Impact of semantic aids on command memorization for on-body interaction and directional gestures — Source link ↗

Bruno Fruchard, Eric Lecolinet, Olivier Chapuis

Institutions: Université Paris-Saclay

Related papers:

- Vocabulary navigation made easier

- Intelligent talk-and-touch interfaces using multi-modal semantic grammars

- Cues, Mental Nodes, Cognitive Overhead, and HTML: A Theoretical Framework for Web Communication

- A pedagogical role for conceptual models in user interface design

- A novel user-based gesture vocabulary for conceptual design

Impact of Semantic Aids on Command Memorization for On-Body Interaction and Directional Gestures

Bruno Fruchard, Eric Lecolinet, Olivier Chapuis

Impact of Semantic Aids on Command Memorization for On-Body Interaction and Directional Gestures

Bruno Fruchard
LTCI, Télécom ParisTech,
Université Paris-Saclay
Paris, France

Eric Lecolinet
LTCI, Télécom ParisTech,
Université Paris-Saclay
Paris, France

Olivier Chapuis
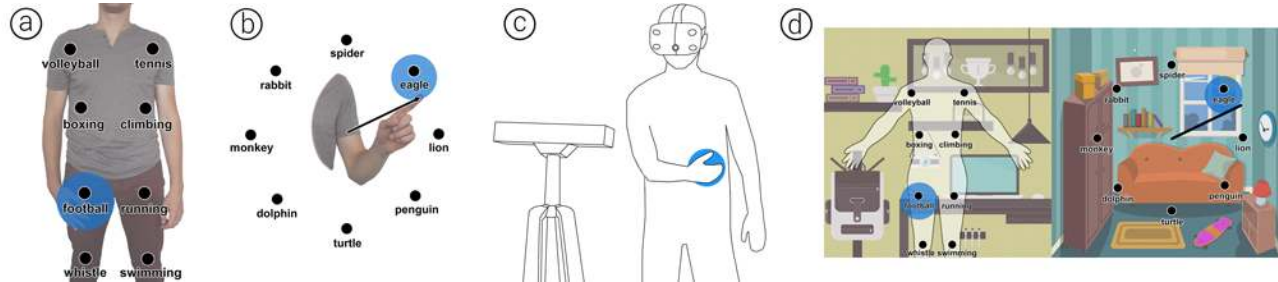LRI, Univ. Paris-Sud, CNRS, Inria,
Université Paris-Saclay
Orsay, France

Figure 1: (a) Using *BodyLoci*, the user selects a command by touching an area on his body; (b) using mid-air *Marking menus*, the user selects a command by performing a directional gesture in front of him; (c) hardware setup in the virtual reality context; (d) novice modes with background images for both techniques.

ABSTRACT

Previous studies have shown that spatial memory and semantic aids can help users learn and remember gestural commands. Using the body as a support to combine both dimensions has therefore been proposed, but no formal evaluations have yet been reported. In this paper, we compare an on-body interaction technique (Body-Loci) to mid-air Marking menus in a virtual reality context. We consider three levels of semantic aids: no aid, story-making, and story-making with background images. Our results show important improvement when story-making is used, especially for Marking menus (28.5% better retention). Both techniques performed similarly without semantic aids, but Marking menus outperformed BodyLoci when using them (17.3% better retention). While our study does not show a benefit in using body support, it suggests that inducing users to leverage simple learning techniques, such as story-making, can substantially improve recall, and thus make it easier to master gestural techniques. We also analyze the strategies used by the participants for creating mnemonics to provide guidelines for future work.

CCS CONCEPTS

• **Human-centered computing** → **Gestural input**;

KEYWORDS

Semantic aids; Memorization; Command selection; On-body interaction; Marking menus; Virtual reality

1 INTRODUCTION

Previous studies have shown that spatial memory and semantic aids can help users learn and remember commands [34, 35]. Such aids can make it easier and more enjoyable to learn the expert mode of command selection techniques, allowing users to interact more efficiently and more fluidly with user interfaces. Such techniques would be especially useful for mobile interaction or virtual reality, as no expert mode is generally available in such contexts.

The body provides natural landmarks that should support spatial memory and provide semantic information that might help to memorize commands (e.g. birthmarks) [5] when gestures are performed on the user's body. In this paper, we report the first study comparing an on-body interaction technique, BodyLoci (Figure 1-a), to a baseline technique, which is a mid-air variation of Marking menus [25, 54] (Figure 1-b). While Marking menus only rely on directional gestures, BodyLoci makes use of the body as an interaction surface, which might help retention for the above-mentioned reasons. This study was performed in a Virtual Reality environment (Figure 1-c) in an attempt to provide expert techniques in this context. Both techniques are well adapted because they do not require the user to see their hands when performing gestures in expert mode.

In a first experiment (n=24), we compared learning and retention for both techniques and found no difference in recall performance between them over two days. In a second experiment (n=24, different participants), we augmented both techniques with semantic aids. These aids are inspired by mnemonic devices such as the method of loci [52], which makes use of elaborative encoding and visual imagery to store information. These methods leverage several memory components (spatial memory, object/image memory, elaborative encoding [3, 7, 34]) which likely interact with each other in a way that enhances memorization [30, 33]. Building on these results, two different kinds of semantic aids were used in this experiment: (1) *Story-making* and (2) *Background images*. In the

first case, we incited users to create stories to strengthen mnemonic encoding [34, 35]. In the second case, we used background images in addition to story-making to see if this additional information would help users in creating stories (Figure 1-d).

Our results showed important improvement when story-making was used, especially for Marking menus (up to 28.5% better compared to the first experiment). While BodyLoci and mid-air Marking menus performed similarly in the first experiment, Marking menus outperformed BodyLoci in this experiment (17.3% better retention). However, background images did not provide a clear benefit.

The most striking result of this study is that just inducing users to leverage simple learning techniques can substantially improve recall. This suggests that such methods, which do not require strong effort and can even be seen as a sort of game, should efficiently help users in using gestural techniques. However, contrary to our expectations, our study does not show a benefit in using body support, probably because this technique was new and unusual for the participants. Moreover, Marking menus gained more performance than BodyLoci when semantic aids were used, which might be the result of a high mental demand as suggested in the discussion.

After presenting the related work and the techniques we used, we report on our two experiments then present an analysis of the strategies used by the participants for creating mnemonics. We propose a classification of these strategies to provide guidelines for future work and discuss our results, then conclude.

2 BACKGROUND

User interfaces mostly rely on *recognition*: the user must typically recognize the command she wants to perform in a set of buttons or in a list of menu items. However, *recall*, which does not require searching commands, is believed to provide better speed performance [38, 40] and better accuracy especially when items are small [37]. This is especially true for repetitive actions, which users might perform very often in their life. In this section, we present techniques designed to enhance recall and discuss their results to situate our work and justify our choices. We then focus on techniques for on-body interaction.

2.1 Spatial Memory

Previous work have focused on leveraging the *spatial positions* of commands and/or suggested that spatially constant interfaces should be favored [15, 21, 36, 38, 39]. To reinforce the aid provided by *spatial cues*, users might rely on implicit or explicit landmarks. To ameliorate accuracy while performing gestures, explicit landmarks might, for instance, be added to input surfaces [17, 51]. Interestingly, the human body provides natural landmarks that might be exploited by users for the same purpose [5, 45].

Recent work by Uddin at al. suggests that landmarks help users memorize command locations [42]. In their study, *spatial landmarks* provided better memorization performance than background images. However, as reported by the authors, participants may not have been aware of the presence of images. We try to avoid this problem in our study by providing participants with explicit instructions to evaluate whether displaying a background image helps mnemonic encoding.

2.2 Method of Loci and Semantic Aids

Combining several input channels or different types of memory is likely to produce better results than just relying on spatial memory, as suggested by Miller's study [30] and [3, 33, 34]. This aspect is also grounded by studies on mnemonic devices, as pointed out in the introduction. Usually, these techniques rely on a combination of strategies, which leverage several memory components [35, 52].

The method of loci is a mnemonic device that has been used since antiquity to acquire vast amounts of knowledge [52]. It requires the user to map the items to recall with locations in a well-known environment, such as the user's home or a famous building. A drawback of this method is that it requires creating a mental image of this environment, which requires important training. However, it has been used as a source of inspiration for an interaction technique where this demanding step is unneeded because the user interacts inside (and thus can see) this environment [34]. Building on the same principle, we propose to use the user's own body as an "environment" where she can map the commands to remember.

The method of loci does not only rely on spatial memory but also on images and on the idea of making stories to enhance memorization. Images may be or not related to the environment, and this method advises users to make use of striking or bizarre images to leverage *image memory* [8, 52]. More generally, the idea of using visual landmarks to leverage memorization has also been investigated in previous research in HCI [34, 36, 42]. Stories provide another kind of semantic aid. They serve as a means to relate these different pieces of information between them, rely on other memory components [3, 33] and involve deeper levels of encoding [12] than incidental learning.

Because most users are not aware of such methods, they are unlikely to use such strategies spontaneously. However, some users may intuitively use resembling strategies. In order to evaluate to which extent these kinds of semantic aids could help memorization, we performed the same experiment with and without semantic aids. In the latter case, during the learning phase, we instructed participants to create stories, either using the normal user interface or an augmented version displaying background images. Importantly, we gave no indications to the participants in the first experiment, but we explicitly informed them that these aids could enhance memorization in the second experiment, to see if this would make a difference.

2.3 Gestures and Incidental Learning

Keyboard shortcuts provide an expert mode on the PC, but rely on a different modality than pointing, thus requiring the expert mode to be learned explicitly. Marking menus [25] solve this problem by relying on gestures that are performed likewise in the novice or expert mode. This type of incidental learning helps users mastering the expert mode with little or no effort. Moreover, gestural shortcuts have been shown to provide better recall rates than keyboard shortcuts [2], whereas rhythmic shortcuts produced similar results [18]. Interestingly, both studies reported that some users created elaborated mnemonics. Yet, these studies did not focus on how participants created them, and, to our knowledge, this subject has not been systematically investigated in the HCI literature.

2.4 On-Body Interaction

To interact with a system, actions are usually performed using the hands without taking advantage of the whole body. Some studies investigated the body as an input surface to interact with by pointing on body areas to trigger actions [1, 19, 47] or store information [10]. These approaches, except the latter, do not require users to hold any particular device and allow eyes-free interaction by leveraging proprioception. This last point is particularly interesting in a virtual reality context where the user does not see her body.

As suggested earlier, natural landmarks can be exploited on the body to enhance the recall of items [5] (e.g. knuckles [49] or birthmarks [5]). Artificial landmarks may also be attached to the skin [48, 49], thus providing tactile input and visual output. The forearm is generally considered as particularly appropriate [27] as it is easy to access.

Interactions on the shoulders, ribs, and hips were also evaluated positively [23, 47]. Using landmarks and proprioception, users may also interact via imaginary interfaces [20, 22]. Finally, it is worth noticing that tactile feedback on the skin may be helpful [27] in the absence of visual feedback (i.e., eyes-free interaction).

Despite the possible benefits presented above, only few studies have investigated the use of on-body interaction to leverage command memorization. To our knowledge, this paper presents the first study that formally compares on-body interaction with a conventional interaction technique.

3 TECHNIQUES AND VR ENVIRONMENT

To compare the learning and retention of commands, we focused our evaluation on two gestural interaction techniques: an adapted version of *MarkingMenus* [25], which acts as a baseline, and a new on-body interaction technique called *BodyLoci*.

This choice was motivated by the following reasons. First, *MarkingMenus* is a well-known technique that provides an efficient expert mode, which makes it a good candidate for a baseline. Second, this choice allows a fair comparison between techniques because they both rely on gestural interaction, provide a novice and an expert mode (respectively involving recognition and recall) and enable incidental learning of the expert mode (the user performs similar gestures when interacting in both modes).

However, while sharing some similarities, these techniques rely on two different types of gestures. *MarkingMenus* use directional gestures, which can be seen as abstract (or arbitrary) gestures [50, 53] in the sense that they do not involve analogies with actions in the physical world (contrary, for instance, to pinch gestures). In contrast, as explained below, *BodyLoci* gestures require memorizing locations on the user's body. As seen above, using the human body may favor spatial memorization [1, 5, 19]. Body parts also involve semantics because they have a function and various characteristics (cf. examples in the Memorization Strategies section). We were thus interested in seeing whether using *body-related* or *abstract gestures* would make a difference.

3.1 Context: Virtual Reality Environments

We chose to perform our study in a VR environment. This context seems particularly interesting for on-body interaction since the user cannot see his body, but *proprioception* allows accurate pointing at
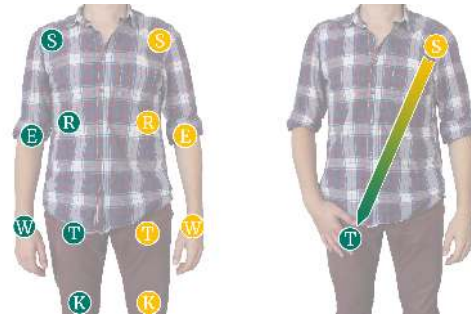

Figure 2: (left) Locations of the areas on the body (wrist, elbow, shoulder, ribs, thigh, and knee). (right) Example of a command selection: the user selects a menu with a 1st touch on an area, then the targeted command using a 2nd touch.

body parts without looking at them. Moreover, VR systems do not generally have shortcuts for command selection, such as hotkeys on a PC, despite interaction techniques have been proposed previously [14]. Thus, gestural interaction seems well adapted to offer fast interaction techniques in such environments.

Technically, we used an HTC Vive system [44]. This system continuously captures the position of the headset and of two controllers (one for each hand) that enable the user to interact with the virtual world. Each controller provides buttons (including a trigger) and a circular touchpad. In addition, we also used a Microsoft Kinect [29] to track the user body.

3.2 BodyLoci

Based on previous studies [23, 47], we selected 12 areas on the body that can be activated (Figure 2). As a sitting position facilitates access to some areas and is likely to reduce fatigue, we conducted our study in this situation. To enable a larger number of gestural shortcuts and favor memorization [6], this technique relies on hierarchical menus. Its design is inspired by [17] where the user must activate two locations on a PC touchpad to perform a selection. A first selection selects a menu and a second selection a command inside it. Both selections are performed by moving the hand close to the desired body location and activating a dedicated trigger (described below). Overall a maximum of 12×12=144 commands can be triggered.

Validation of the selection. To validate a selection we developed a simple solution shown in Figure 3. The HTC Vive controller is attached to the forearm of the user and the index finger is inserted into a sort of thimble (built using a 3D printer) that is attached to the trigger of the controller. By flexing the finger, the user can activate the trigger. This inexpensive solution makes it possible to lay the hand on the body (which provides benefits for accuracy [27] and likely memorization) and avoids asking the user to hold a device. Short vibrations are emitted when the user's hand is close to a body location. This facilitates interaction and avoids errors because the user know if his movements are properly detected by the system.

Novice and Expert Modes. BodyLoci gestures can be performed in both modes. In novice mode, a floating window displays the outline of a human body with the 12 possible body locations (Figure 4). Each location has a label that indicates the name of the associated
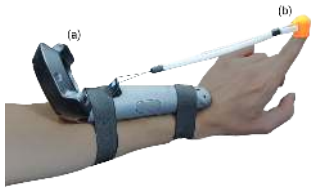
Figure 3: The system used to select areas on the body using a click. The controller (a) is attached to the user's wrist and a 3D printed model (b) is fixed on her finger with a string attached to the trigger to pull it.
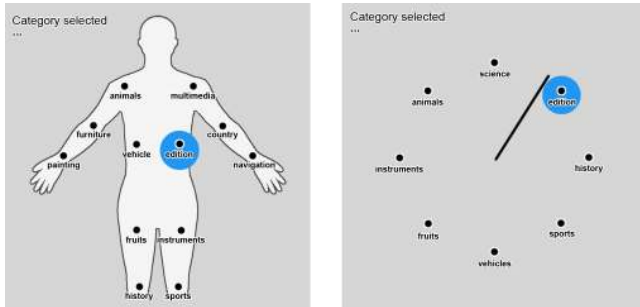
Figure 4: Novice mode for both interaction techniques (BodyLoci on the left, Marking menus on the right). The menu "edition" is highlighted before the user triggers a selection. A black stroke represents the user gesture for MM.

item. When the user points to her body, the closest body location is highlighted with a blue circle (Figure 4). This location (and its associated item) is selected by flexing the finger, as described above. The first selection opens a menu, the second selection activates a command in this menu. The menu is displayed in the same manner as the initial representation, with the menu labels located at the appropriate body locations. To cancel a selection, the user stays in a T-stance for one second.

The novice mode is triggered if the user keeps hovering over a body location for a delay of at least one second. Otherwise, no visual representation is displayed (expert mode). Importantly, as the user is wearing a VR helmet, he cannot see his own body. Our assumption was that proprioception would make it useless for users to look at their body to point to it.

3.3 Mid-air Marking Menus

We adapted *MarkingMenus* for interacting "in the air": The user just has to perform a simple directional gesture in front of her, by moving her arm in the air in the eight possible directions (horizontally, vertically or diagonally) to select the desired item. A gesture can be performed anywhere, without the need to start it from a central position.

As with *BodyLoci*, we used two-level menus to have a sufficient number of shortcuts. Instead of using compound mark [25] we used successive marks as in Multi-stroke Marking menus [54]: one for selecting a submenu, and one for selecting an item in this menu. We used successive straight marks rather than "zigzag" marks because the latter was shown to be less accurate in the 2D case and may be even harder to perform accurately "in the air".

To perform a marking gesture, the user must hold the Vive controller in his hand and pull the trigger while performing the gesture. The desired item is selected when releasing the trigger. As with the *BodyLoci* technique, the user can perform the gesture immediately (expert mode) or wait for a delay of one second to make the menu appear in a floating window (novice mode), as shown in Figure 4. The item that will be selected when releasing the trigger is highlighted with a blue circle and a black stroke is drawn between the center of the menu and this item. As previously, the controller emits short vibrations when an item can be selected.

4 BODYLOCI VS. MARKING MENUS

Our goal in this first experiment was to compare benefits offered by on-body interaction and directional abstract gestures on the memorization process. We thus compared the two above-described techniques (*BodyLoci* and *MarkingMenus* - factor Tech). Since the body provides spatial landmarks and associated semantics, we hypothesized that *BodyLoci* should provide better memorization performance than *MarkingMenus* (H1).

Participants and Apparatus. We recruited 24 participants in our local universities. Overall we had a population aged from 21 to 41 (mean 27) counting 11 women. We ran the experiment using the Unity game engine [43] with the VR system described above.

Design and Procedure. We used a within-subject design with factor Tech. We blocked by Tech, counter-balancing the presentation order: half of the participants started with *BodyLoci* and the other half started with *MarkingMenus*.

As proposed in previous work [2, 32, 36, 46], we used two sessions separated by an interval of 24 hours (Figure 5). The first session was intended to evaluate immediate memory and the second long-term memory. Sessions were made of *learning* and *recall* blocks. In *Learning* blocks, participants could trigger the novice mode of the techniques to learn command locations or perform them faster using the expert mode if they recall their locations. In *Recall* blocks, expert mode is enforced and no information is provided to the participants except whether the selection is correct or not. The first session consisted of three consecutive pairs of *learning*/*recall* blocks (named L1, R1, L2, R2, L3, R3) for each technique (see Figure 5). The second session consisted of a recall block, a learning block and again a recall block (R4, L4, R5). R4 serves to evaluate retention after a 24 hours interval, and R5 how efficiently users re-learn commands in each condition (Figure 5).

As explained in the previous section, commands were organized in a two-level hierarchy to allow the selection and memorization of a reasonably large number of commands (4×4=16 in our experiment as explained below). To match the design of *MarkingMenus* and limit each level to 8 items as recommended [25], we removed the areas on the arms ("E" and "W" on Figure 2).

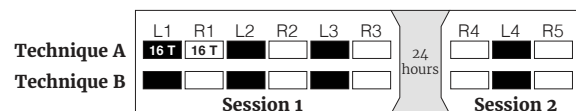|	L1	R1	L2	R2	L3	R3		R4	L4	R5
Technique A	16 T	16 T					24 hours			
Technique B										

Figure 5: Learning and Recall blocks performed by users for each technique on two sessions over two consecutive days. Each block is composed of 16 trials.

ANIMALS	**eagle, monkey, turtle, lion** dolphin, spider, penguin, rabbit		MULTIMEDIA	**eject, volume, play, record** shuffle, mute, next, album
EDITION	**cut, open, delete, save** paste, align, new, find		PAINTING	**canvas, line, palette, spray** layers, brush, curve, ruler
FRUITS	**apple, cherry, melon, orange** lemon, kiwi, grape, peach		VEGETABLES	**olive, mushroom, pumpkin, eggplant** cabbage, artichoke, zucchini, broccoli
SPORTS	**tennis, climbing, football, boxing** volleyball, swimming, running, whistle		NAVIGATION	**traffic, bus, car, orientation** walk, direction, hotel, parking
DISTRACTORS	VEHICLES, INSTRUMENTS, HISTORY, SCIENCE		DISTRACTORS	HOUSE, LITERATURE, COUNTRY, GARMENT

Figure 6: The two sets of categories used during the experiment consist of 4 categories with target items (in bold) and 4 categories used as distractors.

The first level of the hierarchy consisted of 8 categories (4 categories actually used and 4 categories used as distractors). Each category contained 8 items semantically related to this category (e.g. "apple" and "orange" in category "fruits") so that no ambiguity was possible when searching for the category of a command. Two equivalent sets of categories were created for the experiment (see Figure 6).

The positions of the categories and of the commands inside categories were specified before the experiment and were the same for all participants. The two techniques (TECH) and the two sets of categories were counterbalanced over participants. For each technique, participants had to learn the position of 16 commands (4 target commands inside each of the 4 non-distractor categories).

Before the first phase of each technique in both sessions (i.e. L1 or R4), participants started with a training block to get used to it. During training, categories and commands had abstract names such as "menu1" and "item13". The training block lasted until participants felt ready to continue.

To start a trial, participants had to position their dominant hand on a floating window in front of them. This window then disappeared and participants could see the command they had to find and select. When the right category was selected, a sound was played and its name was shown at the top left corner of the floating window displaying the menus (Figure 4). When a command was selected, another sound was played and a green or red square blinked whether the selection was correct or not. Only this last feedback was provided during the recall blocks.

The first session lasted approximately 1 hour and the second session 30 minutes. The first session ended with a questionnaire and the second session with an interview. The goal of this interview was to understand how participants memorized commands and what strategies they adopted.

4.1 Results

For conciseness, we do not report ANOVA results but only paired t-tests (i.e. post-hoc t-tests) with Cohen's d effect size, and indicate 95% confidence intervals on the graphs. We first compare recall rates for the two techniques. Figure 7-a shows that *MarkingMenus* and *BodyLoci* provide similar performance for each phase, except in the first phase where *MarkingMenus* perform better than *BodyLoci*. The difference in performance is significant for R1 ($p = 0.023$, $d = 0.57$), but not for phases R2 to R5 (all p's > 0.70). Thus, hypothesis (H1) is not supported by our results.

Except for the first learning phase (L1, Figure 7-b), the average time of a trial in the learning phases is significantly higher for *BodyLoci* than *MarkingMenus* (all p's < 0.002 and $d > 0.72$, but $p = 0.41$ for L1). The difference is especially important for the second phase (L2) and cannot be solely explained by a longer execution time of *BodyLoci* [1].

However, both results may have been affected by the fact that users needed time to master this novel and unconventional technique, especially at the beginning of the experiment (for instance the difference in time between techniques was more than 3 times higher for L2 than for L5). In comparison, *MarkingMenus* are much more similar to common user interfaces, which participants have been using for years. This explanation is also supported by the participants' interviews (see below).

Subjective results. At the end of the first session, participants were asked to fill a questionnaire. We analyzed these results using non-parametric t-tests. First, participants had to report which technique they preferred (with the possibility to have no preference). Overall 58.3% preferred using *MarkingMenus* and 20.8% *BodyLoci*. This difference is significant ($p = 0.04$), and most of the participants found *MarkingMenus* easy to manipulate and closer to conventional interactions (i.e. pointing).

The rest of the questionnaire consisted of Likert-scales using 7 levels with questions about *comfort*, *enjoyment*, *fatigue*, and *mental demand* (see Figure 7-c). We also asked participants about their *perceived performance* on the last recall phase and their *perceived progression* between the first and last recall phases (Figure 7-c). We found that *MarkingMenus* was significantly better for *comfort* ($p < 0.01$), *fatigue* ($p = 0.01$), and *recall rate* ($p = 0.02$). A trend was observed for *progression* ($p = 0.08$) and *mental demand* ($p = 0.06$). This last trend supports the hypothesis that the on-body interaction technique required more attention than *MarkingMenus*, presumably because participants were not accustomed to this way of interacting.

---

[1] The execution time of a trial is about 4 s for *MarkingMenus* and 5 s for *BodyLoci*, which requires larger movements.
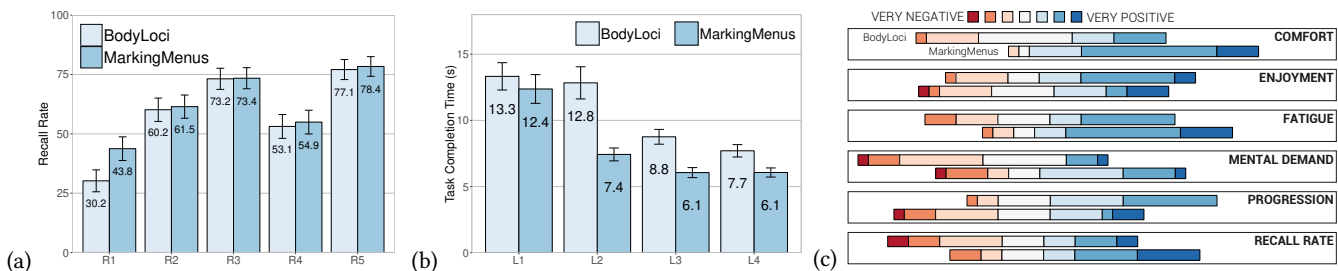

Figure 7: First experiment results by TECH: (a) recall rate for each recall phase, (b) task completion time for each learning phase, (c) results of the Likert-scales from the questionnaire.

Finally, we asked participants if they would use *BodyLoci* without the constraints of the apparatus used for the studies. For this question we used a Likert-scale with levels from 1-"Completely Disagree" to 7-"Completely Agree". The answers were positive (median = 5.5) with mixed justifications. Some of the participants saw in this technique a way to get rid of remote controllers ("it is convenient and you don't need special equipment for that" p12), or found it convenient in the case "your hands are busy [and] you activate a command with your forearm over your chest" (p11). A participant who would disagree using this technique said: "I find counter-intuitive and difficult to associate several functions for each area on the body" (p17).

5 IMPACT OF SEMANTIC AIDS

As said in the Background section, users can memorize a large number of items relatively easily over the long term [8, 28] when provided with adequate strategies (e.g. mnemonic devices). Since users can be creative when explicitly asked to create mnemonics [13], but may not do it spontaneously, we were interested in comparing groups of users instructed, or not, to create mnemonics. For this purpose, we conducted a second experiment, with the same design as the first experiment, but focusing on the effect of semantic aids. Importantly, while we gave no indications to the participants in the previous experiment, in this experiment we explicitly informed them that these aids could enhance memorization.

We used two different kinds of semantic aids, *Stories* and *Stories+Images* (factor AIDS):

(1) *Stories*: we instructed participants to create *Stories* about the command/position couples they had to remember with examples such as "if *rocket* is the command on top of the marking menus, imagine that you are launching it" or "if the command *camera* is on your shoulder, imagine that you are shooting a video with it".

(2) *Stories+Images*: we gave the same instructions but also added a *background image* to the graphical representations of the menus in novice mode (Figure 1-c) in order to provide more materials for users to create mnemonics.

We hypothesized that the participants of this experiment (context *Semantics*) would perform better than the participants of the first experiment (context *Baseline*) (H2), and also that participants aided with images (*Stories+Images*) would perform better than those only instructed to create stories (*Stories*) (H3).

Participants and Apparatus. We recruited 24 new participants in our local universities. Overall we had a population aged from 20 to 49 (mean 27) counting 7 women and 17 men. We used the same apparatus than in the first experiment.

Design and Procedure. We blocked by TECH within participants, counter-balancing the presentation order as in the first experiment and made AIDS a between-subject factor to avoid a transfer from *Stories+Images* to *Stories*.

In other words, we re-run the first experiment, but half of the participants under the *Stories* AIDS, and the other half under the *Stories+Images* AIDS. We used the exact same procedure than in the first experiment and the sessions lasted the same amount of time.

5.1 Results

In order to compare the conditions of our between-subject design, we performed unpaired t-tests. Figure 7-a shows the recall rates of both *Stories* and *Stories+Images*. We can observe that they are very close in all phases (no significant differences, all p's > 0.48). This is also the case for each technique taken alone (all p's > 0.20, most of them large). In other words, we have no interaction between AIDS and TECH. Regarding task completion times in learning phases (see Figure 7-b), they are also very close except for L1 ($p = 0.026$, $d = 0.97$; all p's > 0.6 for L2-L4). The difference for L1 suggests that *images* might help to initiate the creation of mnemonics. Thus, overall, our results do not support (H3): *background images* (used in addition to *stories*) may help at the beginning of the memorization process but do not seem to significantly improve memorization afterwards.

As *Stories* and *Stories+Images* lead to very similar results, we now compare the first and the second experiments (i.e., *Baseline* vs. *Semantics*, two groups of 24 participants) to evaluate H2. Figure 8 and Figure 9 show the results by TECH. First, the overall recall rate is significantly higher for *Semantics* than for *Baseline* for all phases but R1 (from R1 to R5: $p = 0.116$; $p = 0.027$, $d = 0.66$; $p = 0.014$, $d = 0.74$; $p = 0.015$, $d = 0.73$; $p = 0.025$, $d = 0.67$). Moreover, there are no significant differences in task completion time for the learning phases, even when comparing by TECH (all p's > 0.19). Thus, our results support (H2).

More precisely, as can be observed in Figure 8, there is an interaction with TECH on recall rate: the differences between *Semantics* and *Baseline* are more important for *MarkingMenus* than for *BodyLoci*. In fact, differences are not significant for *BodyLoci* and only show trends for R1 and R3 ($p = 0.097$, $p = 0.423$, $p = 0.076$; $p = 0.184$, $p = 0.145$). In contrast, all the differences are significant for *MarkingMenus* except for R1 ($p = 0.328$; $p = 0.006$, $d = 0.83$; $p = 0.014$,
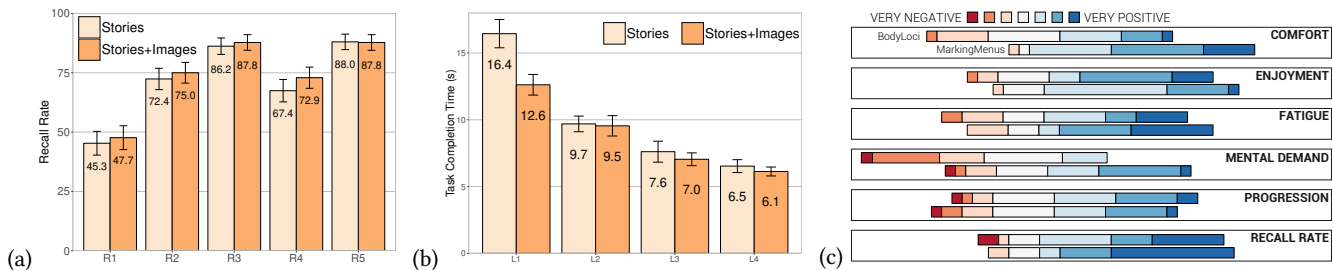

Figure 7: Second experiment results by AIDS: **(a) recall rate for each recall phase, (b) task completion time for each learning phase. (c) Results of the Likert-scales from the questionnaire for each** TECH.
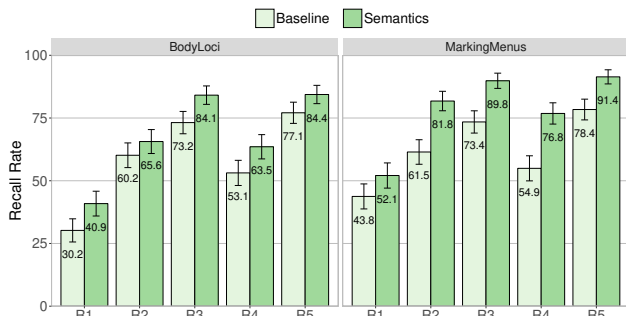
Figure 8: Comparing *Baseline* and *Semantics* recall rate for each TECH.
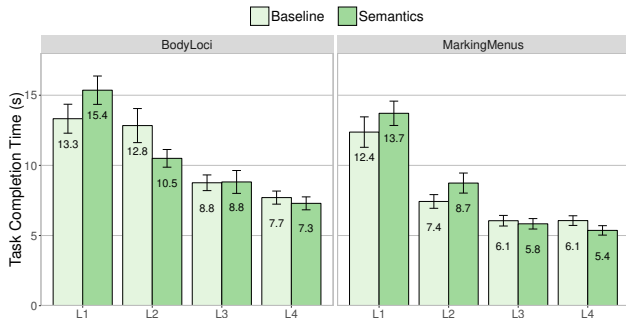

Figure 9: Comparing *Baseline* and *Semantics* average time of a trial in the learning phases for each TECH.

$d = 0.74$; $p = 0.002$, $d = 0.94$; $p = 0.011$, $d = 0.77$) and these differences are large, e.g., an improvement of 18.3% for R3 and of 28.5% for R4 (retention). Thus, inciting users to create *stories* substantially improves memorization with *MarkingMenus*, and there is a trend suggesting that *BodyLoci* also benefits from this mnemonic aid.

In accordance with the previous result, *MarkingMenus* performed better than *BodyLoci* in this second experiment (Figure 8). The differences are significant for R2 ($p = 0.009$, $d = 0.73$) and, interestingly, for R4, the retention phase ($p = 0.012$, $d = 0.60$) and for R5 ($p = 0.027$, $d = 0.51$), but not for R3 ($p = 0.160$). Regarding task completion times in learning phases, the participants took significantly more time with *BodyLoci* than with *MarkingMenus* (all $p's < 0.001$), with differences about $2\,s$ in all phases (see the "semantics" bars in Figure 9). This difference in time is similar to what was observed in the first experiment and can presumably be explained by the same reason (as the second experiment involved different participants, they needed time to master *BodyLoci*).

Subjective results. For this study, 58.3% of the participants preferred using *MarkingMenus* over 33.3% for *BodyLoci* ($p = 0.21$) (instead of 58.3% vs. 20.8% in the previous experiment). Results are shown in Figure 7-c. *MarkingMenus* was again preferred on several aspects (*comfort* $p < 0.01$; *mental demand* $p < 0.01$). Differences for the other aspects revealed some trends in accordance with the previous experiment (*fatigue* $p = 0.08$, *recall rate* $p = 0.09$, *progression* $p = 0.09$). However, overall participants better appreciated *BodyLoci* than in the previous experiment, although they were more efficient using *MarkingMenus* (which was not the case in the previous experiment). Perhaps semantics aids made using *BodyLoci* more enjoyable.

We also asked participants if they would like using *BodyLoci*. Most participants showed interested in using this technique (median = 5.5 on a 7 level Likert scale), but some had mixed feelings. For instance, a participant reported "I find this way of memorizing intuitive" (p2), but another said "I don't see myself using this technique outside a video game context" (P14).

6 MEMORIZATION STRATEGIES

At the end of each experiment, we asked participants about the strategies they used. More precisely we asked them the following questions: (1) *what strategies did you use to memorize commands*, (2) *did you change these strategies during the experiment* and (3) *did you find some commands or categories easier to memorize than others*. The results (summarized in Table 1) show similar patterns between participants and do not seem to depend on techniques (not clear differences were found). We list the main strategies below:

Command grouping. More than two thirds of the participants (67% for Exp. 1; 71% for Exp. 2) created chunks of commands, which could either only include a couple (e.g. "the *eagle* is flying above the *monkey*" Exp1-P2) or all the commands inside a category (e.g. "the positions of the commands formed a shape and I repeated their names sequentially to recall the position of each" Exp1-P7).

Visual imagery. More than half of the participants (58%; 58%) learned the commands by visualizing them (e.g. "I imagined the *lion* attacking my ribs" Exp2-P15) or memorized layouts of commands displayed on the interface. On the other hand, 29% reported having trouble memorizing abstract commands, mainly from the *"Edition"* category. This varied across participants as they were more or less familiar with different commands, depending on their background.

Gesture memorization. With less consistency, 40% of the participants (33%; 46%) used *gestures* to memorize categories and commands. Some of them said that gestures would come to their mind "automatically" when asked to select certain commands. Gestures demanding more effort seemed to be better memorized (e.g. "*Play* was particularly annoying because I had to make a lot of efforts to reach it" Exp1-P2), as remarked in previous work (e.g. [11]).

Remarkable positions. The position of the commands was another important characteristic: 35% of the participants (33%; 38%) said that a command was easier to recall if it was at the same position than its category (or towards the opposite direction with *MarkingMenus*). In addition, about 17% of the participants (4%; 29%) grouped commands that were semantically unrelated, but located nearby (e.g. "the *monkey* is on the bag playing with *oranges*" Exp2-P19).

Finally, we asked participants how they took advantage of background images. Over the 12 participants who performed under this condition, 7 of them said they used them with *MarkingMenus*, but only 2 with *BodyLoci*. Some participants did not pay attention to the images and/or perceived them as a decorative background (despite our instructions). Other participants took into account the color ("the *monkey* has the same color as the *cupboard*" Exp2-P1) or the shape ("the *cactus* has the shape of a *spray*" Exp2-P15). Another common strategy was to relate commands and background images semantically (e.g. "*open* the *window*" Ex2-P1 and "*mushroom* grows in the *forest* outside" Exp2-P8).

Table 1: Memorization strategies adopted by participants retrieved from the interviews of both experiments (e1, e2).

Strategy	Definition	Example	Participant (e1, e2)
Command grouping	*Creation of chunks of commands* [30]	"the *eagle* is flying above the *monkey*" (e1-p2)	69% (67%, 71%)
Visual imagery	*Use of visual imagery to recall commands* [24]	"I paint with the *palette* in my left hand" (e2-p13)	58% (58%, 58%)
Gesture memorization	*Memorizing gestures to reach the command*	"I *cut* something to reach the command" (e1-p17)	40% (33%, 46%)
Remarkable positions	*Relying on similar or opposite positions/directions with the category*	"two times in the same direction to select *sport* and *boxe*" (e1-p12)	35% (33%, 38%)

7 DISCUSSION

Our study compares body-guided gestures and mid-air directional gestures for command memorization. On the other hand, it compares different levels of semantic aids. Overall, it provides comparative results involving on-body interaction and semantic aids that are yet either rare or missing in the current literature. Its most compelling result is that a simple instruction inciting users to create stories substantially improved memorization: up to 13.1% for *BodyLoci* and 28.5% for *MarkingMenus*. This suggests that inducing users to leverage memorization strategies could have an important impact on user interfaces. For instance, providing hints or examples while using a graphical interface could help users to master gestural techniques, and thus to popularize such techniques.

Another interesting result is the efficiency of Marking menus and the fact that they benefited more from semantic aids than BodyLoci. While Marking Menus are often claimed to be very efficient, few studies have actually evaluated their *memory* performance (to our knowledge [4, 34], but in specific cases). This study confirms their efficiency (e.g. 73.4% without aids and 89.8% with semantic aids after only 3 learning phases). The added benefit of semantic aids suggests that this technique relying on "abstract" gestures strongly benefits from the association with concrete concepts (e.g. *visual imagery* [24] was used by 40% of the participants).

BodyLoci obtained similar performance than Marking menus without aids, which also shows the effectiveness of this technique. However, contrary to our hypothesis H1, it did not outperform them. Considering that body interaction should leverage spatial memory, this suggests that *directional* gestures also efficiently take advantage of this memory component. Moreover, as previously said, BodyLoci was hampered by its novelty. It required more effort from participants than Marking menus, especially at the beginning of each experiment, because users had to master this unconventional way of interacting. This may also explain why participants generally preferred Marking menus (although they also mainly said they would like using the BodyLoci technique). However, results may be different in the long term. Marking Menus are quite similar to common user interfaces, which participants have been using for years. It would thus be interesting to conduct a longer study to evaluate the performance of the BodyLoci technique when users are really accustomed to it.

The fact that semantic aids had lesser effect with BodyLoci than with Marking menus can be explained by the previous reason; because of a higher cognitive charge, using these aids efficiently may have been more difficult for users (e.g. some user did not even notice background images, especially when using BodyLoci). Another possible explanation is that these aids may be partly redundant when using BodyLoci because body parts involve semantic information that users may use spontaneously, contrary to Marking menus that rely on abstract gestures.

Finally, adding background images did not yield noticeable improvements, possibly because participants were already using visual imagery or focusing on the command selection task. As a consequence, they may have been overloaded with information [16, 31], or just considered images as a decoration (as mentioned during the interviews), a problem known as *selective attention* [9, 26].

8 CONCLUSION AND FUTURE WORK

In this paper, we presented the first memorization study that compares an on-body interaction technique to a conventional one. We reported the design of *BodyLoci*, an on-body interaction technique favoring gestural shortcuts in VR. This experiment with 24 participants showed the efficiency of both techniques for memorizing gestural commands. However, while on-body interaction has been speculated to specifically favor memorization, both techniques performed similarly.

A second experiment based on the same design with 24 other participants evaluated different levels of semantic aids (creating stories while memorizing commands *with* or *without* background images). Using background images did not provide a significant improvement. However, comparing the two experiments highlighted a compelling benefit of instructing users to create stories. This benefit was particularly strong for the retention phase (24 hours after learning) of *MarkingMenus*, with an improvement of 28.5%. This result suggests that inducing users to leverage memorization strategies could significantly help them in using interaction techniques in expert mode.

Finally, after analyzing the results of the interviews conducted after each experiment, we classified the memorization strategies elaborated by the participants (command grouping, visual imagery, gesture memorization, etc.). The elicitation of these strategies provide guidelines that could help designing techniques that facilitate the creation of mnemonics.

In the future, we would like to better understand the potential of *BodyLoci*. As we performed our evaluations with novice users, we plan to conduct longer evaluations to better assess the impact of user expertise. Such an evaluation may also help evaluating the effect of muscle memory [41], a phenomenon which might explain the "automatic" recall of gestures reported by some participants while performing *BodyLoci* gestures. Finally, we would like to develop techniques inducing users to leverage memorization strategies (such as creating stories), as this could make gestural techniques more efficient and more pleasant to use.

9 ACKNOWLEDGMENTS

REFERENCES

[1] Jussi Angeslevä, Ian Oakley, Stephen Hughes, & Sile O'Modhrain. 2003. Body mnemonics - portable device interaction design concept. *UIST '16 Adjunct.* ACM, 2 pages.

[2] Caroline Appert & Shumin Zhai. 2009. Using strokes as command short-cuts: Cognitive benefits and toolkit support. *CHI '09.* ACM, 2289–2298. **10.1145/1518701.1519052**

[3] Alan D. Baddeley. 2013. *Essentials of Human Memory (Classic Edition).* Psychology Press, East Sussex, UK.

[4] Gilles Bailly, Eric Lecolinet, & Laurence Nigay. 2008. Flower menus: A new type of marking menu with large menu breadth, within groups and efficient expert mode memorization. *AVI '08.* ACM, 15–22. **10.1145/1385569.1385575**

[5] Joanna Bergstrom-Lehtovirta, Sebastian Boring, & Kasper Hornbæk. 2017. Placing and recalling virtual items on the skin. *CHI '17.* ACM, 1497–1507. **10.1145/3025453.3026030**

[6] Gordon H. Bower, Michal C. Clark, Alan M. Lesgold, & David Winzenz. 1969. Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning and Verbal Behavior* 8, 3, 323–343. **10.1016/S0022-5371(69)80124-6**

[7] Gary L. Bradshaw & John R. Anderson. 1982. Elaborative encoding as an explanation of levels of processing. *Journal of Verbal Learning and Verbal Behavior* 21, 2, 165 – 174. **10.1016/S0022-5371(82)90531-X**

[8] Gary G. Briggs, Stephen Hawkins, & Herbert F. Crovitz. 1970. Bizarre images in artificial memory. *Psychonomic Science* 19, 6, 353–354. **10.3758/BF03328856**

[9] Donald Eric Broadbent. 2013. *Perception and communication.* Elsevier, Amsterdam, Netherlands.

[10] Xiang 'Anthony' Chen, Nicolai Marquardt, Anthony Tang, Sebastian Boring, & Saul Greenberg. 2012. Extending a mobile device's interaction space through body-centric interaction. *MobileHCI '12.* ACM, 151–160. **10.1145/2371574.2371599**

[11] Andy Cockburn, Per Ola Kristensson, Jason Alexander, & Shumin Zhai. 2007. Hard lessons: Effort-inducing interfaces benefit spatial learning. *CHI '07.* ACM, 1571–1580. **10.1145/1240624.1240863**

[12] Fergus I. M. Craik & Robert S. Lockhart. 1972. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11, 6, 671 – 684. **10.1016/S0022-5371(72)80001-X**

[13] Herbert F. Crovitz. 1971. The capacity of memory loci in artificial memory. *Psychonomic Science* 24, 4, 187–188. **10.3758/BF03335561**

[14] Raimund Dachselt & Anett Hübner. 2007. Three-dimensional menus: A survey and taxonomy. *Comput. Graph.* 31, 1, 53–65. **10.1016/j.cag.2006.09.006**

[15] Brian D. Ehret. 2002. Learning where to look: Location learning in graphical user interfaces. *CHI '02.* ACM, 211–218. **10.1145/503376.503414**

[16] Robert W. Frick. 1984. Using both an auditory and a visual short-term store to increase digit span. *Memory & Cognition* 12, 5, 507–514. **10.3758/BF03198313**

[17] Bruno Fruchard, Eric Lecolinet, & Olivier Chapuis. 2017. MarkPad: Augmenting touchpads for command selection. *CHI '17.* ACM, 5630–5642. **10.1145/3025453.3025486**

[18] Emilien Ghomi, Guillaume Faure, Stéphane Huot, Olivier Chapuis, & Michel Beaudouin-Lafon. 2012. Using rhythmic patterns as an input method. *CHI '12.* ACM, 1253–1262. **10.1145/2207676.2208579**

[19] Tiago Guerreiro, Ricardo Gamboa, & Joaquim Jorge. 2008. Mnemonical body shortcuts: Improving mobile interaction. *ECCE '08.* ACM, Article 11, 8 pages. **10.1145/1473018.1473033**

[20] Sean Gustafson, Daniel Bierwirth, & Patrick Baudisch. 2010. Imaginary interfaces: Spatial interaction with empty hands and without visual feedback. *UIST '10.* ACM, 3–12. **10.1145/1866029.1866033**

[21] Carl Gutwin, Andy Cockburn, Joey Scarr, Sylvain Malacria, & Scott C. Olson. 2014. Faster command selection on tablets with fasttap. *CHI '14.* ACM, 2617–2626. **10.1145/2556288.2557136**

[22] Chris Harrison, Shilpa Ramamurthy, & Scott E. Hudson. 2012. On-body interaction: Armed and dangerous. *TEI '12.* ACM, 69–76. **10.1145/2148131.2148148**

[23] Thorsten Karrer, Moritz Wittenhagen, Leonhard Lichtschlag, Florian Heller, & Jan Borchers. 2011. Pinstripe: Eyes-free continuous input on interactive clothing. *CHI '11.* ACM, 1313–1322. **10.1145/1978942.1979137**

[24] S. Kosslyn, M. Behrmann, & M. Jeannerod. 1995. The cognitive neuroscience of mental imagery. *Neuropsychologia* 33, 11, 1335 – 1344. **http://www.sciencedirect.com/science/article/pii/002839329500067D**

[25] Gordon Kurtenbach & William Buxton. 1993. The limits of expert performance using hierarchic marking menus. *INTERCHI '93.* IOS Press, 482–487. **http://dl.acm.org/citation.cfm?id=164632.164977**

[26] Nilli Lavie. 1995. Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human perception and performance* 21, 3, 451. **10.1037/0096-1523.21.3.451**

[27] Shu-Yang Lin, Chao-Huai Su, Kai-Yin Cheng, Rong-Hao Liang, Tzu-Hao Kuo, & Bing-Yu Chen. 2011. Pub - point upon body: Exploring eyes-free interaction and methods on an arm. *UIST '11.* ACM, 481–488. **10.1145/2047196.2047259**

[28] Eleanor A Maguire, Elizabeth R Valentine, John M Wilding, & Narinder Kapur. 2003. Routes to remembering: the brains behind superior memory. *Nature neuroscience* 6, 1, 90.

[29] Microsoft. 2017. Microsoft Kinect. Retrieved Sept. 2017 from **http://www.xbox.com/xbox-one/accessories/kinect**

[30] George A. Miller. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review* 63, 2, 81. **10.1037/h0043158**

[31] Seyed Yaghoub Mousavi, Renae Low, & John Sweller. 1995. Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of educational psychology* 87, 2, 319. **10.1037/0022-0663.87.2.319**

[32] Miguel A. Nacenta, Yemliha Kamber, Yizhou Qiang, & Per Ola Kristensson. 2013. Memorability of pre-designed and user-defined gesture sets. *CHI '13.* ACM, 1099–1108. **10.1145/2470654.2466142**

[33] Allan Paivio. 1971. *Imagery and verbal process.* Holt, Rinehart and Winston, New York, USA.

[34] Simon T. Perrault, Eric Lecolinet, Yoann Pascal Bourse, Shengdong Zhao, & Yves Guiard. 2015. Physical loci: Leveraging spatial, object and semantic memory for command selection. *CHI '15.* ACM, 299–308. **10.1145/2702123.2702126**

[35] Michael R. Raugh & Richard C. Atkinson. 1974. *A mnemonic method for the acquisition of a second-language vocabulary.* Inst. for Math. Studies in the Social Sciences – Stanford University, Stanford , CA, USA.

[36] George Robertson, Mary Czerwinski, Kevin Larson, Daniel C. Robbins, David Thiel, & Maarten van Dantzich. 1998. Data mountain: Using spatial memory for document management. *UIST '98.* ACM, 153–162. **10.1145/288392.288596**

[37] Anne Roudaut, Gilles Bailly, Eric Lecolinet, & Laurence Nigay. 2009. Leaf menus: Linear menus with stroke shortcuts for small handheld devices. *INTERACT '09.* Springer-Verlag, 616–619. **10.1007/978-3-642-03655-2_69**

[38] Joey Scarr, Andy Cockburn, Carl Gutwin, & Andrea Bunt. 2012. Improving command selection with commandmaps. *CHI '12.* ACM, 257–266. **10.1145/2207676.2207713**

[39] Joey Scarr, Andy Cockburn, Carl Gutwin, & Sylvain Malacria. 2013. Testing the robustness and performance of spatially consistent interfaces. *CHI '13.* ACM, 3139–3148. **10.1145/2470654.2466430**

[40] Joey Scarr, Andy Cockburn, Carl Gutwin, & Philip Quinn. 2011. Dips and ceilings: Understanding and supporting transitions to expertise in user interfaces. *CHI '11.* ACM, 2741–2750. **10.1145/1978942.1979348**

[41] Richard A. Schmidt, Timothy Donald Lee, et al. 2005. *Motor control and learning: A behavioral emphasis.* Vol. 4. Human kinetics, Champaign, IL, USA.

[42] Md. Sami Uddin, Carl Gutwin, & Andy Cockburn. 2017. The effects of artificial landmarks on learning and performance in spatial-memory interfaces. *CHI '17.* ACM, 3843–3855. **10.1145/3025453.3025497**

[43] Unity. 2017. Unity3D. Retrieved Sept. 2017 from **https://unity3d.com/**

[44] Valve. 2017. HTC Vive. Retrieved Sept. 2017 from **https://www.vive.com/**

[45] Dong-Bach Vo, Eric Lecolinet, & Yves Guiard. 2014. Belly gestures: Body centric gestures on the abdomen. *NordiCHI '14.* ACM, 687–696. **10.1145/2639189.2639210**

[46] Julie Wagner, Eric Lecolinet, & Ted Selker. 2014. Multi-finger chords for hand-held tablets: Recognizable and memorable. *CHI '14.* ACM, 2883–2892. **10.1145/2556288.2556958**

[47] Julie Wagner, Mathieu Nancel, Sean G. Gustafson, Stephane Huot, & Wendy E. Mackay. 2013. Body-centric design space for multi-surface interaction. *CHI '13.* ACM, 1299–1308. **10.1145/2470654.2466170**

[48] Martin Weigel, Tong Lu, Gilles Bailly, Antti Oulasvirta, Carmel Majidi, & Jürgen Steimle. 2015. iSkin: Flexible, stretchable and visually customizable on-body touch sensors for mobile computing. *CHI '15.* ACM, 2991–3000. **10.1145/2702123.2702391**

[49] Martin Weigel, Aditya Shekhar Nittala, Alex Olwal, & Jürgen Steimle. 2017. Skinmarks: Enabling interactions on body landmarks using conformal skin electronics. *CHI '17.* ACM, 3095–3105. **10.1145/3025453.3025704**

[50] Jacob O. Wobbrock, Htet Htet Aung, Brandon Rothrock, & Brad A. Myers. 2005. Maximizing the guessability of symbolic input. *CHI EA '05.* ACM, 1869–1872. **10.1145/1056808.1057043**

[51] Jacob O. Wobbrock, Brad A. Myers, & John A. Kembel. 2003. Edgewrite: A stylus-based text entry method designed for high accuracy and stability of motion. *UIST '03.* ACM, 61–70. **10.1145/964696.964703**

[52] Frances Yates. 1992. *The Art of Memory.* Random House UK, London.

[53] Shumin Zhai, Per Ola Kristensson, Caroline Appert, Tue Haste Anderson, & Xiang Cao. 2012. Foundational issues in touch-surface stroke gesture design – an integrative review. *Foundations and Trends in Human–Computer Interaction* 5, 2, 97–205. **10.1561/1100000012**

[54] Shengdong Zhao & Ravin Balakrishnan. 2004. Simple vs. compound mark hierarchical marking menus. *UIST '04.* ACM, 33–42. **10.1145/1029632.1029639**