

The role of mismatches in the sensory feedback provided to indicate selection within a virtual environment

Raquel Viciano-Abad · Arcadio Reyes-Lecuona ·
Matthieu Poyade · José Escolano

Published online: 11 June 2010
© Springer Science+Business Media, LLC 2010

Abstract It is generally understood that virtual reality simulations have a high computational cost. Hence, they rarely can reduce completely all the incoherence within the cross-modal sensory outputs provided. The main research approaches to date have consisted in technically reducing possible mismatches, however minimal research has been conducted so as to analyse their influence on human capabilities. Thus, the objective of this study is to provide further insights to the designers of virtual reality about the negative influence of simulation lags and interesting design implications. To clearly show this, we have investigated the importance of coherent sensory feedback by incorporating time delays and spatial misalignments in the feedback provided by the simulation as a response to participant's actions to mimic computationally expensive environments. We have also evaluated these misalignments considering two typical interaction setups. In particular, the sensory mismatches influence has been assessed in human factors, such as the sense of presence, task performance and delay perception. Our experimental results indicate that the closer the interaction conditions are to real configurations the higher the sensory requirements are regarding accuracy. The implications of this study offer the designer guidelines to prioritise the reduction of those mismatches in the sensory cues provided depending on the simulations goals.

Keywords Co-location · Delay · Feedback · Mismatches · Performance · Presence · Selection

R. Viciano-Abad (✉) · J. Escolano
Department of Telecommunication Engineering, University of Jaén, Jaén, Spain
e-mail: rviciano@ujaen.es

A. Reyes-Lecuona · M. Poyade
Department of Electronic Technology, University of Málaga, Málaga, Spain

1 Introduction

As virtual reality (VR) devices have become increasingly less expensive, issues related with the use of virtual environments (VE) within different applications and the inclusion of multimodality, user's stimulation through different sensory channels, have become more common in various research areas such as human–computer interaction and experimental psychology [38]. The inclusion of these devices within a high-consumer computational application, such as one based on a realistic 3D environment, yields an increment of computational cost due to the addition of concurrently running operations. This, in turn, means that modern computers still show delayed responses to user input on many occasions. Furthermore, VE are widely used in collaborative applications [4] where geographically distributed users collaborate in real time through network connections. Typical applications can be found relating to activities such as training, planning (medicine, industry, emergency assistance, etc.), telemonitoring and gaming [35]. In these VR systems lag is sometimes inevitable and can be attributed to inappropriate election of sampling/updates rates of input/output devices, low software performance and limitations of the system hardware, due to the existing trade-off between function, time and money. In addition, for networked VE, software overheads (i.e. network communications) may also become a lag source.

Real time systems are considered as systems that fail if its performance criteria are not met; criteria that depend on the application context. The difference between a 'soft real-time system' and a 'hard real-time system' is that soft real-time systems only are expected to meet their objectives on an average basis whilst hard real-time systems have to meet their objectives 100% of the time. As a result these soft real-time systems are commonly more suited to novel applications such as VR ones [10]. Nevertheless, in VR applications the performance criteria rather than only in objective aspects are also based on subjective ones, such as participants' perceptions. Therefore, the definitions of the criteria against which real-time performance successes are measured to date have not yet been clearly agreed upon.

Based on the literature reviewed, we can affirm that the evaluation of thresholds where possible sensory mismatches may degrade the interaction, via diminishing human factors such as presence, performance or perceived system responsiveness is still open to discussion. Attending to multimodality, several studies (i.e. see meta-analysis performed in [8]) have reported the benefits of providing richer sensory outputs in a VE. But despite its benefits, multimodality can become an additional source of mismatches for the design of a VE, due to the synchronization requirements of all the system sensory outputs. For this reason, the characterization of thresholds for possible asynchronies (attending to different factors) is an important issue to address prior to the inclusion of multimodality. Nevertheless, the literature available to date is surprisingly limited in the evaluation of thresholds beyond the perceptible ones. First, as in other fields, this lack of evaluation is due to the assumption that any kind of asynchrony is a negative one and hence, major efforts are paid to minimize them. Second, as described, multimodal information is included assuming that there are several brain mechanisms to tackle incoherent information.

In considering whether or not it would be beneficial to develop special techniques to reduce completely mismatches within a VE, it would be interesting to identify the asynchronies that make human factors more vulnerable. Thus, it is not only necessary to evaluate thresholds of permissible mismatches but also to identify those mismatches which have a more negative influence.

In certain studies, the effect of mismatches on different human factors have been evaluated, but most of them were concerned with effects on performance measures such as accuracy errors and overall times [20, 36]; and attended only to delays in the visual rendering derived from different lag sources [27, 46, 47]. These are typical factors evaluated for the interaction in a 2D graphical user interface (GUI), but design criterions about lag or latency within a VE may account for the degradation of presence [1, 24]. A phenomenon that includes aspects such as realism, involvement, content, perceived quality, sensory fidelity and so on. Furthermore, mismatches in auditory and haptic cues are also relevant in typical 3D applications that usually provide multimodal information.

Lag is also typical in daily activities where people use several devices with technological limitations. Nevertheless, people have the ability of easily adapting to them because the brain integrates information to yield immediate corrections or adaptations in terms of a longer time scale [21]. Indeed, on a typical VR interaction with a desktop setup, hand-eye coordination is performed when we point to objects on a computer monitor by controlling a cursor with a mouse. Thus, to establish the correct relationship between visual and proprioceptive cues, we are forced to adapt to the new visuo-motor arrangement. For instance, when haptic feedback is provided within a VE, it is common to find the existence of a spatial misalignment between the visual and proprioceptive workspaces. Participants do not touch objects in the same place where they visually see them. Considering this interaction set-up, it is still open to research the analysis of the extent to which the assumption of a spatial misalignment may alter the adaptation to typical asynchronies encountered in virtual simulations during a short time period.

In sum, in a virtual simulation with multimodal outputs, where limiting all the possible kinds of asynchronies is difficult, and with different interaction set-ups to those encountered in that of normal interactions, it is necessary to assess the more critical inaccuracies in the feedback provided. One of the main goals of the work presented here consists in analysing the extent to which temporal and spatial disparity among sensory cues provided as selection feedback in a VE affect two factors, the sense of presence and task performance. This goal has been addressed with an experiment based on including temporal and spatial asynchronies in the feedback provided within a VE to indicate an object selection to participants. Furthermore, we have performed the experiment with two interaction set-ups so as to analyse possible differences in the effect of sensory mismatches due to the alignment between visual and haptic workspaces. This study considers selection actions because they are normally the first made to perform other tasks, for instance manipulation and navigation tasks. In particular, the selection technique implemented is based on the classical virtual hand metaphor that consists in mapping the participants' hand movements in a virtual object or avatar using a linear transformation. Techniques based on this metaphor are more natural and intuitive [12]; hence, effect of discrepancies in sensory feedback provided must be more evident.

The rest of this paper is organized as follows. First, a summary of the main conclusions achieved regarding the influence and adaptation to sensory mismatches. In Section 3, the testbed implemented to evaluate the influence of mismatches in the sensory outputs provided within a VE is described. The experimental procedure followed and main results obtained are respectively presented in Sections 4 and 5. In Section 6, the main contributions of this study are discussed within the frame of previous work. Finally, the main conclusions are summarized in Section 7.

2 Background and study rationale

Research about multimodal interaction with mismatches has been mainly addressed within the field of neuropsychology and attending to two types of mismatches. One is a first order asynchrony, included in a long term period, between spatial and proprioceptive cues, which is analysed including spatial misalignments between these workspaces using prisms. The other is a second order temporal asynchrony between sensory cues, which has been mainly focused on auditory and visual cues. In the following sections, we review the main studies about the two categories of mismatches in the sensory cues provided and main results obtained.

2.1 Time asynchronies in sensory information

Wilhelm Wundt, considered one of the modern psychology fathers [6], assessed the thresholds for intramodal simultaneity perception, that is to say, between two stimuli of the same nature. These values still stand today, but thresholds in which intermodal (between stimuli of different nature) simultaneity perception takes place are also being evaluated. In Table 1, we have listed the main contributions made about simultaneity thresholds and mismatches effects by different experimental studies.

In certain circumstances, the time delay between different stimuli is not perceptible due to different resolutions of senses and the brain's ability to coherently integrate this information. While for other perceptible asynchronies, neuropsychology studies have indicated that the cognitive system coherently integrates non simultaneous sensory information. Hence, via the processes involved in this adaptation, a person may continue performing a task, minimizing the chance of annoyance. Different studies [29, 43] have corroborated this theory and have evaluated different processes, such as the voice intelligibility when there are mismatches among voice associated visual and auditory cues (See Table 1). Likewise, thresholds of cross-modal perceptual simultaneity attending to visual and non-speech auditory stimuli have been assessed (review included in [25]). Two main findings have been outlined by these studies. First, that there are high individual differences and second, that there is a higher tendency for a simultaneity perception to occur when the auditory stimuli are provided first. Results have also indicated that the tolerance thresholds for delay among visual and auditory cues are higher in the case of them being associated with speech.

The detection range of mismatches between visual and haptic stimuli has also been evaluated, but mainly attending to the tactile component of haptic perception. Hence, aspects such as the hand movement, information that can be decisive in the perception of delay between haptic and visual stimuli, have not yet been considered. These studies [17] have found that visual stimuli must precede haptic ones so as to estimate simultaneity. Nevertheless, more recent studies have obtained that aspects such as attention may shift the perception of simultaneity. In line with this, Spence et al. [39] have detected that thresholds were biased by attention, being lower when attention was more focused on visual stimuli.

All of the above-mentioned studies have analysed basic aspects of perception processes, and therefore, they have considered basic visual, auditory and tactile stimuli. On the contrary, during the interaction in a VE, further analysis is necessary about the influence of mismatches within sensory information considering more complex stimuli, such as those associated with the lighting of an object, the tactile information perceived as a consequence of an object selection with our hands, etc. Also, in a complex environment, real or simulated, people do not completely pay attention to luminic stimuli or tones, for instance

Table 1 Studies about simultaneity thresholds and effects

Authors	Goals	Main results
Extracted from Boring (1929) [6]	Intermodal simultaneity thresholds.	Sound: 2 ms, Touch: 27 ms, Sight: 43 ms.
Munhall et al. (1996) [29], Vatakis and Spence (2006) [43]	Voice intelligibility.	Not affected by time delay of 80 ms with auditory cues first and 200 ms with visual cues first.
Review made in Levitin et al. (2000) [25]	Cross-modal perceptual simultaneity between visual and auditory (non speech) cues.	High individual difference and thresholds of 75 ms (auditory first) and 90 ms (visual first).
Loomis (1992) [26]	Simultaneity thresholds among visual, auditory and haptic stimuli associated with collision awareness.	Visual and auditory cues: thresholds of 41 ms (auditory first) and 45 ms (visual first). Auditory and visual cues: thresholds of 223 ms (auditory first) and 42 ms (haptic first).
Spence et al. (2001) [39]	Influence of attention in stimuli perception.	Attentional biases toward the visual modality and toward the right side of space. Simultaneity thresholds with visual cues first of 155 ms (attention in tactile stimuli) and 22 ms (attention in visual stimuli).
Rothrock et al. (2006) [34], Simpson et al. (2007) [36], Szameitat et al. (2009) [41]	Effects of brief delays in system responsiveness (2D GUI) in performance and emotions.	Delays of 1.5 s (within a range of [500–2,800] ms) in visual awareness of target movement influenced negatively reactions times, error rates, emotions (dislike and annoy) and workload.
MacKenzie and Ware (1993) [27]	Effects of lag (visual updating) in performance during a 1D targeting task.	Measured effect with lag of 75 ms and performance degradation with 225 ms.
Jay and Hubbold [19, 20]	Influence in performance and perceived task difficulty of delays in visual and haptic feedback during a 2D targeting task.	Less negative influence of delay in haptic feedback. Movements time were significantly slowed by a lag of 69 ms in visual feedback (cursor movement) and did not increase in response to haptic latency until it approached 200 ms.
Lee and Choi [24]	Average discrimination of delays in visual and haptic feedback in a 3D environment.	During an activity that consisted in moving a block, average discrimination thresholds of delay were 150 ms for haptic feedback and 100 ms for visual feedback.

they perceive information of different nature and from different locations. Levitin et al. [25] have also evaluated the intermodal simultaneity thresholds in a task that consisted in striking a pad with a mallet. All the stimuli analysed (visual, haptic and auditory) were related with collision awareness. As pointed out in Table 1, thresholds were higher between auditory and haptic stimuli, and when auditory cues were emitted prior to haptic ones.

Beyond the thresholds of simultaneity, within the field of 2D human–computer interaction, several studies (see [39] for a complete review) have evaluated the extent to which the delay in typical stimuli included in a PC interface may affect negatively on different emotional factors (i.e. stress, tiredness, annoyance, etc.) and performance. Attending to interaction on a 2D graphical user interface, delays of approximately 1.5 s

in computers responsiveness have negative effects in performance and mental workload [34, 36, 41]. Mackenzie and Ware [27] stated that within the field of VR, where the requirement of a tight link between the user's view of the environment and their actions was essential, the evaluation of lag impact on human performance was rare. This study results in the characterization of lag influence on performance for a 1D targeting acquisition task using a mouse, reflecting that lags of 225 ms substantially degraded performance. For 2D targeting tasks [19, 20], the influence on performance of delays in visual and haptic feedback has been also evaluated. These studies indicated that lag in visual cues was more negative than that of haptic latency, considering movement times, mistakes and perceived task difficulty. A similar study was performed in a 3D environment [24], but only considering discrimination thresholds. While moving a block from one to another place, participants' average discrimination thresholds of delay in haptic feedback were also higher than visual delay thresholds.

In spite of these studies, within a VE where human computer interaction and its consequences are closely related with daily interaction, further research is needed to evaluate the influence of sensory mismatches. As we have previously mentioned, besides the ranges of perceived intramodal and crossmodal simultaneity, there are also recalibration processes or brain processes activated to tackle small temporal mismatches [21] that allow for higher thresholds. A person may overcome asynchronies out of the perceived simultaneity range via two mechanisms: (1) a short term mechanism with immediate corrections [28, 45] and (2) adaptations in a longer-time scale [15, 21, 30, 45]. These processes may contribute within a VE to reduce possible negative effects of sensory mismatches in certain human factors.

2.2 Spatial mismatches in sensory stimulation

In spite of inherent differences in the senses related to the spatial resolution, it is important to characterize the extent to which spatial mismatches in auditory, visual and haptic cues may influence certain factors during the interaction within a VE. One of the main advantages of VE is the flexibility of graphically changing the space in which a person is interacting. In addition, workspaces attending to visual, auditory and haptic information are not normally coincident, but through different techniques they could be perceived as coherent [50]. As an example, there are simulations with scale differences between visual and auditory workspaces, or where workspaces are not coincident. This last case is usually found within a typical interaction set-up that includes a force feedback device, where visual and proprioceptive workspaces are not coincident. The proprioceptive workspace is limited by the work area of the haptic device whilst the visual changes are graphically depicted in a vertical display.

In perception research, certain studies (see review in [33]) have detected that visuo-motor coordination changes if proprioceptive and visual cues are misaligned. In these studies, the experimental procedure consists in asking participants to point their hand at a specific object while they see the environment through a prism. This prism is used to shift the visual information from the proprioceptive one obtained from muscles and joints. Rossetti et al. [33] suggested that there are two mechanisms to allow spatial recalibration in presence of contradictory stimuli. A long term mechanism that allows a remapping between the perceptual spaces not aligned and, a second one activated immediately. The latter is able to quickly optimize accuracy in conditions of temporary decoupling. Nevertheless, these possible adaptations happen within certain ranges and with lateral deviations, but they may require months or may not even take place at all if the visual workspace is completely inverted [16].

In the field of VR, several studies have reported an adaptation process to spatial misalignment between proprioceptive and visual workspaces [9, 13, 23], but mainly in cases without haptic feedback. Moreover, this adaptation ability is used to create two powerful illusions when participants interact with their hands within the VE. The first of these illusions, referred to as the pseudohaptic technique, consists in creating certain haptic perceptions using visual constraints. The second illusion is referred to as the rubber-hand illusion [18] and this is based on the ability of perceiving as own a virtual hand, discarding certain proprioceptive information from one's real hand.

Other researchers have analysed the effects of spatial co-location when participants used a haptic device, such as a PHANToM (from Sensable Inc.) These studies have considered the influence of spatial misalignment in the provided feedback, attending also to the existence or not of haptic feedback. The effect of a correct perspective in a VE based on a typical fishtank configuration has been evaluated in [3] for a selection task. Spatial coincidence yielded a similar reduction in selection times than that of when haptic feedback was included. The influence of spatial co-location on performance was also analysed for three kinds of tasks with different complexity levels [40]: accuracy selection, spatial manipulation (moving objects within a maze) and spatial response (object manipulation with simulated gravity so as to avoid falls). This study has demonstrated that coincidence leads into a positive influence for spatial selection and spatial response. In the case of haptic feedback, this improvement was also more evident for a spatial response. In contrast, Congedo et al. [11] did not report the lack of coherence between these workspaces as a negative effect, in both conditions, with or without haptic feedback. This study has analysed the performance of a manipulation task that consisted in positioning an object so as to make it coincide with a model.

Several studies [32, 42, 49] have also found that in the case of spatial discrepancies between two sensory modalities, one of them, normally the less dominant one, gets shifted to the other one, a phenomena known as intersensorial predisposition. In the case of spatial mismatches between visual and auditory cues, the visual information is the more dominant one, a phenomenon known as 'ventriloquist effect' [5].

2.3 Study rationale

A great deal of research has been conducted on the role of mismatches among stimuli attending to isolate cues and within the field of neuropsychology. Considering stimuli typically provided during human–computer interaction, typical delays due to background processes or network communications have been analysed regarding mainly task performance but also, considering emotional states. Nevertheless, within a VE where the mental model involved in the interaction is more closely related to that of a real experience, these effects have not explicitly been evaluated. Indeed, in this last situation it is an important issue to evaluate the extent to which possible mismatches in the feedback provided may disrupt the users' feeling of being within the environment, referred in the literature as the sense of presence [26]. In other words, the extent to which the devices used to interact are transparent for the participants' experience

The main goal of this study is to assess the influence of the typical mismatches encountered during the interaction within a VE in typical selections tasks. Thus, our main research question is:

1. Which are the sensory mismatches that influence more negatively presence and task performance within a selection task?

Due to the nature of the task tested, selection, where haptic stimuli are the more natural information that one usually perceives when selecting an object, we hypothesize:

Hypothesis 1. While performing a selection task, asynchronies in haptic feedback are the most negative in presence and performance.

Typical interaction within a VE displayed on a desktop is usually provided in a condition with misalignment between visual and proprioceptive workspaces. Nevertheless, in everyday interaction, selection is made in conditions of coherence. Thus, the second question raised in this work is:

Question 2. Does the influence of sensory mismatches change depending on the degree of coincidence between visual and proprioceptive workspaces?

Our hypothesis in this case is that in presence of a first level incoherence, such as that included between workspaces, the influence of brief mismatches in the sensory feedback provided is lower than in a condition of coherence. Thus, the second hypothesis is stated as follows:

Hypothesis 2. Mismatches in the selection feedback are more negative in a condition of coincidence between workspaces.

3 Testbed based on the Simon game

3.1 Description

The system used as a testbed reproduces a virtual version of the game “Simon” (see Fig. 1). This testbed allows interaction via basic selection actions and it is used to assess the influence of mismatches in the feedback provided to indicate selections. The game consists of four differently coloured buttons and, by lighting them and emitting a different sound, the Simon game shows a random sequence. Participants must then try to reproduce this sequence correctly by pushing the respective buttons. Button selections were made through the stylus of a desktop PHANToM (from Sensable), which is modelled in the VE as a wooden pencil.

The testbed provided multimodal interaction. First, during the sequence emission, participants could touch the device, see the device and the buttons lighting, and hear the different sounds emitted. Second, the feedback provided to indicate a button selection always consisted of: (1) the button lighting while it was pressed, (2) a typical beep sound (different for each button) of 200 ms, (3) the force feedback provided by the PHANToM device as a programmable function of participants’ position.

3.2 Implementation and testbed configuration

The testbed system was implemented in Microsoft Visual C++ 6.0 using the freely available OpenGL libraries and the 3DTouch SDK of Sensable Technologies. The collision detection was carried out during the graphical rendering using the library SWIFT ++ [14].

Once a button was selected, depending on the experimental condition, mismatches included in the selection feedback could be: (1) a spatial misalignment in the haptic feedback, (2) a time delay in the button lighting, and (3) a time delay in the beep emission.

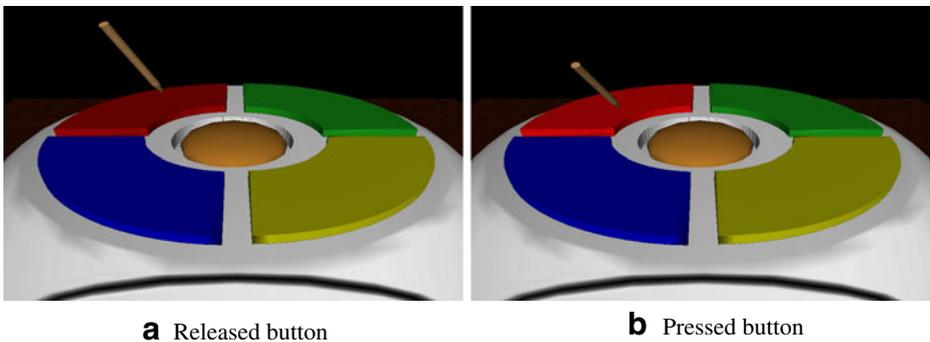


Fig. 1 Images of the Simon game used in the testbed, with all the buttons released (a) and with one button pressed (b)

The spatial misalignment was included by introducing a vertical displacement between the visual and haptic models of the button. Thus, in conditions with a misalignment in the haptic feedback, the button surface was not felt until the wooden pencil went through its visual model a programmable distance. Time delay was referred to the visual awareness of a collision between the wooden pencil and the button. Hence, in conditions with time delay, participants felt a delay in the button lighting and/or the sound emission since they visually perceived the collision.

Attending to the degree of coincidence between the visual and proprioceptive workspaces, participants played the Simon game within the platform in one of two possible configurations. In one configuration, referred to hereafter as co-located (C), participants interacted using the Reachin display developed by Reachin Technologies (Fig. 2a). This device consists of a monitor, a mirror, and a standard pair of LCD shutter glasses (Crystal Eyes from Stereo Graphics Inc.). Visual and proprioceptive workspaces were aligned by using the mirror, so that participants could see the changes derived from their selections in the same place where they did them. In the other configuration, referred to hereafter as non co-located (NC), only the PHANToM device was used. In this condition, the Simon game was shown on a vertical display (Fig. 2b) while participants moved their hand within the PHANToM workspace. Stereoscopic vision in a backward condition (inside the screen) was provided in both configurations.

4 Experimental method

The experiment concerns itself on the effect of sensory mismatches in task performance and the sense of presence, with two interaction configurations attending to the degree of coincidence between proprioceptive and visual workspaces. A measurement of perceived delay in the system responsiveness was also included so as to analyse the differences among the participants perception of mismatches and their influence in these two factors.

4.1 Participants

Thirty-two participants (23 men and 9 women) were recruited without any compensation or reward. Participants were aged between 21 and 30 ($\mu=23.1$; $\sigma=2.5$). They were students of bachelor (12) and master (15) degrees within the field of Telecommunication Engineering,

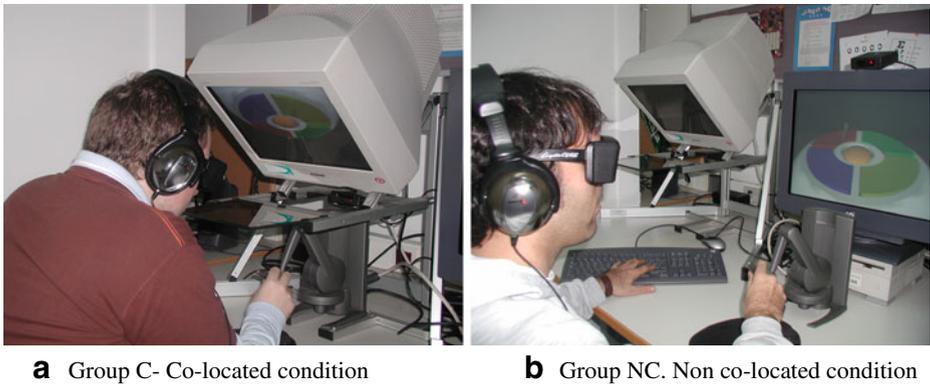


Fig. 2 Interaction set-ups used in the experiment. Participants of group C interacted in a set-up (a) with visual and proprioceptive workspaces coincident. Participants of group NC interacted in a set-up with a spatial misalignment between workspaces (b)

and researchers (5). All participants were located in the University of Málaga and were quite experienced in computer usage. However, they were all novice users of both VR applications and force feedback devices.

A week previous to the experiment date, all the participants interacted in the same testbed and in the same co-location condition during another experiment. As reported in Ref. [44], during that previous experiment the influence of different feedback modalities was evaluated, and they did not always perceive the stimuli provided in this experiment.

4.2 Experimental conditions

The sequences emitted were configured with a length of five random steps and an interval of 300 ms between the buttons switching from on to off. The interval elapsed among steps was that of 500 ms. As described in Section 3.2, the feedback provided to indicate the selection had three possible sources of mismatches. These were: 1) a misalignment of the haptic feedback was included decoupling visual and haptic models to a distance of 0.25 cm or 0.75 cm; 2) a delay in the lighting of the button that was set at 200 ms; and 3) a delay in the button sound of 400 ms.

These values were chosen from a pilot experiment made with 10 participants. During that stage, we tested ten different configurations of asynchrony, to establish perceptible values for each feedback modality. The configuration chosen were those rated with a higher number of participants. Regarding auditory and visual delays in the feedback provided to indicate selection, the majority of the participants chose one value as annoyance. Nevertheless, regarding haptic feedback the participants reported as annoyance different thresholds that held values among 0.25 cm and 0.75 cm. The participants' sensitivity, in terms of annoyance levels, to the time delay included in the auditory stimuli was lower than that of the sensitivity to the delay in lightning of the button. Hence, a higher value of delay was included in the buttons sound emission. As for the lightning of the button, a delay of 200 ms was already perceived as a system drawback. The mismatch in the haptic feedback was detected from 0.25 cm, however this was considered as permissible for half of the participants. Therefore, we also chose a second value of 0.75 cm that could make its effect even more evident.

In the studies reviewed [28, 45], the tolerance thresholds evaluated for time asynchronies were included between two stimuli, such as a flash or a tone, obtaining different results depending on the respective emission order. On the contrary, in this platform chosen delays were referred to the moment in which participants visually perceived a button selection. Nevertheless, to evaluate the influence of providing first visual (lighting) or auditory cues (beep sound), we considered conditions showed in Fig. 3. These asynchrony conditions allowed for evaluating the differences in their effect due to their simultaneity with other mismatches and their order.

4.3 Design

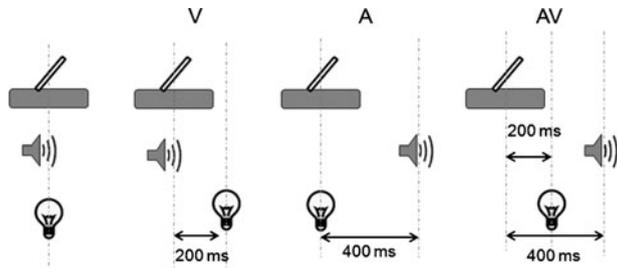
The experiment performed had a $2(\text{co-location}) \times 3(\text{haptic misalignment}) \times 2(\text{visual delay}) \times 2(\text{auditory})$ -mixed design with co-location as a between-subjects factor and the three conditions of feedback mismatches as within-subject factors. As can be seen in Fig. 4, the co-location factor had two possible values: (1) co-located condition (group C) and (2) non co-located condition (group NC). Therefore, participants were divided into two groups. The sixteen participants of group C interacted with visual and proprioceptive workspaces aligned, using the Reachin display; and the other 16 participants, in group NC, interacted in a typical desktop set-up. The three within-subject factors and values were: time delay in auditory feedback with values 0 (NA) or 400 ms (A); time delay in visual feedback with values 0 (NV) or 200 ms (V); and a spatial misalignment in haptic feedback with three values 0 (NH), 0.25 cm (H1) and 0.75 cm (H2). The combinations of these values corresponded with twelve possible conditions. Hence, participants interacted in the twelve blocks described below in a randomized order and reproducing five sequences in each one:

- **Block S.** All the sensory stimuli provided to indicate a button selection were synchronized.
- **Blocks A, V, H1 and H2.** Four blocks with mismatches in only one modality of sensory feedback.
- **Blocks AV, H1A, H2A, H1V and H2V.** Five blocks with mismatches in two modalities of sensory feedback.
- **Blocks H1AV and H2AV.** Two blocks with mismatches in the three modalities of feedback.

As can be also seen in Fig. 4, the dependent factors considered were the sense of presence, the perceived delay in the system responsiveness and task performance.

Presence was measured via three items extracted from the Slater-Usuh-Steed (SUS) questionnaire [37] (see Appendix I). Answers to this questionnaire are anchored in a seven Likert scale and were operationalized computing the number of items rated over 5. The

Fig. 3 Conditions of time delay between visual and auditory cues referred to the visual awareness of a button selection: synchronized stimuli, V (visual delay), A (auditory delay) and AV (auditory and visual delays)



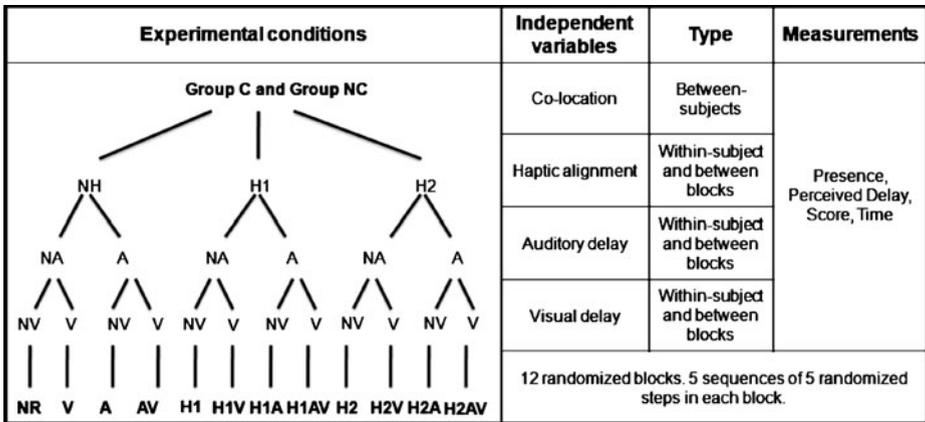


Fig. 4 Experimental design considering four factors: co-location (group C and NC), haptic misalignment (NH: without misalignment, H1: 0.25 cm and H2: 0.75 cm), auditory delay (NA: without delay, A: 400 ms) and visual delay (NV: without delay, V: 200 ms)

presence score was taken as the number of answers r out of n (the number of questions) that have a score of ‘6’ or ‘7’. This measure is referred to hereafter as the SUS factor. The perceived delay was an additional item extracted from the PQ [48] (see Appendix I). Regarding task performance, we measured the number of correct sequences reproduced (referred to as the Score) and the time elapsed between button pressings (referred to as Time). We measured through the Score, participants’ movement accuracy and their ability to reproduce sequences despite the different mismatches in the feedback. Likewise, the Time measurement allows for assessing the influence of delayed feedback in the participants’ reactions.

4.4 Procedure

Upon arrival to the research laboratory, participants fulfilled a consent form and answered personal questions about their experience with computers, games and any possible sight illness, etc. As mentioned above, these participants were familiar with the Simon game and the answer procedure. We explained to them that in this new experiment they would always feel the three kinds of feedback cues but that they would notice the existence of different mismatches within that feedback. To control the possible influence of their previous experience in the testbed, participants interacted in the same condition attending to the co-location factor. Thus, the same sixteen participants interacted in group C and the other half in group NC.

After all the information was given, they performed the experiment without any intervention. As previously explained, the task consisted in reproducing five sequences in each condition. After every block of trials, they completed a four itemed questionnaire displayed on the screen, using the keyboard to rate each item.

5 Results

We performed a repeated measures analysis of variance (ANOVA) with each dependent variable including the four factors considered: co-location and the three conditions of

mismatches in audio, visual and haptic stimuli. This analysis reported those differences between experimental conditions with a statistical significant value. It computes an F statistic that can be roughly described as a ratio of differences due to experimental manipulations to the error one finds in the sample, a p value which is the probability that the difference observed between experimental conditions occurred due to chance, and partial η^2 which is an approximation of how much variance in the overall dataset, the experimental manipulation accounts for [4].

Results were also analysed separately within each group so as to study in detail all the differences in the effect of mismatches between the interaction conditions (co-located and non co-located) and paired t-tests were used to determine significance differences among the blocks. For one participant of group NC, the output data file with the recorded information was corrupted. Therefore, results analysed for this group included the data of only fifteenth participants.

5.1 Presence

5.1.1 Global analysis

The ANOVA performed with presence (SUS factor) reported main negative effects of the haptic misalignment ($F_{2,28}=30.8$, $p<.001$, $\eta^2=.68$) and the auditory delay ($F_{1,29}=24.4$, $p<.001$, $\eta^2=.45$). The co-location factor was not found to be significant but its interaction with the other factors was significant (C*V: $F_{1,29}=4.1$, $p=.05$, $\eta^2=.12$) or nearly significant (C*A: $F_{1,29}=3.0$, $p=.09$, $\eta^2=.09$). Significant interaction was identified between auditory delay and haptic misalignment (H*A: $F_{2,28}=7.2$, $p=.003$, $\eta^2=.34$), which was also influenced by the co-location condition (C*H*A: $F_{2,28}=4.3$, $p=.02$, $\eta^2=.12$). Thus, in the following sections we analyse the cause of these interactions.

5.1.2 Influence of asynchronies in group C and group NC

Results of the presence analysis performed separately in each group are listed in Table 2. The spatial misalignment included in haptic feedback was significant in both groups (NC: $F_{2,14}=24.9$, $p<.001$, $\eta^2=.82$; C: $F_{2,13}=11.3$, $p<.001$, $\eta^2=.63$), but for group C differences among conditions were higher. Indeed, for this last group, a misalignment of 0.25 cm (H1) yielded a reduction in presence of nearly 50% higher than that found in group NC. Therefore, when haptic and visual workspaces were aligned, the sense of presence was more sensitive to haptic mismatches. Time delay in the auditory feedback was found to also have influenced significantly the presence results for both groups, but was found to be more negative in group NC (NC: $F_{1,15}=33.1$, $p<.001$, $\eta^2=.62$; C: $F_{1,14}=4.5$, $p=.05$, $\eta^2=.40$), as can be seen (Table 2) in the differences between average values (NA-A). The delay in the lighting of the button was the feedback mismatch with less influence in presence. Indeed, for group C, this delay had a positive effect.

We illustrate in Fig. 5 the average presence results obtained in each experimental condition. As can be seen, H1 condition had a more negative effect in presence for group C than it did for group NC. Nevertheless, in the co-located condition, the negative effect of H1 misalignment (0.25 cm) was compensated slightly with the presence of visual mismatches, as can be seen by comparing results of blocks H1V and H2V with results of blocks H1 and H2, respectively. For this reason, the delay in lighting the button was found as a positive influence in the main components analysis. In contrast, this compensation was not found within group NC.

Differences between trials were also numerically analysed with a Student T test. This last analysis corroborated (see Fig. 5) that the sensitivity to haptic misalignments was higher for group C. In this group, presence rates decreased significantly in blocks H1 (H1-S=-1.20, $p=.004$) and H2 (H2-S=-1.73, $p<.001$) compared to that achieved in block S, while for group NC, only conditions of block H2 (H2-S=-1.73, $p=.001$) yielded a significant reduction.

5.2 Task performance

5.2.1 Score

The ANOVA indicated that haptic ($F_{2,28}=8.4, p=.001, \eta^2=.37$) and visual ($F_{1,29}=4.7, p=.03, \eta^2=.14$) mismatches influenced negatively and significantly altered the score achieved. The time delay in the emitted button sound (auditory feedback) affected the score achieved to the lowest extent, although its absolute value was higher than that included in the visual delay.

On average, the score achieved was higher for group C (C: $M=3.9, SD=1.0$; NC: $M=3.5, SD=1.2$), but co-location was not a significant factor. Nevertheless, a nearly significant interaction was found between co-location and haptic factors ($F_{2,28}=2.8, p=.07, \eta^2=.16$), due to the less influence of haptic misalignment in the score achieved in group NC. Indeed, the analysis performed in each group (results listed in Table 3) indicated that none of the three kinds of mismatches influenced score results for group NC. A possible explanation for this finding is that a spatial misalignment of a higher order, such as the non coincidence between haptic and visual workspaces, reduced the possible effect in the score achieved of slight mismatches included in the feedback provided. Indeed, mismatches in the haptic feedback reduced significantly score results in the condition of coherence between workspaces ($F_{1,13}=6.96, p=.009, \eta^2=.51$).

Surprisingly, although without a significant main influence (See Table 3), time delay in the button lighting affected score results positively in both groups. Though ad hoc, a potential explanation for this could be that the delay in visual feedback compensates the negative effect of haptic feedback misalignments. This effect can be seen for both groups, in Fig. 6a.

By comparing these results with feedback mismatches influence in presence, the main concern raised is that the allowable thresholds of inaccuracy in the feedback provided are different. Thus, a further analysis of these thresholds should evaluate separately selection accuracy or sense of presence.

Table 2 Influence of selection feedback mismatches in the sense of presence (SUS) for group C and group NC. Average values of presence (M) in conditions with (H1 and H2) and without haptic misalignment (NH), with (A) and without (NA) auditory delay and with (V) and without (NV) visual delay

Group	Measure	Haptic			Auditory		Visual	
		NH	H1	H2	NA	A	NV	V
C	<i>M</i>	1.78	0.98	0.45	1.18	0.95	1.03	1.11
	<i>p</i> / η^2	.001^a / .63			.05 / .24		.32 / .07	
NC	<i>M</i>	1.48	1.09	0.53	1.28	0.79	1.12	0.94
	<i>p</i> / η^2	.001^a / .62			.001^a / .82		.09 / .17	

^a Values of significance lower than .001 ($p<.001$)

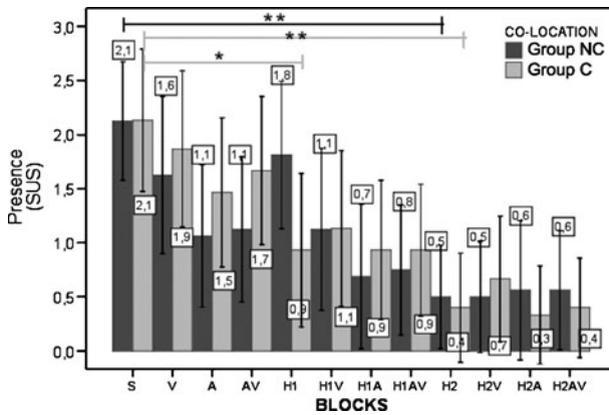


Fig. 5 Average results and errors (95%) of presence (SUS) obtained in each block of trials, for group C (co-located) and NC (non co-located) and certain remarkable significant differences with block S (** $p < 0.001$, * $p < 0.01$). Mismatches conditions are: S (synchronized), V (200 ms of visual delay), A (400 ms of auditory delay), H1 (0.25 cm of haptic misalignment), H2 (0.75 cm of haptic misalignment) and their combinations

5.2.2 Time between buttons pressings

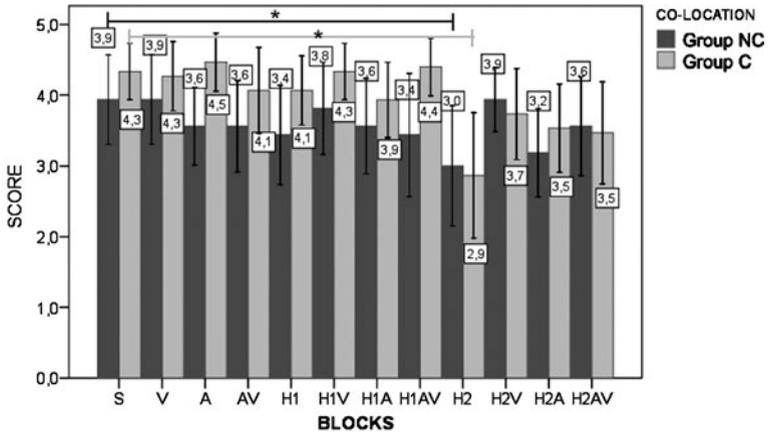
The global analysis performed with the time measurement indicated that the delay included in the auditory feedback had the most negative effect ($F_{1,29}=24.2, p < .001, \eta^2 = .66$). The delay included in the button lighting was also negative but less significant ($F_{1,29}=5.9, p = .02, \eta^2 = .17$). Surprisingly, the misalignment included in the haptic feedback did not yield a significant reduction of selection time. As in score results, performance was also better for group C, that is to say, lower selection times (C: $M=928$ ms, $SD=240$ ms; NC: $M=975$ ms, $SD=297$ ms), but without a significant difference.

The analysis performed separately for each group (see Table 4) reported that the delay in the auditory feedback was the most negative for both groups (C: $F_{1,14}=24.2, p < .001, \eta^2 = .63$; NC: $F_{1,15}=34.5, p < .001, \eta^2 = .69$) and that the effect of visual delay was only significant for group C ($F_{1,14}=9.7, p = .008, \eta^2 = .41$).

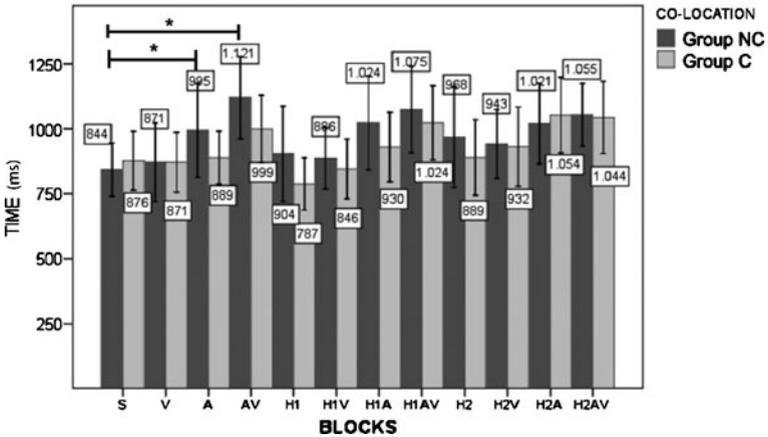
Average time results obtained in all the blocks for both groups are shown in Fig. 6b. In this figure, blocks with auditory delay can be clearly identified as those where participant’s responses were slower. This negative influence was more evident for group NC. Thus, the responsiveness of participants that interacted with incoherence between visual and

Table 3 Influence of selection feedback mismatches in the Score achieved for group C and group NC. Average values of presence (M) in conditions with (H1 and H2) and without haptic misalignment (NH), with (A) and without (NA) auditory delay and with (V) and without (NV) visual delay

Group	Measure	Haptic			Auditory		Visual	
		NH	H1	H2	NA	A	NV	V
C	M	4.28	4.18	3.40	3.93	3.97	3.86	4.04
	p / η^2	.009 / .51			.75 / .01		.08 / .19	
NC	M	3.75	3.57	3.42	3.66	3.49	3.44	3.70
	p / η^2	.25 / .17			.21 / .10		.15 / .13	



a Score: Number of sequences reproduced properly



b Time: Elapsed time between buttons pressings

Fig. 6 Average results and errors (95%) of task performance in each block of trials, group C (co-located) and NC (non co-located); and certain remarkable significant differences with block S (* $p < 0.01$). Mismatches conditions are: S (synchronized), V (200 ms of visual delay), A (400 ms of auditory delay), H1 (0.25 cm of haptic misalignment), H2 (0.75 cm of haptic misalignment) and their combinations

proprioceptive workspaces relied more in the coherence of audio feedback. Indeed, for group NC, the increment in the time elapsed between button pressings referred to values obtained in block S was significant in all the blocks with delay in the sound emission (for instance, A-S=150 ms, $p = .02$; AV-S=277 ms, $p < .001$).

5.3 Perceived delay

Results obtained about the perceived delay indicated that the three types of mismatches were detected by participants (Haptic: $F_{2,28} = 7.1$, $p = .003$, $\eta^2 = .33$; Auditory: $F_{1,29} = 41.9$, $p < .001$, $\eta^2 = .59$;

Table 4 Influence of mismatches in selection feedback in the time elapsed between button pressings for group C and group NC. Average values of presence (M) in conditions with (H1 and H2) and without haptic misalignment (NH), with (A) and without (NA) auditory delay and with (V) and without (NV) visual delay

Group	Measure	Haptic			Auditory		Visual	
		NH	H1	H2	NA	A	NV	V
C	<i>M</i> (ms)	908	896	979	866	989	904	952
	<i>p</i> / η^2		.12 / .27		.001^a / .63		.008 / .41	
NC	<i>M</i> (ms)	957	972	996	902	1,048	959	991
	<i>p</i> / η^2		.59 / .07		.001^a / .69		.27 / .07	

^a Values of significance lower than .001 ($p < .001$)

Visual: $F_{1,29}=34.7$, $p < .001$, $\eta^2=.54$). The misalignment included in haptic feedback was not perceived as the most negative influence, in spite of its more negative influence in presence and score rates. As expected, auditory delay was the most negative influence because it had the highest value in absolute terms. If we calculate absolute values of haptic asynchrony in time terms, its value was lower¹ than those included in auditory and visual feedback. Thus, perceived delay did not correspond with results obtained about the effect of these mismatches in presence and score measurements.

The difference between groups was neither significant, although the perceived delay was higher for group C than that for group NC (Group C: $M=3.2$, $SD=1.7$; Group NC: $M=2.7$, $SD=1.3$). This corroborated results obtained about the more negative influence of these feedback mismatches in a condition closer to natural interaction, with haptic and proprioceptive workspaces spatially coincident.

This analysis for groups also showed the existence of a significant interaction between haptic and auditory factors ($F_{2,14}=9.0$, $p=.003$, $\eta^2=.56$) for group NC. This interaction indicated that the negative influence of auditory delay was not altered by the haptic feedback condition. In contrast, for group C, the inclusion of more than one modality of asynchrony in the feedback increased perceived delay. These results can also be seen in Fig. 7 and indicate two aspects. First, the auditory delay in group NC captured the possible influence of the misalignment in haptic feedback. Second, for group C, the addition of more than one kind of mismatches in the feedback provided increased the delay perception, but it affected more presence and performance in relative terms than in absolute terms. Hence, the coexistence of more than one kind of mismatches led into a less negative influence.

6 Discussion

6.1 Influence of sensory mismatches

The analysis performed about the influence of sensory mismatches in the feedback provided to indicate selection has not confirmed hypothesis 1. We hypothesized that mismatches in the

¹ Note that average time elapsed between buttons pressings without any mismatch was around 800 ms and the average displacement over the Simon surface was around 10 cm. Thus, assuming a constant speed of 12.5 cm/s a spatial lag of 0.25 cm and 0.75 cm correspond with a time delay of around 20 ms and 60 ms respectively.

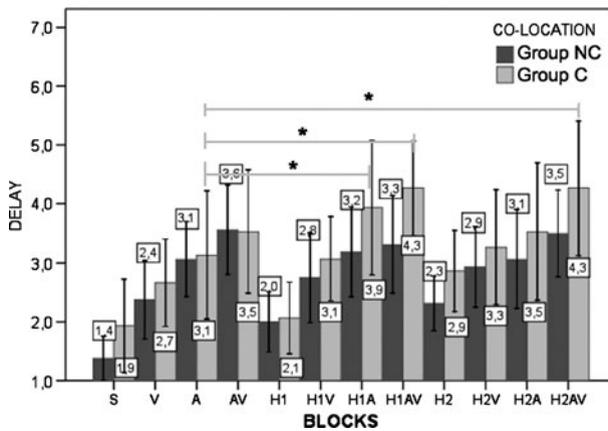


Fig. 7 Average results and errors (95%) of perceived delay obtained in each block of trials, group C (co-located) and NC (non co-located); and significant differences with block A ($*p < 0.01$). Mismatches conditions are: S (synchronized), V (200 ms of visual delay), A (400 ms of auditory delay), H1 (0.25 cm of haptic misalignment), H2 (0.75 cm of haptic misalignment) and their combinations

haptic feedback were the most negative influence in both, presence and performance, in a selection task.

This hypothesis has been confirmed by presence results. Mismatches in haptic feedback have appeared as the most negative influence, despite that of auditory cues being perceived as the main lack in the system responsiveness. The analysis performed has also highlighted the higher sensitivity of presence to sensory mismatches than task performance metrics. On the contrary, this hypothesis has not been confirmed attending to task performance. Participants' reactions measured through the time elapsed between button pressings have been clearly more influenced by the delay in auditory cues. These results are in accordance with those obtained in [19], where movement times were significantly increased by haptic delay only when values were higher than 200 ms. In contrast, the perceived delay has been negatively affected by inaccuracies in haptic feedback. These results agree with those found in [24], where detection thresholds of haptic delay were found to be lower than those values of performance deterioration reported by [19]. Nevertheless, score results obtained differ from those obtained in [19]. They did not find that any of the haptic delays analysed (within a range from 20 ms to 170 ms) significantly influenced results. Indeed, they stated that although haptic feedback had previously been shown to enhance performance in such a task, delaying it failed to have any effect because it was not crucial to the task in question. In contrast, the score achieved in this game, based on selecting correctly five buttons, was influenced by mismatches included (within the same range). This difference can be explained by the different task complexity, as concluded in [27] for lag in the visual feedback; the harder the task, the greater the detriment caused. In Jay [19] the task consisted of tapping between two targets in a 2D environment, while the tapping in our work is made in 3D conditions and consisted of selecting in order four different buttons.

Regarding audio lag, simultaneity thresholds between auditory-visual cues and between haptic-visual cues were analysed in [25]. The task considered in that study was based on evaluations made by two participants when one was hitting a drum. A blindfolded actor struck it while hearing the impact through headphones, with an offset from actual impact time, and an observer watched from an isolated room, hearing the same sound. These experiments indicated that the intermodal simultaneity thresholds were in both cases

(auditory-visual and haptic-visual) fewer than 100 ms. Thus, as the absolute value of auditory delay included in our experiment was higher (400 ms), we expected that the dependent factors analysed were affected negatively. This was confirmed by measurements of time per response and sense of presence. Nevertheless, audio lag did not influence the score, despite being a task very dependent on sound (sequence emitted with sound). Thus, auditory lag seems to be a decisive factor in the selection but not in the accuracy.

Visual lag during selection tasks has been evaluated in previous studies attending to the movement of the cursor displayed. As the lag included in this study corresponds with the visual feedback displayed to indicate selection (lighting of a button) and it is not referred to the tool rendering, we cannot directly relate results obtained. The experiment carried out has outlined that the inclusion of a delay in the lighting of the buttons may act by reducing the effect of mismatches in other feedback modalities. Thus, although the visual mismatch has been perceived as a delay in the system responsiveness, slowing down participants' responses, it has yielded a positive influence in presence when it was provided together with auditory cues and in the score achieved with haptic cues. So, it seems that the influence of mismatches acts more as a function of relative values among feedback modalities than in terms of absolute values. Based on a similar phenomenon, there is a technique based on including visual artefacts to reduce the impact of network latency within collaborative VE [7]. This technique consists in making the users aware about the state of the network lag changing for instance the colour of the cursor used.

6.2 Alignment between visual and proprioceptive workspaces

As stated in hypothesis 2, the co-location factor has arisen as a decisive factor in the influence of feedback mismatches in presence and task performance. First, results have indicated that when haptic and visual workspaces are coincident negative effects of synchrony among sensory cues provided to indicate selection are higher. Second, depending on this factor, the modality of feedback in which participants relies more to perform their interaction is different.

Despite not finding conclusive results about differences in the influence of mismatches in the system responsiveness between co-location conditions, perceived delay has always been higher in the co-located set-up in those conditions with significant difference between groups. This finding has become more evident regarding performance metrics. Thus, in the setup with a first order spatial misalignment, only auditory delay has appeared as a drawback and has only affected the selection time. In contrast, in a co-located condition both, the score achieved and the time elapsed, have been more negatively influenced by mismatches. These results bring us to conclude that in conditions of coherence between workspaces, the lack of quality in the feedback provided is more evident. As a step forward on conclusions previously obtained about the influence of spatial coincident on performance in different tasks [3, 11, 40], this study has provided empirical evidence that in selection tasks this correspondence, although desirable, relies on higher exigencies about feedback.

As for the sense of presence, results have stressed that the mental model of participants' interaction within the VE is different in the two setups evaluated. In a setup where participants' conditions of selection are closer to those of a real environment, the sense of presence relies more on haptic feedback coherence. In contrast, in a setup closer to interaction conditions with a 2D GUI, synchronization of auditory cues becomes more relevant for the sense of presence.

6.3 Sensorial recalibration

The score results obtained have shown a compensation effect between visual and haptic mismatches. Thus, in conditions where both kinds of mismatches were provided the score was higher than in conditions where only one was provided. As we have previously stated, the score has been influenced more by relative asynchronies than by absolute values. Nevertheless, this effect has not been found in selection times and perceived delay, measurements in which effects of delay in the auditory cues were the most negative ones. These results may indicate that this process of intersensorial recalibration happens only in those measurements where the haptic misalignment effect becomes more negative.

As described in Section 2.1, the effects of intersensorial mismatches between visual and auditory cues have been more extensively documented [17, 25, 43], within a range between 50 ms and 100 ms but mainly attending to very isolated stimuli as tones and light pulses. One of this study issues was the evaluation of the recalibration processes with a more typical stimuli of a VE. For the temporal values of delay analysed (200 ms and 400 ms), this experiment has not shown an intersensory recalibration process between auditory and visual stimuli, similar to that previously reported. In none of the measures, an auditory delay of 400 ms has captured the influence of a visual delay of 200 ms. Indeed, the sense of presence has not been altered or has slightly been increased due to the simultaneity generation of delayed visual and auditory cues. Therefore, this study provides slight evidence for the necessity for further study about not only intermodal simultaneity thresholds but also about the influence of simultaneous delay of different sensory cues in human factors as those analysed.

6.4 Implications for human–computer interaction design

Mismatches as those analysed within this study that occur in contemporary human–computer interaction have, up to the author’s knowledge, not been investigated before. The most relevant aspect extracted from the analysis performed about participants’ perceived delay between their selection actions and the system responses is that of the higher tolerance of mismatches in the feedback provided when visual and haptic workspaces are not spatially coincident. These results corroborate the conclusion obtained from a previous experiment in which we found that participants have a most critical attitude about deviations within the VE from typical behaviours and features of a real environment [44].

In the experiment reported here, the mismatches influence in the perceived delay has been very similar in the two set-ups considered, but average results in the different conditions have shown a more negative influence in the set-up with spatial correspondence between visual and haptic workspaces. Indeed, in the condition without induced mismatches participants were more critical with the perceived delay in the coincident condition.

Different studies [2, 4, 22, 31] have focused on sensory information quality within a VE quantifying it via participant’s perception instead of only considering technical parameters. In the study presented here, the measure included about participants’ perception of delay has shown that their awareness about the existence of mismatches in the sensory cues provided is not directly related with their influence in task performance and the sense of presence. This finding is not in the line with one of Pelligrini’s [31] assessments according to which, once the thresholds of human delay discrimination are established, they must not be broken by the simulation parameters. Our work is more in line with a second formulation made in that study, where it was stated that if the simulation accuracy and resolution cannot

be very high, mismatches in stimuli should be under annoyance thresholds, which in turn usually depends on the task performed within the environment.

In an environment as simple as the one used in this testbed, the simulation loop did not have to manage a high number of events and the number of objects to be rendered haptically and visually were very limited. Thus, in these environments the implications extracted from this study are not decisive because the stimuli can be generated without asynchronies being placed out of the perceptible range. Nevertheless, actual tendencies in VE design require several threads for managing the input and output information in several data formats and related not only to one user. In these situations, findings related with the sensory mismatches permissiveness beyond the perceptible thresholds are an interesting topic. In terms of potential computational cost savings, the findings of the study described may act in relaxing the critical points of synchronization among threads devoted to graphic and haptic rendering, or audio stimulation, such as in networked environments where special measures have been taken to thwart the possible effects of lag in real time interaction.

Indeed, results of this experiment have shown that although participants perceived all the source of mismatches included as a drawback in the system responsiveness, their performance and their sense of presence have not been influenced in the same way for these mismatches. As the auditory time delay introduced is the highest one, this has been the source of delay highlighted as the most negative for the system responsiveness. Nevertheless, misalignments introduced in the haptic feedback were more negative in the sense of presence and in the participants' ability of reproducing a sequence than the auditory delay.

7 Conclusions

The underlining issue of this work is to isolate the extent to which possible sensory mismatches may diminish the interaction within a VE, detecting the sensory modality of feedback that has to be provided more accurately, which in turn requires more design efforts.

One of the main findings obtained from the experiment made is that, for a selection task, there is a difference in the tolerance to mismatches depending on the correspondence condition between visual and proprioceptive workspaces. A difference that is more evident in the task performance. Thus, an interaction setup with spatial correspondence requires higher design efforts regarding the accuracy of the sensory feedback provided. Based on this finding, for a particular application without the necessity of reproducing exactly the interaction conditions of a real environment, efforts should not be paid in making coincident visual and proprioceptive workspaces if possible mismatches in the feedback provided as the system response to different actions cannot be reduced under the perceptible thresholds. Nevertheless, for a VE developed with the main goal of being used as a training platform to perform accuracy actions, the existence of correspondence between visual and haptic workspaces is a decisive factor, and therefore, possible mismatches in the feedback provided must be minimized.

The analysis of the sense of presence has outlined that if eliciting this factor is one of the simulation goals, haptic and visual renderings must be accurately synchronized to avoid possible mismatches in the haptic feedback provided. In contrast, for a simulation where the final goal is associated with the learning of skills directly related with selection actions made within the environment, the possible mismatches in the auditory

stimuli gain more relevance in measures associated with participants' time responses. Furthermore, those simulations with high lag in both visual and auditory cues provided as feedback to users' selections should give more priority to reduce delay in auditory cues.

This study has also revealed empirical evidence that human capacity in detecting different sensory mismatches is not completely associated with their influence in human factors, such as the sense of presence and the task performance. In other words, a person may detect the presence of mismatches in the perceived stimuli, but at the same time, their effect may not alter his/her sense of presence or ability to perform a task because they are assumed as inherent to the system or not enough annoyance to disturb these factors. Therefore, tolerance thresholds of mismatches in the sensory feedback provided can be higher than those found with studies focused in human detection thresholds.

Acknowledgements The authors wish to thank the participants in the study for their collaboration and comments. This work has been partially supported by DIANA group (University of Málaga) and by the University of Jaén through project UJA2009/12/12.

Appendix I. Presence Items

Appendix I. Presence Items

- I had the sensation of touching the Simon device which appears in the virtual environment

1	2	3	4	5	6	7
(Not at all)						(Very much)

- There were times during the experience when the virtual environment was a reality for me

1	2	3	4	5	6	7
(At no time)						(All the time)

- The virtual environment seems to me more like

1	2	3	4	5	6	7
(Images)						(A device)

- The delay that I experienced between my actions and the expected outcomes was:

1	2	3	4	5	6	7
(None)						(Large)

References

1. Allison RS, Harris LR, Jenkin M, Jasiobedzka U, Zacher JE (2001) Tolerance of temporal delay in virtual environments. In Proc. of IEEE Virtual Reality, pp 247–254. doi:<http://doi.ieeecomputersociety.org/10.1109/VR.2001.913793>
2. Ardito C, Costabile MF, Angeli A, Pittarello F (2007) Navigation help in 3D worlds: some empirical evidences on use of sound. *Multimed Tools Appl* 33:201–216. doi:[10.1007/s11042-006-0060-0](https://doi.org/10.1007/s11042-006-0060-0)
3. Arsenaault R, Ware C (2000) Eye-hand co-ordination with force feedback. In Proc. of the SIGCHI conf. on Human factors in computing systems, pp 408–414
4. Bailenson JN, Yee N (2008) Virtual interpersonal touch: haptic interaction and copresence in collaborative virtual environments. *Multimed Tools Appl* 37:5–14. doi:[10.1007/s11042-007-0171-2](https://doi.org/10.1007/s11042-007-0171-2)
5. Begault DR (2000) 3-D sound for virtual reality and multimedia. Facsimile reprint of 1994 edition, Preface to the NASA Technical Memorandum version, Ames Research Center, Moffett Field, CA, USA
6. Boring EG (1929) A history of experimental psychology. Appleton Century Crofts, New York
7. Boukerche A, Shirmohammadi S, Hossain A (2006) Moderating simulation lag in haptic virtual environments. In Proc. of the 39th Annual Symposium on Simulation, pp 269–277, IEEE Computer Society, Washington, USA. doi:[10.1109/ANSS.2006.31](https://doi.org/10.1109/ANSS.2006.31)
8. Burke JL, Prewett MS, Gray AA, Yang L, Stilson, FRB, Coovert MD, Elliot LR, Redden E (2006) Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis. In Proc. Multimodal interfaces, pp 108–117. ACM, NY, USA
9. Burns E, Razaque S, Whitton M, McCallus M, Panter A, Brooks FP (2005) The hand is slower than the eye: a quantitative exploration of visual dominance over proprioception. In Proc. of IEEE Virtual Reality, pp 3–10
10. Buttazzo G, Lipari G, Abeni L, Caccamo M (2005) Soft real-time systems: predictability vs. efficiency (Series in Computer Science). Springer, Berlin
11. Congedo M, Lécuyer A, Gentaz E (2006) The influence of spatial delocation on perceptual integration of vision and touch. *Presence-Teleoper Virtual Env* 15:353–357
12. De Boeck J, Raymaekers C, Coninx K (2005) Are existing metaphors in virtual environments suitable for haptic interaction. In Proc. of Virtual Reality, pp 261–268. Laval, France
13. Durlach PJ, Fowlkes J, Metevier CJ (2005) Effect of variations in sensory feedback on performance in a virtual reaching task. *Presence-Teleoper Virtual Env* 14:450–462
14. Ehmann S, Lin MC (2001) Accurate and fast proximity queries between polyhedral using convex surface decomposition. In Proc. of Eurographics, 20, pp 500–510
15. Fujisaki W, Shimojo S, Kashino M, Nishida S (2004) Recalibration of audiovisual simultaneity. *Nat Neurosci* 7:773–778
16. Harris CS (1965) Perceptual adaptation to inverted, reversed, and displaced vision. *Psychol Rev* 72:419–444
17. Hirsh IJ, Sherrick CE (1961) Perceived order in different sense modalities. *J Exp Psychol* 62:423–432
18. Ijsselstein WA, deKort YAW, Haans A (2006) Is this my hand I see before me? The rubber hand illusion in reality, virtual reality, and mixed reality. *Presence-Teleoper Virtual Env* 15:455–464. doi:[10.1162/pres.15.4.455](https://doi.org/10.1162/pres.15.4.455)
19. Jay C, Hubbard R (2005) Delayed visual and haptic feedback in a reciprocal tapping task. In Proc. of World Haptics, pp 655–656. doi:<http://doi.ieeecomputersociety.org/10.1109/WHC.2005.29>
20. Jay C, Hubbard R (2006) Quantifying the effects of latency on sensory feedback in distributed virtual environments. In Proc. of Virtual Images Seminar, pp 9–16. Paris, France
21. Keetels M, Vroomen J (2007) No effect of auditory-visual spatial disparity on temporal recalibration. *Exp Brain Res* 182:559–565. doi:[10.1007/s00221-007-1012-2](https://doi.org/10.1007/s00221-007-1012-2)
22. Larsson P, Västfjäll D, Kleiner, M (2001) Ecological acoustics and the multi-modal perception of rooms: real and unreal experiences of auditory-visual virtual environments. In Proc. of the Conf. on Auditory Display, pp 245–249, Espoo, Finland
23. Lécuyer A, Burkhardt JM, Coquillart S, Coiffet P (2001) “Boundary of illusion”: an experiment of sensory integration with a pseudo-haptic system. In Proc. of IEEE Virtual Reality, pp 115–122
24. Lee I, Choi S (2007) Discrimination of virtual environments under visual and haptic rendering delays, frontiers in the convergence of bioscience and information technologies, pp 554–562. doi:<http://doi.ieeecomputersociety.org/10.1109/FBIT.2007.124>
25. Levitin DJ, MacLean K, Mathews M, Chu L, Jensen E (2000) The perception of cross-modal simultaneity. In AIP Conf. Proc. 517, pp 323, doi:[10.1063/1.1291270](https://doi.org/10.1063/1.1291270)
26. Loomis JM (1992) Distal attribution and presence. *Presence-Teleoper Virtual Env* 1:113–119
27. MacKenzie IS, Ware C (1993) Lag as a determinant of human performance in interactive systems. In Proc. of the INTERACT '93 and CHI '93, pp 488–493, NY, USA. doi:<http://doi.acm.org/10.1145/169059.169431>

28. Morein-Zamir S, Soto-Faraco S, Kingstone A (2003) Auditory capture of vision: examining temporal ventriloquism. *Cogn Brain Res* 17:154–163. doi:10.1016/S0926-6410(03)00089-2
29. Munhall KG, Gribble P, Sacco L, Ward M (1996) Temporal constraints on the McGurk effect. *Percept Psychophys* 58:351–362
30. Navarra J, Soto-Faraco S, Spence C (2007) Adaptation to audiotactile asynchrony. *Neurosci Lett* 413:72–76. doi:10.1016/j.neulet.2006.11.027
31. Pellegrini RS (2001) Quality assessment of auditory virtual environments. In *Proceedings of ICAD, Espoo, Finland*
32. Radeau M, Bertelson P (1977) Adaptation to auditory-visual discordance and ventriloquism in semirealistic situations. *Percept Psychophys* 22:137–146
33. Rossetti Y, Koga K, Mano T (1993) Prismatic displacement of vision induces transient changes in the timing of eye-hand coordination. *Percept Psychophys* 54:355–364
34. Rothrock L, Barron K, Simpson T, Frecker M, Ligetti C, Barton R (2006) Applying the proximity compatibility and the control-display compatibility principles to engineering design interfaces. *Hum Factors Ergon Manuf* 16:61–81. doi:10.1002/hfm.v16:1
35. Ryu S, Kim H, Park J, Kwon Y, Jeong C (2007) Collaborative object-oriented visualization environment. *Multimed Tools Appl* 3:209–234. doi:10.1007/s11042-006-0066-7
36. Simpson TW, Barron K, Rothrock L, Frecker M, Barton RR, Ligetti C (2007) Impact of response delay and training on user performance with text-based and graphical user interfaces for engineering design. *Res Eng Des* 18:49–65. doi:10.1007/s00163-007-0033-y
37. Slater M, Usoh M, Steed A (1994) Depth of presence in virtual environments. *Presence-Teleoper Virtual Env* 3:113–119
38. Sousa Santos B, Dias P, Pimentel A, Baggerman J, Ferreira C, Silva S, Madeira J (2009) Head-mounted display versus desktop for 3D navigation in virtual reality: a user study. *Multimed Tools Appl* 4:161–181. doi:10.1007/s11042-008-0223-2
39. Spence C, Shore DI, Klein RM (2001) Multisensory prior entry. *J Exp Psychol Gen* 130:799–832
40. Swapp D, Pawar V, Loscos C (2006) Interaction with co-located haptic feedback in virtual reality. *Virtual Real* 10:24–30. doi:10.1007/s10055-006-0027-5
41. Szameitat AJ, Rummel J, Szameitat DP, Sterr A (2009) Behavioral and emotional consequences of brief delays in human–computer interaction. *Int J Hum Comput Stud* 67:561–570. doi:10.1016/j.ijhcs.2009.02.004
42. Theile G (1993) Trends and activities in the development of multichannel sound systems. In *Proc. of AES International Conf.*, pp 180–187
43. Vatakis A, Spence C (2006) Audiovisual synchrony perception for music, speech, and object actions. *Brain Res* 1111:134–142. doi:10.1016/j.brainres.2006.05.078
44. Viciano-Abad R, Reyes-Lecuona A (2008) Effects of Co-location and crossmodal interaction between haptic, auditory and visual cues in presence. *Haptics: Perception, Devices and Scenarios, LNCS 5024: 832–837*, Berlin: Springer. doi:10.1007/978-3-540-69057-3
45. Vroomen J, Keetels M, de Gelder B, Bertelson P (2004) Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Cogn Brain Res* 2:32–35. doi:10.1016/j.cog-brainres.2004.07.003
46. Watson B, Walker N, Ribarsky B, Spaulding V (1999) Managing temporal detail in virtual environments: relating system responsiveness to feedback. In *Proc. of CHI '99*, pp 280–281, NY, USA. doi:http://doi.acm.org/10.1145/632716.632888
47. Watson B, Walker N, Woytiuk P, Ribarsky W (2003) Maintaining usability during 3D placement despite delay. In *Proc. of IEEE Virtual Reality Conf.*, pp 133–140. doi:http://doi.ieeecomputersociety.org/10.1109/VR.2003.1191131
48. Witmer BG, Jerome CJ, Singer MJ (2005) The factor structure of the presence questionnaire. *Presence-Teleoper Virtual Env* 14:298–331. doi:10.1162/105474605323384654
49. Woszczyk W (1993) Quality assessment of multichannel sound recordings. In *Proc. of AES Int. Conf.*, pp 197–218
50. Yang U, Jounghyun KG (2004) Increasing the effective egocentric field of view with proprioceptive and tactile feedback. In *Proc. of IEEE Virtual Reality*, pp 27–34



Raquel Viciano-Abad received the PhD degree in 2009 and the M.Sc. degree in Telecommunication Engineering in 2001 from the University of Málaga. In 2002, she joined as a research member the Advance Human-Machine Interfaces and Virtual Reality Laboratory (DIANA group), University of Málaga. Since 2005, she is an assistant professor in the department of Telecommunication Engineering at the University of Jaén. Her specific research interests include user centre design and evaluation of interaction techniques for virtual environments and training applications based on virtual reality, mainly regarding haptic feedback, multimodality and the sense of presence.



Arcadio Reyes-Lecuona received the PhD degree from the University of Málaga, Spain, in 2001 and his M. Sc. degree in Telecommunications Engineering from the same University. He is a research member at the Advance Human-Machine Interfaces and Virtual Reality Research group (DIANA group), and head of the Virtual Reality and Human Computer Interaction Lab. at the University of Málaga. He is a member of the Spanish association of Human Computer Interaction (AIPO) and takes part in the European Network of Excellence INTUITION. His specific research interests include 3D interaction and evaluation of interaction techniques for virtual environments, including haptic feedback.



Matthieu Poyade received the Master degree in Telecommunications Engineering from the University of Málaga (Spain) in 2007 and a Master of Sciences degree in investigation in Telecommunications and Networking from the University of Lyon 1 (France) in 2005. He graduated a Diploma of Engineering in IT from the Mikkeli Polytechnics School (Finland) in 2004. He is currently realizing his Ph.D. studies in the School of Telecommunication Engineering, University of Málaga, supervised by Ph.D. Arcadio Reyes Lecuona. He is a research member at the Advance Human-Machine Interfaces and Virtual Reality laboratory (DIANA group). His current investigations deal with haptic manipulation in Virtual Environments.



José Escolano received the B.S. and M.S. degrees in telecommunications engineering in 2002 and 2004, respectively, and the Ph.D. degree in 2008 from the Technical University of Valencia, Spain. In 2004, he joined the Multimedia Communications and Signal Processing Group, Erlangen-Nürnberg University, Germany, for three months as a Visiting Researcher. During 2005–2007, he was with the Acoustic Technology Department, Technical University of Denmark, for one year as part of the European Doctorate in Sound and Vibration Studies (EDSVS) Marie Curie grant. Since 2006, he has been with the University of Jaén, Spain, as an Assistant Professor. His main research interests are focused in room acoustic simulation, numerical modeling, multichannel audio techniques, and electroacoustic transduction. He has published about 30 papers in journals and international conferences. Dr. Escolano is a member of the Audio Engineering Society (AES) and the Acoustical Society of America (ASA).