and sweeps in front of the torso in order to clear the area and push to the side an eventual tsuki punch. At the same time, the other arm retracts in armed position. As for the tsuki, there is a rapid rotation of both wrists at the very end of the gesture.

### 3.3. Initial Evaluation

At the beginning of the study, the starting level of the learners was assessed by performing a simple test. Each of the three karate gestures was presented individually once on a video sequence of ten repetitions performed by an expert. The learners then had to try to reproduce these by performing two sequences of ten repetitions of each gesture. These sequences were recorded both on standard video and with a 3D full body motion capture system. This initial evaluation served two purposes. First, it ensures that none of the learners already know the gestures and, second it establishes the starting level of the learners. The starting level was evaluated by a karate expert who attributed a score to each learner by watching their performances on video.

### 3.4. Training

Starting the subsequent week, all learners had to fulfill three weeks of training. The weekly training consisted of a one hour session divided in 2 parts. The first part was a 15 minutes standardized warm-up recorded on video by a karate expert. After this warm-up, each

Figure 2: Traditional group learning environment



Figure 4: VR learning environment



Figure 3: Video learning environment

group proceeded with the 3 gestures training in their attributed learning environment. The traditional group worked with a karate expert teaching them as shown in figure **??**. The video group watched a prerecorded lesson with the same expert but on video (figure **??**). The virtual reality group trained in an immersive environment where a virtual teacher was performing a prerecorded lesson. In that environment, the motion of the head of the learner was tracked to adapt his viewpoint. Stereovision was used to perceive the 3D environment. The virtual environment was reduced to its simplest expression in order to evaluate the impact of the avatar only. The virtual world was a representation of a japanese dojo and the master was a representation of the real master himself wearing the traditional karate suit (gi) built from photos (figure **??**).

During this training, each of the three gestures were taught for 15 minutes divided in two sections. The first eight minutes or so were dedicated to explanations of the gesture and the remaining part of the 15 minutes was dedicated to exercises related to the gesture. The teaching was designed in a pedagogical progression where each gesture explanations and exercises were divided in phases in which the gesture was decomposed and analyzed.

All groups followed the same lesson but in different conditions. Indeed, the lessons were concurrently cap-

tured on video and in 3D. The lesson was the same for the three weeks. For the traditional group, the teacher was asked to follow the lesson as rigorously as possible. Obviously, in that context it is impossible to strictly avoid adaptation of the content and interaction with the class. In our experiment, the traditional group served as a control group to evaluate to what extent the three gestures could be learned.

### 3.5. Final Evaluation

In the week following the last training, the learners were evaluated in a similar manner as at the beginning. As it was the case for the initial evaluation, the learners were evaluated outside of their learning environment. For the final evaluation, the experimenter only calls the name of each gesture and the learner performs it in 2 series of ten repetitions. These repetitions were captured both on video and in 3D full body motion capture and evaluated by the expert.

## 4. Results

In order to evaluate the performance of the learning environments we statistically analyzed the evolution of the scores between the first and the second evaluation for the three groups. In a first time, we performed a two-ways ANOVA (group and performance) with repetition. No statistical difference was found between the groups for the initial evaluation for the tsuki ($p=0.887$) and the soto uke ($p=0.217$). However, the VR group had a significantly different score for the mae geri ($p=0.022$). That can be explained by the fact that one of the learners of that group already score high at the inital evaluation compared to the other learners. Despite this limitation, the statistical test confirms that all the groups were similarly beginners in karate.

Moreover, the same test demonstrated that there was a significant difference between the scores of the initial and final evaluation ($p<0.001$). It clearly demonstrates that all groups have improved their performance on the

three gestures. Finally, a one-way ANOVA (groups) was used to see if that improvement occurred in the same manner in all three groups. There was no significant difference between the groups in improvement for tsuki (p=0.887), mae geri (p=0.074) and soto uke (p=0.217). That means that there is no significant effect of the experimental set-up on the score improvement.

In conclusion, these results tend to show that the virtual teacher was as efficient as the physical or video teacher to deliver relevant information. Learners were equally able to improve their performance troughout learning indifferently of the learning environment they used. Therefore, VR can be used as a learning tool for that kind of complex motions.

## 5. Conclusions

Our contributions are in the creation of a virtual karate lesson based on pedagogical premises and of an experimental protocol to evaluate learning in different environments. This protocol permitted to isolate the impact of a model-based avatar in the learning process and to evaluate it.

In fact, one might argue that the tasks presented in the Chua et al. [?] and Patel et al. [?] were more mimicking tasks than learning tasks. In the first case, the users only performed 12 repetitions of each gestures and were evaluated on the fourth lasts. In the second case three gestures were shown in a one hour experiment including an outside environment evaluation. In neither case oral instructions nor progressive exercises about the gestures were proposed to the users.

In our case, oral explanations are provided to the learners together with exercises demonstrating the gestures in progressive learning phases. The completed gestures are repeated in 4 series of 10 repetitions at two different paces and in 6 to 8 series of 10 repetitions of exercises during every lesson. The entire experiment lasted five weeks since all training and evaluations are performed on a weekly basis. That more closely corresponds to the definition of learning.

Since all groups equally performed on the final evaluation, it is possible to conclude that the virtual teacher has no negative impact on learning. From that premise, it is possible to continue to study VR environments as new learning tools. Future steps of this project will consist on studying different feedback interfaces and their impact on the learning both in term of presence and motivation to learn in the environment and in term of quality of the learned gesture and its possible transfer in a real sportive situation. Emotional states, internal learning process, perception and feelings of the learners confronted to different learning environments will also be studied throughout the use evocation interviews.

## References

[1] P. T. Chua, R. Crivella, B. Daly, N. Hu, R. Schaaf, T. Camill, J. Hodgins, and R. Pausch. Tai chi: Training for physical task in virtual environments. In *VR '03: Proceedings of the IEEE Virtual Reality Conference 2003*, page 87, Washington, DC, USA, 2003. IEEE Computer Society.

[2] K. Patel, S.-H. Jung, R. Diankov, R. Bajcsy, and J. N. Bailenson. The effects of fully immersive virtual reality on the learning of physical task. In *Proceedings of the 9th Anual International Worshop on Presence*, 2006.

[3] N. Vignais, R. Kulpa, C. Craig, S. Brault, F. Multon, and B. Bideau. Influence of the graphical levels of detail of a virtual thrower on the perception of the movement. *Presence: Teleoperators and Virtual Environments*, 19(3):243–252, 2010.