A 3D Descriptive Model for Designing Multimodal Feedbacks in any Virtual Environment for Gesture Learning

Djadja Jean Delest Djadja[a], Ludovic Hamon[b] and Sébastien George[c]
LIUM, Le Mans University, Le Mans, France

Keywords: Virtual Reality, Pedagogical Feedback, Design, Motion Learning.

Abstract: This paper addresses the problem of creation and re-usability of pedagogical feedbacks, in Virtual Learning Environments (VLE), adapted to the needs of teachers for gesture learning. One of the main strengths of VLE is their ability to provide multimodal (*i.e.* visual, haptic, audio, etc.) feedbacks to help the learners in evaluating their skills, the task progress or its good execution. The feedback design strongly depends on the VLE and the pedagogical strategy. In addition, past studies mainly focus on the impact of the feedback modality on the learning situation, without considering others design elements (*e.g.* triggering rules, features of the motion to learn, etc.). However, most existing gesture-based VLEs are not editable without IT knowledge and therefore, failes in considering the evolution of pedagogical strategies. Consequently, this paper presents the GEstural FEedback EDitor (GEFEED) allowing non-IT teachers to create their multimodal and pedagogical feedbacks into any VLE developed under Unity3D. This editor operationalises a three dimensional descriptive model (i.e. feedback virtual representation, its triggering rules, involved 3D objects) of a pedagogical feedback dedicated to gesture leaning. Five types of feedbacks are proposed (i.e. visual color or text, audio from a file or a text and haptic vibration) and can be associated with four kinds of triggers (i.e. time, contact between objects, static spatial configuration, motion metric). In the context of a dilution task in biology, an experimental study is conducted in which teachers generate their feedbacks according to pre-defined or chosen pedagogical objectives. The results mainly show : (a) the acceptance of GEEFED and the underlying model and (b), the most used types of modalities (*i.e.* visual color, vibration, audio from text), triggering rules (*i.e.* motion metric, spatial configuration and contact) and (c), the teacher satisfaction in reaching their pedagogical objectives.

1 INTRODUCTION

Virtual Environments (VEs) dedicated to learning technical gestures has been used in many fields such as sports, health, biology, *etc.*(Lee and Lee, 2018; Le Naour et al., 2019; Mahdi et al., 2019). One of the main advantages of Virtual Learning Environments (VLEs) lies in their multimodal (*i.e.* visual, audio, haptic, etc.) feedbacks characterizing the dynamic of performed motions as well as assisting the learners in their self-evaluation, the task progress and/or its good execution.

For example, (Le Naour et al., 2019) show a virtual avatar reproducing the throw of a rugby ball of the teacher. (Wei et al., 2015) used the same principles for physical therapy, and added a textual guidance (*e.g.* "Arm too high, too slow"). (Lin et al., 2018) colored the skeleton bones of the 3D avatar to show the error in the plantar pressure during a Tai-chi exercise. A prerecorded voice can describe the next motion to perform for cooking (Mizuyama, 2010), while a brief sound can indicate the grabbing or the assembly of an object in industry (Chen et al., 2019). Furthermore, haptic feedbacks can generate an attractive force to reorient motions for calligraphy learning (Nishino et al., 2011) or for surgery (Wang et al., 2018).

Gesture learning is contextual to the task and the teaching strategy as this expression can be related to: i) the observation and imitation of successive postures (ii), the learning of a sequence of actions or (iii) the building of a motion respecting geometric, kinematic or dynamic features (Larboulette and Gibet, 2015; Djadja et al., 2020). The vast majority of the past gesture-based VLEs are, therefore, specific to the task to learn, with no or few functionalities in terms of feedback creation or edition.

However, designing efficient feedbacks is not triv-

[a] https://orcid.org/0000-0001-7923-0690
[b] https://orcid.org/0000-0002-3036-0854
[c] https://orcid.org/0000-0003-0812-0712

Djadja, D., Hamon, L. and George, S.
A 3D Descriptive Model for Designing Multimodal Feedbacks in any Virtual Environment for Gesture Learning.
DOI: 10.5220/0012081000003538
In *Proceedings of the 18th International Conference on Software Technologies (ICSOFT 2023)*, pages 84-95
ISBN: 978-989-758-665-1; ISSN: 2184-2833

ial. Indeed, (Sigrist et al., 2012) deeply investigated the impact of feedback modalities for motor learning in VE and Real Environment (RE). Those modalities must be carefully chosen according to the complexity of the task and the cognition abilities of the learners. Nonetheless a feedback cannot be reduced to its modality, and other design elements are crucial such as, its virtual representation, its triggering rule or the motion metric to monitor. For this last point, even if the teacher is involved in the metric initial choice, the efficiency of the evaluation system is not guaranteed (Senecal et al., 2002). Consequently, a system must be built to allow any teacher, without IT knowledge, to (re)design and (re)implement all feedback elements, to make them efficient and adapted to the learning situation. To our knowledge, no such a system exists, except the work of (Lo et al., 2019), limited to one VE not built for learning purposes.

This paper proposes a three-dimensional descriptive model (virtual representation, triggering rules, involved 3D objects) of a pedagogical feedback and its operationalization through the GEstural FEedback EDitor (GEFEED). This editor allows any teacher to create and integrate feedbacks, characterizing the performed technical gestures, in any VLE developed with the unity engine. Five kinds of feedbacks are available (*i.e.* visual color, visual text, audio from a file, audio from a text and haptic vibration) and can be associated with a set of four kind of triggering rules (*i.e.* time, contact between 3D objects, spatial configuration or threshold of a motion metric to reach).

Section 2 reviews the past studies regarding VLE for gesture learning, feedbacks, their design and their potential re-usability. The next section presents the 4 dimensional descriptive model. The architecture and Human Computer Interface (HCI) of GEFEED are described in section 4. A first VLE dedicated to the dilution in biology is considered in section 5, with the creation of a feedback example. Section 6 is dedicated to an experiment where the teachers must create their feedbacks with this VLE. The usability and usefulness of GEFEED, its main functonalities are put to the test. The results of the experiment are discussed in section 7 while perspectives end this paper.

2 RELATED WORKS

For learning gesture-based tasks or motor skills, various VLE have been built integrating real-time feedbacks to: guide learners in correcting their motions (Luo et al., 2011; Cannavò et al., 2018; Liu et al., 2020), following the protocol made of an action sequence (Mahdi et al., 2019; Mizuyama, 2010), performing a self-evaluation (Cannavò et al., 2018), improving engagement (Adolf et al., 2019) or enhancing the overall pedagogical experience (Mizuyama, 2010).

All those VLE are, by design, specific to their pedagogical and research objectives including the provided feedbacks. In this work a pedagogical feedback is considered as a pedagogical information, previously defined by a teacher, provided to the learner through a virtual representation, during the task or after it, to: (i) assist learners in the evaluation of the task (ii), its progression or (iii), guide them in its good execution. By defining the motion features of the 3D object to monitor (*e.g.* geometric, kinematic or dynamic features, collisions, etc.), triggering rules (*e.g.*, threshold for features, time step, etc.) and a virtual representation with which the pedagogical information will be conveyed (*e.g.* an arrow the motion must follow, a hand vibration to avoid reaching a dangerous area, etc.), a strategy for operationalising the pedagogical information is defined. However, given the learning context and the pedagogical objective, one can ask for the best design strategy of such a feedback.

(Sigrist et al., 2012) investigated the impact of feedback modalities (*i.e.* visual, haptic, audio, multimodal) for motor learning in VE and RE. Visual feedbacks are mainly used, intuitive and efficient. A first type of visual feedback relies on a color change, for example, of specific joints of the body to help in adjusting its position and orientation to learn tai-chi, *i.e.* green when the learner motion is close to the expert one, red otherwise (Liu et al., 2020). In addition, a textual score can be added that points out the body position (correct or not) to assesse the overall performance of a basketball throw (Cannavò et al., 2018). A last recurrent type of visual feedbacks is the replay of teachers' motions through a 3D avatar and the motion trajectory, displayed during or after the performance, to guide learners and for self-evaluation (Cannavò et al., 2018; Le Naour et al., 2019; Djadja et al., 2020).

Audio feedback can support visual ones as they are easily interpretable (Sigrist et al., 2012). A first strategy is to add recorded voices to displayed texts advising learners to, for example, handle a Chinese frying pan (Mizuyama, 2010) (*e.g.* "don't move your left wrist", "push the contents forward with the ladle"). A brief sound can also be heard for the completion of a good action or an inappropriate one. In the context of tenon structure training, (Chen et al., 2019) provided a collision sound when the hand touches a tenon part. A prompt, read by a prerecorded voice, can then deliver the related knowledge. Those feedbacks were added to a green light, appearing for 2 seconds, if the assembly of two parts is successful between the matching surfaces.

Haptic feedbacks, usually appropriate for navigation and orientation, need specific and sometimes costly hardware (Sigrist et al., 2012). They are often combined with visual and audio ones, to decrease the cognitive charge in a reinforcement way of the same pedagogical message. Otherwise, the risk of a cognitive overload is significant. One of the less costly and cumbersome strategies implies vibration motors. For example, the vibration intensity, implemented by (Adolf et al., 2019), represents the force of a ball for juggle learning. This feedback is completed with the volume and pitch of a specific sound and the display of the ball's trajectory. (Luo et al., 2011) used vibrotactile motors attached to body parts to correct user's wrong posture in yoga training. A yoga instructor gave audio instructions (*e.g.* "Left Arm Up") while a text displayed the same message and a red arrow pointed to the targeted body part.

Triggering rules, often defined by spatial and temporal conditions of 3D artefacts to monitor, are not studied in past studies to our knowledge. However, they can be guessed by considering the uses cases, and one can point out three categories: a) reaching thresholds of geometric, kinematic or dynamic features of 3d object or body parts (Luo et al., 2011; Liu et al., 2020) b), collisions between those artefacts (Adolf et al., 2019; Chen et al., 2019) and c) specific steps or times of the task to learn (Mizuyama, 2010; Cannavò et al., 2018; Adolf et al., 2019; Le Naour et al., 2019). However no tendencies can be highlighted in terms of frequency. To go further, feedback design is strongly contextual on several aspects (e.g. application domain, task, pedagogical objective and strategy, learner cognition abilities). The pedagogical strategy can strongly vary from one teacher to another for the same task. The re-use and adaptation of existing VLEs is therefore necessary for their sustainable integration and adoption in any curriculum.

In terms of re-engineering aspects, few VLE dedicated to gesture learning have edition or adaptation functionalities for the final users (*i.e.* teachers and learners), such as: i) the definition of the pedagogical scenarios as a sequence of predefined actions (Mahdi et al., 2019) and ii) the capture and replay of gestures to learn, if one does not consider the heavy process behind the motion capture (Le Naour et al., 2019). In our previous work, we proposed the MEVEL (Motion Evaluation in Virtual Environment for Learning) system with which teachers can record their own replayable gesture-based task to learn, and divide it into a set of ordered actions (Djadja et al., 2020). The textual feedbacks on motion features (i.e. speed, acceleration, jerk, Dynamic Time Warping) were not editable.

The strong context-dependent aspect of feedbacks make shard the definition of consensuses in term of efficient design principles. To address this problem, this study considers that a continuous design and implementation loop of feedbacks assisting the gesture learning, that integrates the teachers in all creation steps, enhance the efficiency of the VLE, given a pedagogical objective. However, there is no indication in past studies that teachers can modify or incorporate new feedbacks into existing VLEs without a re-engineering process requiring IT knowledge. One editor can be noted in the architecture domain, outside any learning context (Lo et al., 2019). The authors proposed the basis of an "action trigger" creation and edition system to help the final users in managing (*i.e.* add, edit, remove) their feedbacks. Nevertheless, the list of the possible triggers and actions is not clearly stated.

Research Positioning and Contributions. This paper proposes a descriptive feedback model and its implementation through the GEstural FEedback EDitor (GEFEED), to help non-IT teachers in creating their own feedbacks for learning gestures. This editor can add feedbacks in any VLE made with unity engine.

A first methodological contribution is the descriptive model of pedagogical feedback, tested and validated, from the acceptance, usefulness point of view and for reaching pedagogical objectives in a specific context. The model is generic and can be applied to any gesture-based task to learn in VE and extended to better formalized the pedagogical intentions (cf. section 7). The GEFEED system allowing any teachers, without IT skills, to design and implement multimodal feedbacks in an existing VLE made with Unity Engine, is a second technological contribution. Finally, the experimental study brings new data and tendencies in a specific frame, useful for the feedback design *i.e.* the most used: i) types of feedbacks belonging to a same modality, where most of the past studies focus only on the impact of modalities, (ii) triggers rules and (iii), their correlation with the feedback modalities. This last point is the third scientific contribution.

A three-dimensional model to design all steps of feedbacks is proposed in the next section.

3 A 3-DIMENSIONAL DESCRIPTIVE MODEL OF PEDAGOGICAL FEEDBACKS

Considering the strong contextual aspect of a gesture-based task and the variation of pedagogical strategies (Djadja et al., 2020), a system must be built to allow teachers defining all the functional elements of a feedback *i.e.*: (i) the temporal and spatial conditions of its triggering (ii), its virtual representation conveying the pedagogical information and (iii) the 3D artefacts implies in its virtual representation or its triggering rules. Figure 1(a) presents the UML diagram of a descriptive model made of a three main classes: *VRObject*, *Trigger* and *Feedback*.

A *VRObject* is a virtual body part of the user or a any other 3D object in VE, whose motions or states is monitored by the triggering rules. It can also be the element on which the virtual representation must be applied. A *Trigger* checks if the spatial and temporal conditions of the motion or state of a VRObject (except for the time trigger, cf. section 3.2) are met, to turn on the *Feedback*. A *Feedback* is a visual, audio or haptic representation of the pedagogical information to convey. A *Feedback* can be attached to one or several *Triggers* and a *Triggers* is linked to only one *Feedback*.

3.1 Feedbacks

The implemented *Feedback* are, “visual color”, “visual text”, “audio from a file”, “audio from a text” and “haptic vibration”.

The “visual color” consists in changing the color of the outline of the chosen *VRObject e.g.* the ball outline becomes yellow when it reaches the target.

Table 1: List of implemented *Feedbacks* and their parameters.

Types	Parameters
Visual text	Duration, text, position, orientation, size, 2D or 3D, “VRObject”
Visual color	Duration, color, “VRObject”
Audio from a file	Audio file
Audio from a text	Text
Haptic vibration	Duration, amplitude, frequency, device (*e.g.* HTC Vive controller)

The “visual text” consists in showing a text attached to a *VRObject* or in a predefined 3D position. The “audio from a file” and “audio from text” respectively read an audio recording provided by teachers or use a Text-To-Speech API [1] to read the provided text

The “haptic vibration” makes vibrating the specified compatible hardware device (*e.g.* controllers of the HTC Vive), using a predefined duration, amplitude and frequency. Table 1 resumes all *Feedbacks* and their parameters.

3.2 Triggers

The four considered *Triggers* are respectively named, “time trigger”, “contact trigger”, “spatial trigger” and “metric trigger”. The “time trigger” consists in starting the *Feedback* at each pre-defined time step.

The “contact trigger” starts the *Feedback* when two defined virtual 3D objects collide.

The “spatial trigger” turns on the *Feedback* when a 3D virtual object enters into a defined radius, the center being defined by the teacher or being the 3D position of a *VRObject*.

The “metric trigger” check if some thresholds or intervals of geometric or kinematic features are reached to trigger the “Feedback”. For this work, the following metrics are proposed: the horizontal/vertical orientation of a *VRObject*, the velocity and the jerk of its motion, the Dynamic Time Warping (DTW) to compare the shape of the trajectory of the teacher and the learner motion (*i.e.* the lower it score is the closer the two motions are (Djadja et al., 2020))

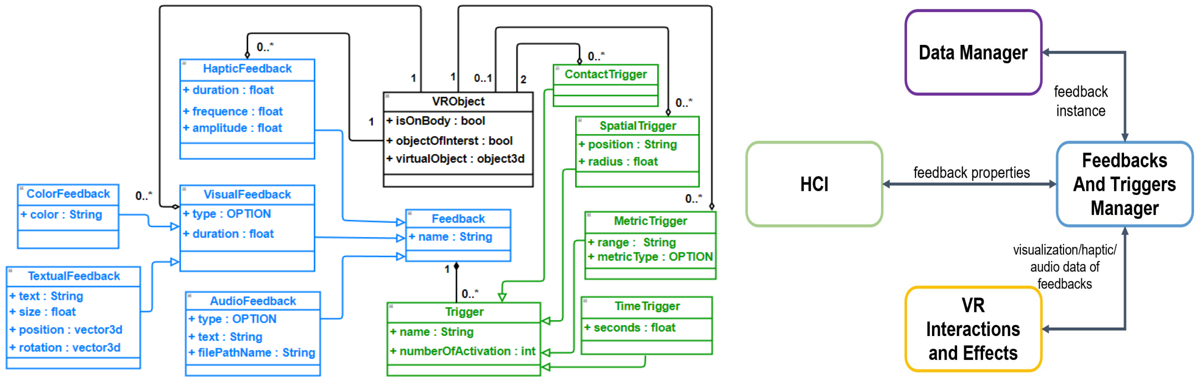
Table 2: List of considered *Triggers* and their parameters.

Types	Parameters
Time	Minutes, seconds
Contact	Two “VRObject”
Spatial	“VRObject” whose position is analysed, radius, 3d position of the center or 3d position of a “VRObject”
Metrics	“VRObject”, metric type (horizontal, vertical, velocity, jerk, DTW), its threshold or interval

All *Triggers* have a parameter that allows them to run indefinitely or a specific number of times. Table 2 resumes all the “Triggers” and their parameters.

All types of *Feedbacks* and their parameters were chosen according to those encountered during the review of past studies (cf. section 2). However, the motion capture and replay of body parts were not integrated as this kind of *Feedbacks* still requires an heavy processing chain, that non-IT teachers can ot handle

[1] Click here for more information on the Test-To-Speech API



(a) (b)

Figure 1: 3D Descriptive Model of pedagogical feedbacks (a) and GEFEED architecture (b).

nowadays. The *Trigger* types and their parameters are deduced from the use cases of past studies.

The next section operationalises the proposed model through the GEFEED system to allow teachers creating their own feedbacks in any VE developed with Unity engine.

4 GEstural FEedback EDitor

To configure pedagogical feedbacks in VE, a descriptive model is proposed, based on the dual use of visual, audio or haptic *Feedbacks* associated with one or several *Triggers*, the overall applied to *VRObjects* if required. The following sections describes the GEstural FEedback EDitor (GEFEED) through its architecture, the existing VLE integration, its main functionalities and HCI for the creation of gesture-based pedagogical feedbacks.

4.1 Architecture

This editor was implemented using Unity version 2019. Figure 1 (b), shows the architecture made of four modules: "Feedbacks and Triggers Manager (FTM), HCI, Data Manager (DM), VR Interactions and Effects (VRIE)". The FTM module (blue part, figure 1 (b)) represents an instance of the descriptive model (cf. 3). The instance is used in the HCI module (green part, figure 1 (b)) for building the pedagogical feedback in VLE *i.e.* selecting *VRObjects*, configuring the virtual representation, *Triggers* and their properties. The DM module (purple part, figure 1 (b)) store, in JSON files, the properties of the *Feedback* and their *Triggers*. The data are then used to load the *Feedbacks* in VE thanks to the VRIE module (yellow part, figure 1 (b)). This last module offers to users VLE interactions to choose the involved *VRObjects* (*e.g.* on which a color could be apply), set up the spatial properties (*e.g.* placing a text), navigate, show and test the feedback in VE.

4.2 Integration of an Existing VE

GEFEED can import any existing VE if this last one is exported as an assetbundle[2], a set of platform-specific non-code assets (*e.g.* 3d models, audios, images, etc.) that Unity can load at runtime (Djadja et al., 2020). However, the interaction scripts cannot be exported unless they are pre-built. Nevertheless their links to the 3D objects are still missing. To counterbalance this last issue, a solution relies on a "csv" file inventorying each link between the 3D objects and their interaction scripts, through a dedicated plugin coded for this work. This file is therefore made during the VE exportation by the plugin and read during the importation by GEFEED. A video is available [3] in which an exported VE and its importation in GEFEED thanks to our method is shown.

4.3 Interface

The GEFEED interface can been seen in figure 2. Figure 2 (a) and (b) are respectively used for *Feedbacks* and *Triggers* *i.e.* their creation, deletion and selection. They are identified by their unique names, and respectively configured with the menus in figure 2 (d) and (e). Figure 2 (c) allows saving and loading *Feedbacks* and their *Triggers*. It can also activate the 3D interaction system of the VRIE module, for interacting in the VE using a mouse and a keyboard (cf. section 4.1). With the three buttons of Figure 2 (f), teachers can respectively perform a translation, a rotation and

[2] Click here for more information on assetbundles

[3] Click here to access to videos of imported VLEs, created feedbacks, questionnaires and collected data

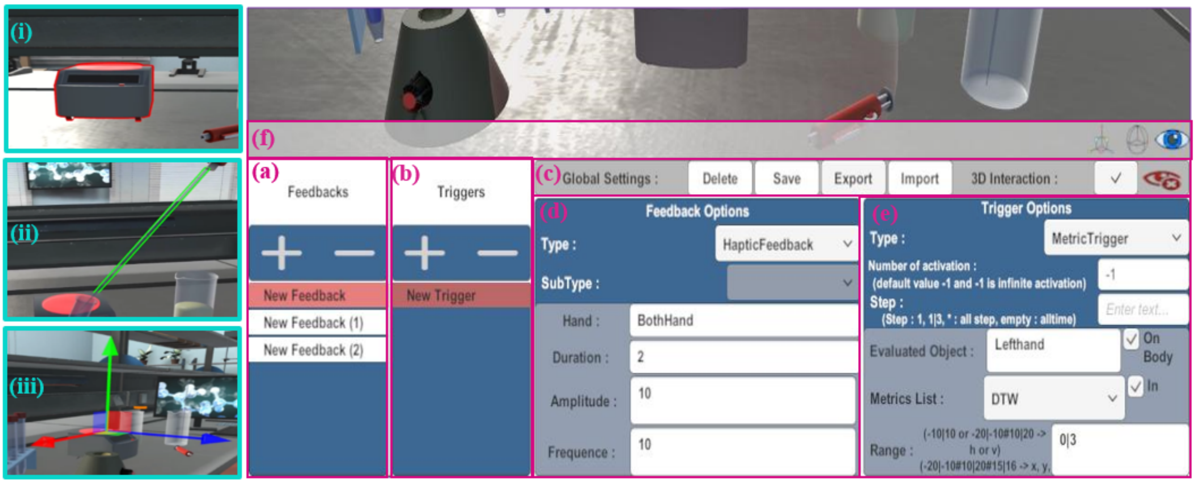


Figure 2: Human Computer Interface (HCI) of GEFEED.

chose among several angles of view of the *VRObject* (*i.e.* up, down, left and right). For example, the end user can define, in the VLE, the location and orientation of a text to display (*e.g.* figure 2 (iii)). Figures 2 (i) and (ii) show examples of "visual color" *Feedbacks*.

The next section presents the VLE considered for this study and an example of feedback designed and implemented thanks to GEFEED.

5 THE DILUTION VLE WITH A FEEDBACK EXAMPLE

In biology, dilution is the action of adding a liquid to a dangerous or unstable solution to lower its concentration before making an analysis. This process follows a strict protocol and mainly consists in getting a part og the initial solution in a test tube, and drop off this part in another test tube filled with water.

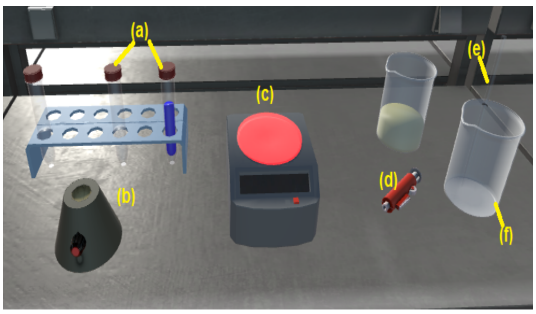


Figure 3: Example of a biology dilution VLE.

Making a dilution implies to have: one test tube containing the initial solution and the other one the diluted solution (figure 3 (a)), a tool to homogenize solutions (figure 3 (b)), an electric heat source to sterilize the tubes opening and the pipette extremity (figure 3 (c)), a rubber bulb attached to a pipette to get/release a part of the solution (figure 3 (d) and (e)) and a container (figure 3 (f)) to drop off the pipette once the task is done.

One of the requirements to perform such a task is the following: the user's hand holding the rubber bulb attached to the pipette, must not leave the sterile zone, *i.e.* a half sphere whose center is the electric heat source. Therefore, a "visual color" *Feedback* was created to change the color of the rubber bulb. This *feedback* is attached to a "spatial trigger" with a position at the center of the electric heat source and a radius defined by the teacher. Figures 4 and 5 respectively shows the setup describing this *Feedback* and *Trigger*.

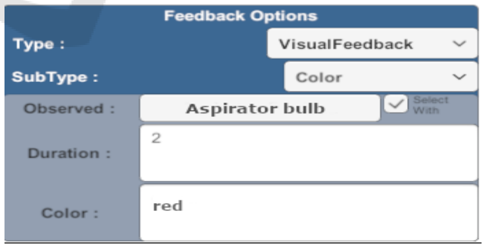


Figure 4: "Visual color" setup example.

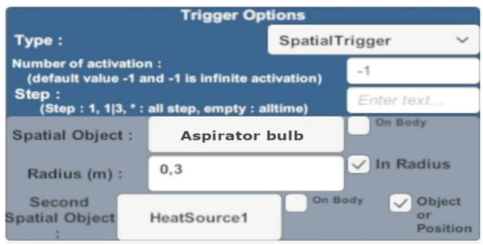
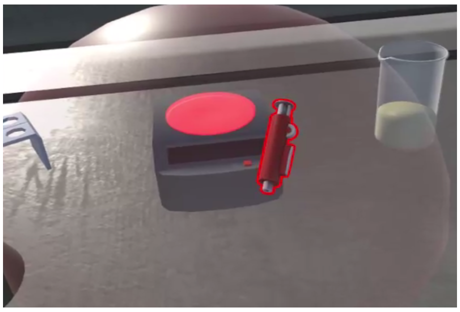
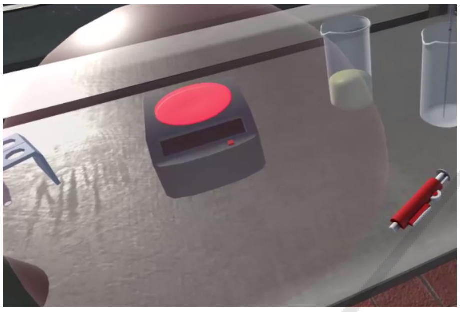


Figure 5: "Spatial trigger" setup example.

The figure 6, shows the rubber bulb when it is in the sterile zone and outside of it. A demonstration video can be found by following the footnote[3].



(a)



(b)

Figure 6: An example of a visual color feedback: the rubber bulb is in the sterile zone (a) or outside of it (b).

The validation, usefulness and acceptance of the GEFEED system and its descriptive model are studied in the next section.

6 EXPERIMENTAL STUDY

For this study the point of view of teachers is adopted as it is a first mandatory step to its acceptance, before evaluating the efficiency of VLEs for gesture learning, enhanced with teachers' feedbacks. The protocol made to evaluate the usability and the utility of GEFEED, in the specific context of dilution task, is explained in this section.

6.1 Protocol

Ten teachers in biology, aged from 26 to 54 years (average 42.8 years) were volunteers. Each of the participants received explanations about: (i) the VLE and its main functionalities (ii), the dilution process and (iii) GEFEED. For this last point a video[3] was provided showing an example of a feedback creation.

The two feedbacks add respectively a green color thanks to a "time trigger" (*i.e.* every 5 seconds) and yellow color thanks to a "spatial trigger" that measures the distance between the test tube and the heat source (*i.e.* when a test tube enters a radius of 0.5 meters from the heat source).

The main goal of this protocol is to create *Feedbacks* and *Triggers* to achieve the pedagogical objectives. The protocol, divided into three parts, aims to progressively familiarize teachers with GEFEED before letting them acting freely to reach their own objectives.

In the **first part (tutorial)**, the pedagogical objectives, *Feedbacks* and *Triggers* types and parameters are given. These three pedagogical objectives are: (a) learners must sterilize the opening of a tube containing the initial solution without being too close to the heat source, (b) hands must not touche themselves and (c) hand motions must be smooth.

For (a), a "haptic vibration" is applied on the left-hand controller, during 1 second with an amplitude of 10 Pa and frequency of 10 Hz. A "spatial trigger" must monitor the distance between the test tube with the initial solution and the heart source. The maximum radius between these two objects is 0.3 meters.

For (b), a "visual red color" is applied on both virtual hands representing each HTC Vive controllers, during 1 second. A "contact trigger" must monitor the collision between both hands.

For (c), an "audio from a text" is created and must read the following statement: "Take your time, slow down your motions". A "metric trigger" must monitor the jerk of the right-hand motion (Larboulette and Gibet, 2015). The activation threshold value must be adjusted between 50 and 100. The feedback must be activated twice.

The **second part (technical assessment)** provides the three following pedagogical objectives: (i) the test tube must be held vertically (ii), the user's hands be in front of the user at all time and (iii), the pipette must be held horizontally. Teachers must define their own *Feedbacks* and *Triggers*.

The **third part (pedagogical assessment)** proposes to the participants to define a maximum of three pedagogical objectives and create their associated *Feedbacks* and *Triggers*.

The first part acts as a tutorial, the second part tests if teachers can technically use GEFEED by giving them some operationalisable and pedagogical objectives. The last part assesses the GEFEED abilities in reaching teachers' pedagogical objectives. The third part is not mandatory and allows to see if the interest of teachers keep going regarding the use of GEFEED

The time was recorded and, once the protocol finished, the teachers completed questionnaires on the usefulness and the usability of GEFEED.

6.2 Results and Analysis

All participants completed the tutorial and technical parts, and 8/10 the pedagogical part. The average completion time was 26.991 minutes (std 5.663 minutes). In the following paragraphs, the results regarding the use of *Feedbacks* (modalities and types) and *Triggers* (types) for part 2 and 3 are presented. The original data can be found here[3]. One can note that no participant was interested in recording their voice to make an “"audio from file"” *Feedback*.

6.2.1 Technical Assessment

Tables 3 and 4 show the number of *Feedback* and *Triggers* for each pedagogical objective. The visual modality is the most used (12 times), followed by the “audio” and “haptic” modalities (both 9 times). “visual color” is more considered than “visual text” (8/12). For objectives 1 and 3, the haptic modality has the best score (with the visual ones for objective 1), while visual feedbacks are the most considered for objective 2.

Table 3: Number of *Feedbacks* by modality and type for the 2nd part.

		Pedagogical objectives		
Modalities	Properties	1	2	3
Visual	color	2	4	2
	text	2	1	1
Haptic	vibration	4	1	4
Audio	from a text	2	4	3

Regarding *Triggers*, the “metric” type appears 15 times, followed closely by the “spatial” one (13 times). In addition the “metric trigger” dominates objectives 1 and 3, while the “audio” one is the most considered for objective 2. The “time trigger” was not used at all.

Table 4: Number of *Triggers* by type for the 2nd part.

		pedagogical objectives		
		1	2	3
Types	Contact	1	0	1
	Metric	7	1	7
	Spatial	2	9	2
	Time	0	0	0

Relationships between *Feedbacks* and *Triggers* can be seen in table 5. The couples “visual/metric” and “audio/spatial” appear more frequently (7 and 6 times) over all objectives, followed by the “haptic/metric” couple (5 times).

Table 5: Relationships between *Feedbacks* and *Triggers* for the 2nd part ("V" for visual, "H" for haptic, "A" for audio feedback, "C" for contact, "M" for metric, "S" for spatial trigger).

		Pedagogical objectives								
		1			2			3		
		V	H	A	V	H	A	V	H	A
Types	C	1							1	
	M	3	2	2	1			3	3	1
	S		2		4	1	4			2

The most frequent couple for objective 1 is “visual/metric” (3 times), for objective 2 “visual/spatial” and “audio/spatial” equally (4 times), and for objective 3 “visual/metric” and “audio/metric” equally (3 times).

6.2.2 Pedagogical Assessment

In this non-mandatory part, 2 participants proposed 3 pedagogical objectives, 4 participants suggested 2 objectives and 2 participants formalized only 1 objective. In total, 8/10 participants explored 16 pedagogical objectives.

Table 6: Number of *Feedbacks* by modality for the 3rd part.

Modalities	Properties	Pedagogical objectives
Visual	color	2
	text	0
Haptic	vibration	6
Audio	from a text	8

Table 7: Number of *Triggers* by type for the 3rd part.

		Pedagogical objectives
Types	Contact	9
	Metric	1
	Spatial	5
	Time	1

Tables 6 and 7 show the number of *Feedback* modalities and *Trigger* types. The audio modality is the most used (8 times) followed by the haptic one (6 times). It seems that teachers abandoned the visual modality to focus on audio and haptic ones. “Contact trigger” was chosen 9 times and the “spatial” one 5 times. This part suggests an investigation of “contact triggers” at the cost of “"metric"” ones.

The relationships between *Feedbacks* and *Triggers* can be found in table 8. The couple “audio/contact” is the most used (4 times) followed by “haptic/contact” and “audio/spatial” (both 3 times).

Table 8: Relationships between feedbacks and triggers for the 3rd part ("V" for visual, "H" for haptic, "A" for audio feedback, "C" for contact, "M" for metric, "S" for spatial trigger).

		Pedagogical objectives		
		V	H	A
Types	Contact	2	3	4
	Metric			1
	Spatial		3	2
	Time			1

In brief, different strategies can be seen in the *Feedback* and *Trigger* choices, even for the same learning objective. In terms of frequency, the visual modality prevailed in part 2, while the haptic and audio were more investigated in part 3. In terms of trigger types, the “spatial trigger” is regularly chosen in the whole experiment.

The “metric trigger” leaded part 2 while the “contact trigger” was more studied in part 3. The participants curiosity in exploring more *Feedbacks* (except the “audio from a file”) and *Triggers* (except the “time trigger”) can be noted. In terms of couples, no tendency can be observed. Indeed, part 2 highlighted “visual/metric” (the most used) but also “visual/spatial” and “audio/spatial” couples. In part 3 “audio/contact” (the most used) but also “haptic/contact” and “haptic/spatial” can be observed.

6.2.3 Utility Questionnaire

This questionnaire is divided into two parts to estimate: (i) the usefulness of the descriptive model in the context of the dilution activity and (ii), its usefulness if those concepts will be used in any other VLE.

(i) and (ii) are made of two parts, one for *Feedbacks* with 5 questions (q1 visual text, q2 visual color, q3 audio from a file, q4 audio from a text, q5 haptic vibration) and the other one for *Triggers* with 4 questions (q1 time, q2 contact, q3 spatial, q4 metric), based on a 5-Likert scale (1 very useless, to 5 very useful). Figures 7 (a) and (b) respectively present the results regarding the usefulness of *Feedbacks* and *Triggers* for the dilution activity.

“Visual color” was considered the most useful by the participants (9/10) followed by “haptic vibration” (8/10), “audio from a text” and “visual text” (7/10). These results are consistent with the analysis of *Feedbacks* in part 2 (cf. section 6.2.1), while part 3 (cf. section 6.2.2) helped teachers in defining their preferences between audio and haptic *Feedbacks*.

Regarding triggers, 9/10 participants thought “spatial trigger”, closely followed by “contact” and “"metric"” ones, are useful. If part 2 of the experiment mainly studies the “metric” and “spatial triggers”, part 3 makes an investigation of “contact triggers” while maintaining a notable interest in “spatial” ones. “spatial triggers” seem to be unquestionable as well as “contact” and “metric” ones.

Figure 8 (a) and (b) respectively present the results regarding the estimated usefulness of *Feedbacks* and *Triggers* if they will be used in other VLEs.

Same tendencies can be observed regarding the success of “visual color” (8 times), close to “haptic vibration” (7 times) followed by “visual text” (7 times). The *Triggers* considered most useful were the “contact” and “metric” ones (both 9/10). “Spatial triggers” have notable results too (8/10). This confirms the potential interest of those three feedback modalities and triggers, no matter what the VLE is.

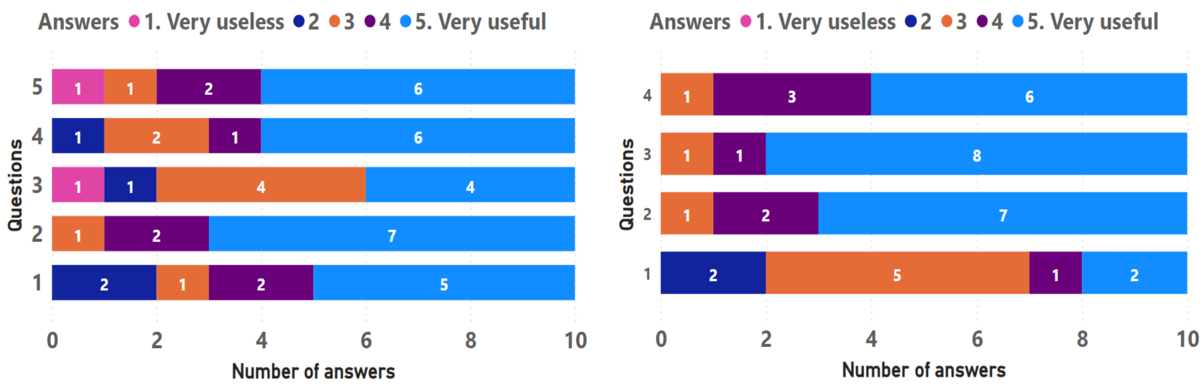
6.2.4 Reaching Pedagogical Objectives

Two questions were asked to get the agreement of participants regarding the fact that the created *Feedbacks* (question 1) and *Triggers* (question 2) allow them to achieve the pedagogical objectives. These questions were based on a Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Except for 1 participant, all of them agree with this two assertion (figure 9). The descriptive model of pedagogical feedbacks seems to respond to their pedagogic needs in this specific experimental context.

6.2.5 Usability Questionnaire

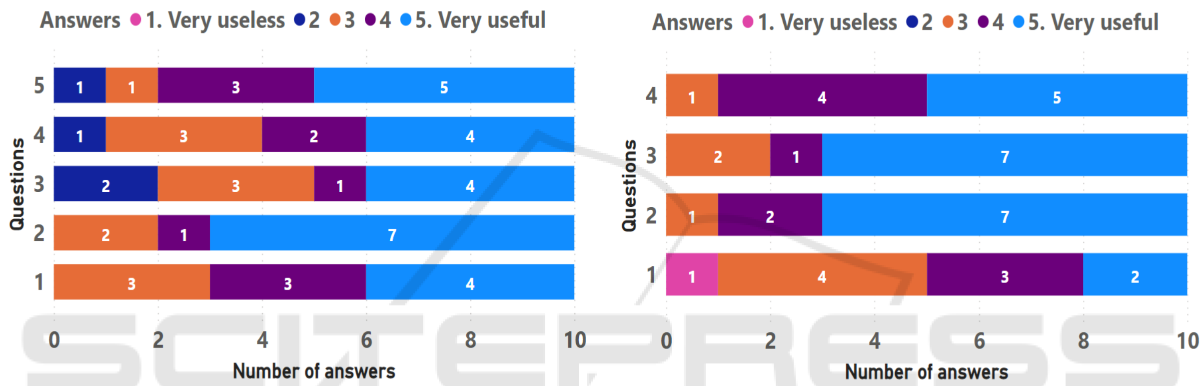
The usability of GEFEED and its HCI were assessed with the SUS questionnaire made of 10 assertions (Brooke, 2013): (1) the willingness to use the system frequently (2), its complexity (3), its ease of use (4), the need for technical support to use it (5), the well integration of its functionalities (6), its number of inconsistencies (7), its fast learning curve (8), the cumbersome aspect in its use (9), the user confidence in using it and (10), the number of things to learn before being able to use the system. The participants must state their agreement with a Likert scale of 1 to 5 (from "strongly disagree" to "strongly agree").

From the answers of those questions an average score is computed to estimate the acceptability the system (Brooke, 2013). A system is considered "acceptable" from 50.9/100 points. The average score obtained for GEFEED is 60.75 points (std 15.57) with 10 teachers, that makes it belong to the strong probable acceptance class (High marginal, figure 10). Therefore, even if its functionalities, HCI and ergonomics can be improved, the current usability level of GEFEED make its descriptive model assessable.



(a) (b)

Figure 7: Responses to the usefulness questionnaire on *Feedbacks* (a) and *Triggers* (b) in the context of the dilution VLE.



(a) (b)

Figure 8: Responses to the usefulness questionnaire on *Feedbacks* (a) and *Triggers* (b) in other contexts.

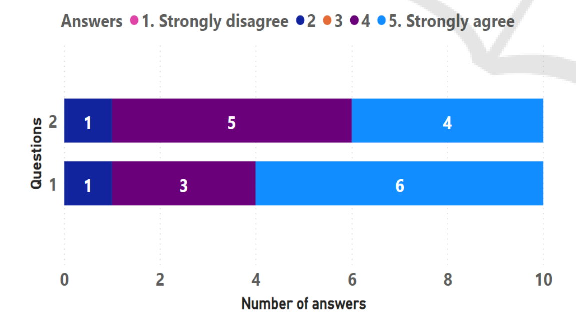


Figure 9: Responses to the questionnaire on the achievement of pedagogical objectives.

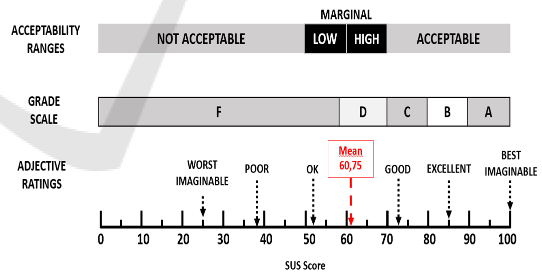


Figure 10: Average SUS score (in red) of GEFEED.

7 DISCUSSION OF FINDINGS

All the participants completed the whole protocol (except 2 for part 3) with an acceptable time. The presented results and tendencies must be considered in the specific context of one VLE. Other experiments must be done implying more teachers and other VLEs to deeply investigate these results.

A first interesting result, is the capacity of GEFEED to satisfy the pedagogical objectives (predefined or not) for 9/10 teachers.

In terms of *Feedbacks*, the visual modality and especially the color type is, from a quick look, the most interesting feedback that confirms the results of (Sigrist et al., 2012). However the audio and haptic ones were also considered at a significant and sometimes close level. For testing purposes, or by confidence, the observed practices were confirmed by the teachers' opinion on the effective usefulness of three modalities (visual, audio, haptic) and their specifically types ("visual color", "visual text", "audio from a text", "haptic vibration"). The absence of interest for "audio from a file" feedback cannot currently be explained, showing the need of a study group or interviews of participants after the experimentation.

For the triggering rules, a clear exploring strategy occurs by studying the "metric" and "spatial" triggers in the second part of the experiment, and then, the "contact" and "spatial" ones in the last part. The interest of "spatial trigger" appears as unquestionable, while the metric and contact triggers were also useful for the majority of the participants. However, the interest in the time trigger was limited. An explanation can be the absence of a time constraint requirement in the dilution task and the pedagogical objectives.

In terms of operationalisable and pedagogical strategies, even if the couples "metric/visual", "audio/visual" and "audio/contact" can be noted, no tendency clearly appears regarding the best *Trigger/Feedback* couple to use for one pedagogical objective. Nevertheless, this shows the diversity of the strategies of the teachers and/or their curiosity in exploring the functionalities of GEFEED.

During the third part of the experiment, the pedagogical objectives of the teachers were formalized such as "homogenize the solution before getting the sample", "place the pipette in the container after releasing the sample in the test tube", "the pipette must be attached to the rubber bulb", etc. However those statements were more closed to operational instructions than a clear pedagogical objective or intention (*e.g.* learning a specific action or motion, a sequence of actions, avoiding a dangerous gesture, discover the effect of an action, use a specific knowledge, etc.). Consequently, one cannot study the relation between the pedagogical objectives and the design elements of the pedagogical feedbacks. The descriptive model must be extended by incorporating some fields to record a formalization of the pedagogical objectives or intentions, if possible categorize them, and distinguish them from operational instructions given to the learners, that must also be saved in the model.

8 CONCLUSION AND FUTURE WORK

In this paper, a three-dimensional descriptive model of pedagogical feedbacks for learning gesture-based tasks was proposed as well as its operationalization through the GEFEED system. GEFEED allows any teacher to create multimodal feedbacks in any VLE developed with Unity engine.

The pedagogical feedbacks in VLE being a crucial elements for assisting the motion learning, this study aims to propose a full-processing chain made of useful design elements to reach their various pedagogical objectives.

This processing chain relies on the definition of: a virtual representation of the pedagogical information to convey, the triggering rules of the feedback, and the 3D objects implied in the virtual representation or the triggering rules.

An experimental study allow analyzing teachers' practices regarding the feedback creation for a dilution simulation. Among the proposed feedbacks and triggers provided by GEFEED, the first results highlight: (a) for feedbacks, the visual color, visual text, audio form a text and haptic vibration for feedbacks and (b), the contact between two objects, spatial static configuration and thresholds of motion metrics for triggers, as the most used and useful design elements to reach the pedagogical objectives.

However, those tendencies must be deeper studied with more participants and others VLEs. This can also be crucial for the identification of the best association "feedback/trigger" as no tendencies can be currently observed given the current task.

Finally, the current three-dimensional model must be extended to better distinguish, formalize and save the pedagogical objectives, intentions and instructions. With those pieces of information, an analysis method will be proposed to identify efficient design elements for creating a feedback able to reach a targeted pedagogical objective, given a specific learning situation. This is essential before evaluating the efficiency of a VLE enhanced by GEFEED pedagogical feedbacks, on gesture learning.

REFERENCES

Adolf, J., Kán, P., Outram, B., Kaufmann, H., Doležal, J., and Lhotská, L. (2019). Juggling in vr: Advantages of immersive virtual reality in juggling learning. In *25th ACM Symposium on Virtual Reality Software and Technology*, VRST '19, New York, NY, USA. Association for Computing Machinery.

Brooke, J. (2013). Sus: a retrospective. *Journal of usability studies*, 8(2):29–40.

Cannavò, A., Pratticò, F. G., Ministeri, G., and Lamberti, F. (2018). A movement analysis system based on immersive virtual reality and wearable technology for sport training. In *Proceedings of the 4th International Conference on Virtual Reality*, ICVR 2018, page 26–31, New York, NY, USA. Association for Computing Machinery.

Chen, L., Wu, L., Li, X., and Xu, J. (2019). An immersive vr interactive learning system for tenon structure training. In *2019 2nd International Conference on Data Intelligence and Security (ICDIS)*, pages 115–119.

Djadja, D. J. D., Hamon, L., and George, S. (2020). Design of a motion-based evaluation process in any unity 3d simulation for human learning. In *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, VISIGRAPP 2020, Volume 1: GRAPP, Valletta, Malta, February 27-29, 2020*, pages 137–148. SCITEPRESS.

Larboulette, C. and Gibet, S. (2015). A review of computable expressive descriptors of human motion. In *2nd International Workshop on Movement and Computing*, pages 21–28. IEEE.

Le Naour, T., Hamon, L., and Bresciani, J.-P. (2019). Superimposing 3d virtual self + expert modeling for motor learning: Application to the throw in american football. *Frontiers in ICT*, 6:16.

Lee, G. I. and Lee, M. R. (2018). Can a virtual reality surgical simulation training provide a self-driven and mentor-free skills learning? investigation of the practical influence of the performance metrics from the virtual reality robotic surgery simulator on the skill learning and associated cognitive workloads. *Surgical Endoscopy*, 32.

Lin, H. H., Han, P. H., Lu, K. Y., Sun, C. H., Lee, P. Y., Jan, Y. F., Lee, A. M. S., Sun, W. Z., and Hung, Y. P. (2018). Stillness moves: Exploring body weight-transfer learning in physical training for tai-chi exercise. In *Proceedings of the 1st International Workshop on Multimedia Content Analysis in Sports*, MMSports'18, page 21–29. Association for Computing Machinery.

Liu, J., Zheng, Y., Wang, K., Bian, Y., Gai, W., and Gao, D. (2020). A real-time interactive tai chi learning system based on vr and motion capture technology. *Procedia Computer Science*, 174:712–719. 2019 International Conference on Identification, Information and Knowledge in the Internet of Things.

Lo, T. T., Xiao, Z., and Yu, H. (2019). Designing 'action trigger' for architecture modelling design within immersive virtual reality. *CUMINCAD*.

Luo, Z., Yang, W., Ding, Z. Q., Liu, L., Chen, I.-M., Yeo, S. H., Ling, K. V., and Duh, H. B.-L. (2011). "left arm up!" interactive yoga training in virtual environment. In *2011 IEEE Virtual Reality Conference*, pages 261–262.

Mahdi, O., Oubahssi, L., Piau-Toffolon, C., and Iksal, S. (2019). Assistance to scenarisation of vr-oriented pedagogical activities: Models and tools. In *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*, pages 344–346.

Mizuyama, H. (2010). A semiotic characterization of the process of teaching and learning a skilled motion taking wok handling as an example. In *11th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems*, volume 43, pages 448–453.

Nishino, H., Murayama, K., Shuto, K., Tsuneo, K., and Utsumiya, K. (2011). A calligraphy training system based on skill acquisition through haptization. *Journal of Ambient Intelligence and Humanized Computing*, 2:271–284.

Senecal, S., Nijdam, N. A., and Thalmann, N. M. (2002). Motion Analysis and Classification of Salsa Dance Using Music-related Motion Features. In *Proceedings of the 11th Annual International Conference on Motion, Interaction, and Games*, pages 1–11. IEEE.

Sigrist, R., Rauter, G., Riener, R., and Wolf, P. (2012). Augmented visual, auditory, haptic, and multimodal feedback in motor learning : a review. *Psychonomic bulletin & review*, 20(1):21–53.

Wang, Haoyu, W., Sheng, J., and Jianhuang, W. (2018). A Virtual Reality Based Simulator for Training Surgical Skills in Procedure of Catheter Ablation. In *International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, pages 247–248. IEEE.

Wei, W., Lu, Y., Printz, C. D., and Dey, S. (2015). Motion data alignment and real-time guidance in cloud-based virtual training system. In *Proceedings of the Conference on Wireless Health*, WH '15. Association for Computing Machinery.