Embodied Hybrid Bodystorming to Design an XR Suture Training Experience

Elena Márquez Segura
elena.marquez@uc3m.es
Department of Computer Science and Engineering, Universidad Carlos III de Madrid
Madrid, Spain

José Manuel Vega-Cebrián
josemanuel.vega@uc3m.es
Department of Computer Science and Engineering, Universidad Carlos III de Madrid
Madrid, Spain

Andrés A. Maldonado Morillo
andres.maldonado@salud.madrid.org
Department of Plastic Surgery, University Hospital Getafe
Getafe, Spain
Department of Medicine, Faculty of Biomedical Science and Health, Universidad Europea de Madrid
Madrid, Spain

Lara Cristóbal Velasco
lara.cristobal@salud.madrid.org
Department of Plastic Surgery, University Hospital Getafe
Getafe, Spain
Department of Medicine, Faculty of Biomedical Science and Health, Universidad Europea de Madrid
Madrid, Spain

Andrea Bellucci
abellucc@inf.uc3m.es
Department of Computer Science and Engineering, Universidad Carlos III de Madrid
Madrid, Spain

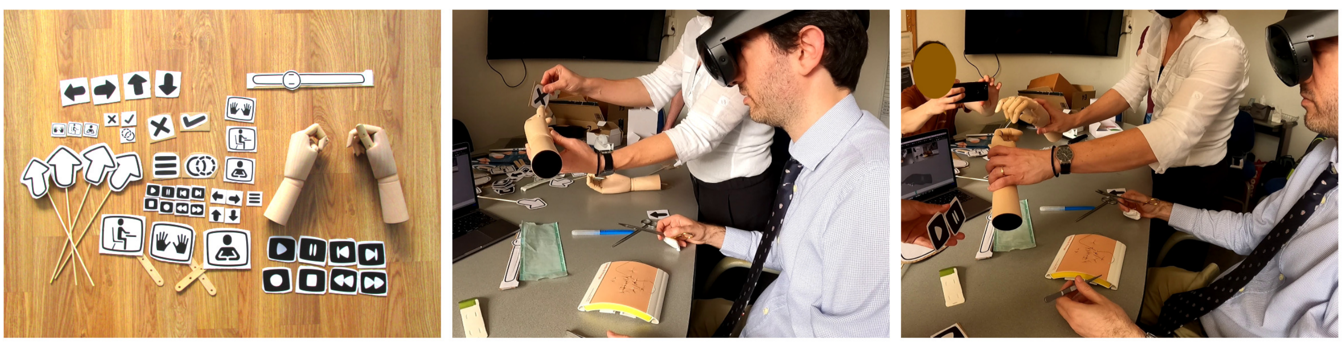

Figure 1: Left: Props used during the bodystorming session. Middle and Right: Photos of the embodied and hybrid bodystorming

ABSTRACT

Extended Reality (XR) technology offers promising results to support skill training. In the field of surgical education, Virtual Reality (VR) has long been explored, showing the potential to foster improved skill development and learning. However, XR in this domain is still underinvestigated, and there is a lack of design knowledge, design methods, and guidelines to inform how to best design XR experiences for effective surgical training. Here, we focus on suture training and show how participatory embodied design activities with experienced surgeons can help open the design space and arrive at interesting design solutions. We report on a hybrid bodystorming combining physical props with XR headsets with passthrough capabilities, supporting rich embodied explorations, a better understanding and articulation of key steps of suturing, uncovering essential design requirements and features, and arriving at an interesting design concept proposal that can be inspiring for future works in the domain.

TEI '24, February 11–14, 2024, Cork, Ireland

ACM ISBN 979-8-4007-0402-4/24/02
https://doi.org/10.1145/3623509.3633362

CCS CONCEPTS

• **Human-centered computing** → **Mixed / augmented reality**; **Interaction design process and methods**.

KEYWORDS

Embodied Design Methods, Bodystorming, Embodied Interaction, Explorative Design, Suture Training, Extended Reality, Mixed Reality, Participatory Design.

ACM Reference Format:

Elena Márquez Segura, José Manuel Vega-Cebrián, Andrés A. Maldonado Morillo, Lara Cristóbal Velasco, and Andrea Bellucci. 2024. Embodied Hybrid Bodystorming to Design an XR Suture Training Experience. In *Eighteenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '24), February 11–14, 2024, Cork, Ireland.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3623509.3633362

1 INTRODUCTION

The emergence of eXtended Reality (XR) technology that spatially integrates virtual cues aligned to the haptic feedback from physical objects has opened up new possibilities in the field of surgical education. Recent findings show potential for XR experiences to foster improved skill development and afford more consistent learning outcomes compared to conventional teaching methods for fundamental surgical skills such as suturing [10, 33]. While the literature on Virtual Reality (VR) simulations with haptic feedback devices is abundant [1, 2], the use of XR technology for surgical training is underexplored, possibly due to the very recent availability of more advanced, ergonomic and affordable headsets and software development kits that make it easier to explore the design space of mixed reality interactions. Therefore, there is a lack of design knowledge, in the form of empirical research, design methods and guidelines to inform how best to craft XR experiences for effective surgical training, and understanding what augmentations of physical elements would contribute to facilitate learning.

Motivated by the potential of XR technology to support motor and cognitive skill acquisition, we engaged in a series of embodied co-design sessions with two experienced surgeons and professors of medicine to design an immersive training tool to help medical students better understand the process of stitching during sutures, including essential hand-object manipulation aspects. We highlight sensory and somatic considerations and reflect on the potential of our hybrid methodology and technological setup, combining physical props with XR headsets with passthrough capabilities, to support exploration, uncovering essential design features that would not have been possible to capture otherwise. We contribute knowledge about the impact of important elements and hybrid co-design strategies with regard to future immersive technology for training.

2 RELATED WORKS AND BACKGROUND

This work draws upon previous literature on experiential learning and the use of educational technology in surgical training, immersive systems and tools to support surgical activities as well as embodied design methods.

2.1 Experiential, Hands-on Learning in Surgical Training

Basic suture training is a fundamental skill for medical doctors, which requires hours of hands-on practice to master different aspects of the stitching process, such as the correct sequence of actions as well as accurate hand movements and manipulation of the instrumentation. Acquiring surgical skills is an embodied activity [29]: Medical students learn not only from observations but from engaging in active practice under expert guidance [7]. Hands-on activities, sensory experiences and simulations play a crucial role in acquiring surgical skills such as motor control and precision [30] as well as procedural knowledge [7]. Additionally, the learning experience involves a sensory semiosis in which, through repetitive practice, the students acquire the ability to interpret visual and haptic cues [8], which are fundamental indications to direct trainees' attention and support motor learning [5]. The hands-on teaching model in medical education, however, is often challenged by several constraints, such as time limitations for in-class practice, limited access to training resources, lack of consistent and objective feedback [43], the difficulty in providing in-class personalized guidance [22] as well as the integration of novel technology in the teaching practice [9].

The use of videos as educational media for surgical training has shown a potential to enhance guided physical training by providing the underlying logic behind expert actions, which is often difficult to explicitly convey in guided sessions [9]. The linear and static nature of videos, however, hampers their full integration as effective supplementary tools. Interactive XR technology, meaningfully integrated into the learning environment, can help tackle the aforementioned issues, providing contextual and spatially-aligned visualizations for interactive guidance [9] and immersive practice, augmented visual and haptic cues, as well as personalized and adaptive experiences that could better enable students to acquire proficiency in suturing procedural steps and technique in a safe and controlled environment [33].

2.2 Extended Reality for Surgical Training

Virtual Reality simulations with haptic force feedback for training and evaluating suturing technique has long been the subject of research [25]. The literature on the effectiveness of immersive VR, compared to other forms of training such as traditional desktop-based 3D images or physical training in different surgical training contexts is not conclusive [1, 31]. However, immersive technology providing contextual visualizations as well as spatial interactivity is more effective than videos [2] and showed potential to increase motivation for training as well as strengthen confidence for task execution [1], enhance context-aware learning [27] and spatial understanding (e.g., the distance between critical structures) [32]. Moreover, immersive systems can be used as risk-free learning environments to foster hand-eye coordination skill development, and, at the same time, they can serve as an environment for objective assessment of trainee performance [19]. Proposed VR systems focus on technical aspects, such as the fidelity of surgical tools simulation using high-end force feedback devices, physics-based 3D simulations of tissues and the development of evaluation metrics to assess the trainee's performance during the training sessions, such as needle placement accuracy, surface tissue damage or time to complete the task [25]. One exception is the work of Chellali et al. [6] that reports methodological implications of their design process for a laparoscopic VR simulator, reflecting on the importance of engaging expert surgeons in hands-on activities with the technology as part of the design process to uncover richer design considerations. Most of the research on VR training has been conducted in the context of laparoscopic suturing [2] and the kinaesthetic devices used do not allow to generalise and adapt the implementations to other training contexts such as basic suture training, in which more fine-grained manipulations of different instrumentation are needed.

There are a few examples of immersive systems for basic suture training. Lia et al. [15] used the HoloLens see-through device to display mixed-reality instructional videos. Tashiro et al. [39] developed a gamified XR system for microsuture training employing a microscopic camera and a Head-Mounted Display (HDM) to provide real-time augmented feedback to the trainee. Rojo et al. [33] augmented the standard suture training material available in medical classrooms with contextual visual cues using the see-through MagicLeap device. The design drivers and rationale of those systems, however, are not described, and there is a lack of design knowledge on how to integrate spatially aligned instructional material (e.g., visual cues) with the haptic feedback of the physical tools, taking into account cognitive theories of multimedia learning [26].

2.3 Embodied Design Methods and Immersive Technologies

Embodied design methods foreground the bodily, sensorial, and first-person experience of the stakeholders to elicit early insights into the embodied experience of participants of a design [20, 21, 24, 34, 40, 42]. They can be adopted in the different stages of a design process, from sensitizing to evaluation. For instance, Bodystorming refers to a generative design method that, in contrast to brainstorming, uses full-body engagement with objects, the space and other people to kindle ideas. It has been developed in a variety of directions. For example, Oulasvirta et al. [24] proposed bodystorming as the practice of implementing ideation sessions in the physical context where designs will be used. Schleicher et al. [34] presented bodystorming as three different approaches: prototyping using enactment; physically emulating the spatial environment in which technology will be used to generate and evaluate ideas in context; and employing actors and props to play out expected use case scenarios. Later on, Márquez Segura et al. [20] appropriated the method to playfully design considering, and using as design resources the physical, collocated, and social aspects of an activity. This method is eminently activity-centred, focusing ideation around key actions of a future situation involving technology, i.e. "embodied core mechanics" [20].

Relatedly, Embodied Sketching [21] encompasses ideation methods that prioritize the lived experience of participants-designers and include the social and spatial context of a scenario as design resources, along with digital and non-digital artefacts that aid in idea generation. Embodied sketching is also based on harnessing play and playfulness as the principal way to instigate creative physical engagement. On a similar note, Embodied Design Ideation practices [42] focus on the researchers' and participants' first-person experiences throughout the design process by relying on *estrangement* in the dimensions of body, material and context.

Embodied design methods are recently being explored in the context of immersive experiences in the hands of recent advances in immersive technologies, including VR, augmented reality (AR), and XR. These technologies often facilitate perspectives, transitions, and even the blending of physical and digital elements and realities that support understanding of present and future situations and technologies, fuelling creativity and innovation. VR has been used in the context of product engineering and design, to design cars [47]; spatially-aware interactive spaces [13] and large scale shape-changing interfaces [3]. Often these works present interesting novel VR facilitated embodied design methods. For example, in Simeone et al.'s Immersive Speculative Enactments [36], users engage in virtual bodystorming in Virtual Reality to reflect on possibilities for design in the tangible world. Relatedly, in McVeigh-Schultz et al.'s Immersive Design Fiction [18], designers are immersed in a virtual narrative world acting as characters (future designers) engaging with future work practices by their design company. In the context of Mixed Reality (MR), Weijdom's Performative Prototyping [41] proposed bodystorming methods, puppeteering (a Wizard of Oz technique in MR), and performative arts techniques and practices, which are used in collaborative mixed-reality environments and to design performative MR experiences in higher-art education. Then, works as that of Zhou et al. [44] or Lund et al. [16] leveraged the use of embodied design methods to explore the affordances of collocated play in VR. In Zhou et al. [44], designers use physical props and subvert VR equipment (i.e. controllers) to imagine a social dance game in MR. Similarly to our work, these props are brought in the VR realm through using the passthrough capabilities of the VR headset.

3 METHODOLOGY

In this work, we follow a Research through Design (RtD) approach [45, 46], which values the design process as much as the resulting design, the 'ultimate particular' [38]. This approach considers knowledge and outcomes obtained *as part of* the process of *designing* invaluable to shape the design, but also and notably relatedly, to better understand the design problem at hand. Typically in RtD processes, both problems and solutions emerge and evolve hand in hand. Further, our RtD process is highly participatory, involving expert surgeons ("expert" from here onwards) as co-designers. We also employ embodied design methods [20, 21, 40, 42] to sensitize [21] designers and surgeons and to ideate and test potential designs. This work represents a first step to include XR technology in the design of augmented training tools, in particular, to improve current practices of suture teaching. In our design process, we took a strong teacher-centred perspective, aiming at surfacing key embodied design knowledge from expert surgeons and teachers conducting the suturing practice, and explaining suturing to novice practitioners. Our co-design process only included two expert surgeons due to: i) the early stage of the design process, targeting a first design concept; ii) the teacher-centred approach taken, focusing on improving *current* teaching practices with XR; and iii) the type of research methods used, eminently qualitatively and focused on micro video-analysis (See section 3.2). Still, while we have currently only included two expert surgeons as co-designers, our participatory design method plans to include learners (i.e., medical students) in later stages, such as iterating and testing our design concepts and prototypes, which will likely help us identify potential learning obstacles, and polish our design.

In the following, we briefly explain the design process we followed (See Summary in Table 1 in the Supplementary Materials), to then focus in particular on one embodied co-design activity (Session 4), which was particularly useful to i) further understand, refine, and materialize the problems targetted by our design as well as design requirements in our application domain; and ii) come up with potential design solutions deemed useful by our expert.

We relate the fruitfulness of this session to the embodied and hybrid nature—mixing VR and physical elements, blurring the barriers between VR and the physical world—of the co-design session, and focus this paper on the method employed, and the emerging and resulting design and application domain insights.

3.1 The Design Process

The design process involved four distinctive phases including 3 debriefing sessions with an expert surgeon (the expert; and third author) running the Department of Plastic Surgery in the University Hospital Getafe, Spain, who also teaches future surgeons surgical techniques, such as those involved in skin sutures (subcutaneous sutures, intradermal sutures, etc.) during theoretical and practical classes. In the latter, students would typically use a suture kit, which includes a piece simulating the skin and fat (See object in between the surgeon's hands in the middle and right pictures in Fig. 1), basic surgical instruments (e.g. scalpel, needle holder, forceps, etc), and the sutures (nylon or vicryl). During the debriefing sessions, the design team developed a basic understanding of overarching problems the surgeons in training typically face, such as the little guided practice time they have, and the challenges of personalized assistance and dedicated teacher's time during practice.

Together, we set to design a mixed reality experience featuring virtual expert "ghost hands,"[1] which would guide students during their suture training with the physical suture kit. This virtual visualization would be designed to guide, model, and correct the students in their practice. Next, we engaged in two embodied design sessions. The first was mainly planned to record data from two expert surgeons (third and fourth authors) performing the different suturing phases, as well as common mistakes a student would typically make. For this, we used the Meta Quest 2—with integrated hand-tracking capabilities and passthrough—which allowed the surgeons to see through the headset and interact with the suturing kit. However, the passthrough displayed a low-resolution grayscale of the surrounding world, which hindered the precision and dexterity of the surgeons. Still, it allowed us to capture data on the gestures performed by the surgeon during the suturing process, as detected by the XR device from an egocentric perspective. We collected hand data using the Unity Hand Tracking API[2] provided by the Meta Quest Unity SDK. For each frame, we collected the 3D position and orientation of each joint of the hand skeleton (25 in total, refer to the documentation for further details on the joints[3]). This data, together with video recording data from two GoPro cameras positioned in front and at about 45 degrees to one side of the expert allowed us to build our first preliminary prototypes, such as a 3D model of the ghost hands guiding the suturing process that could be watched in VR, and as part of a video prototype (See https://shorturl.at/rUW08 or a screenshot in Figure 1 in Supplementary Material).

Further, this first embodied design activity served as a sensitizing one [21] for the experts, to familiarize themselves with the capabilities of the technology; and for us designers, to better understand and flesh out the problems and mistakes students make through the expert's enactment.

Then, we engaged in another embodied design activity—an embodied hybrid bodystorming session, which is the design activity at focus in this paper and is covered in the next subsection. All the design activities are described in the Supplementary Materials, including a brief description and summary of the main outcomes. See a summary of the design process in Table 1 in the Supplementary Material.

3.2 An Embodied Hybrid Bodystorming

This session was planned as a co-design activity to jointly think about potential design concepts to correct and guide students in their practice. From the onset, we planned the session to work in the hybrid physical-virtual space, using the Meta Quest Pro VR headset with passthrough capabilities, which can display virtual content and at the same time allow the user to see a colour and higher resolution 3D view of the physical world than Meta Quest 2. This allowed us to include in the "virtual world" physical elements brought to the session. In particular, it allowed the surgeon to better interact—with more precision, and dexterity—with the suturing kit. Further, it allowed us to include physical props from our "bodystorming basket" (see Figure 1) representing the ghost hands, and physical visualizations prompts to envision and explore potential alternative corrections and guidance for the ultimate design. Further, we also brought our two early prototypes: one in VR, featuring expert ghost hands guiding sutures; and a video prototype (See a screenshot in Figure 1 in Supplementary Material or a demo video in https://shorturl.at/rUW08). This, together with the use of the new Meta Quest Pro was meant to sensitize the expert to the possibilities of the technology.

The bodystorming session began with the video prototype, which triggered overarching discussions about what to see, how, and where; as well as how to control and interact with the visualizations. Later, we went through - and refined - the different suturing phases. The expert surgeon enacted them using the Meta Quest Pro and the suturing kit, walking the designers through and thinking aloud [23] during each phase. For each step of each phase, the expert also enacted the different errors. Each of them in turn, triggered bodystorming using the physical props, which centred on how to signal the mistake and guide the future surgeons' practice.

The enactment of the different suturing phases, errors, and proposed were captured with two GoPro cameras, one recording slightly behind and over the surgeon's shoulder; and another about 90 degrees from the first camera, capturing suturing from one side. Mobile phone devices were also used to take snapshots of key moments and additional video recordings from alternate planes. This material was later reviewed through a qualitative video analysis [11] by the first author, who has extensive experience with video

[1]This is a commonly used game feature in which there is a non-collidable version of the game currently played rendered as an overlay, typically a replay. It is called "ghost" because the game version rendered is usually semi-transparent or in grey-scales, to indicate a footprint condition rather than an interactable game element.
[2]https://developer.oculus.com/documentation/unity/unity-handtracking/
[3]https://developer.oculus.com/reference/unity/v57/class_o_v_r_skeleton/

analysis and transcribing movement at a micro and macro scale, using knowledge in movement analysis and basic biomechanics. The analysis involved several passes through the video material. A first coarse pass focused on 'what' happened, identifying phases, steps within each phase, and typical errors. Subsequent passes through the video material were conducted to do a more detailed micro-analysis focused also on 'how' each suturing step was conducted, which allowed the transcription of detailed procedures, errors, and visualizations. These transcriptions were later reviewed by the third and fourth authors (expert surgeons) for procedural and conceptual validation.

4 OUTCOMES

The hybrid bodystorming yielded varied outcomes. First, it helped us designers understand, and the surgeons reflect, and better explain through enactment the different phases as well as the typical mistakes the future surgeons make. While we already had a preliminary list of phases, key procedures, and mistakes from earlier sessions, this list was completed and substantially enriched in this session. This allowed us designers to better document and update earlier accounts of the suturing phases and mistakes. A posterior video analysis [11] articulated these phases (Covered in subsection 4.1, and transcribed in Tables 3 to 8 in Supplementary Material).

The hybrid bodystorming also allowed all of us to explore different design possibilities, settle on particular design requirements and features, and arrive at concrete ways to signal errors and visualize guidance for each of them (Covered in subsection 4.2 and Tables 3 to 8 in Supplementary Material). A posterior video analysis of the design proposals further classified the resulting designs depending on what they visualized, where, when, and how, which resulted in interesting patterns that can provide interesting directions for future designs. In the following subsections, we describe the main outcomes of our hybrid methodology, first covering the phases of the suturing process as emerging from the surgeon's description during the bodystorming sessions and then reporting on the proposed design solutions with illustrative examples, as well as reflecting on emerging patterns. But first, we recommend the reader to go through the Supplementary Material to have a better picture of the phases, steps, and sequences of actions involved in the suturing process; and be able to better decode rather 'complex' action from a non-expert perspective.

4.1 The Steps of the Suturing Process

The suturing process involved six phases: (1) Preparing the materials; (2) First stitch; (3) Knot; (4) Cutting the knot; (5) Repetition of phase 1: preparing the materials; and (6) Second stitch: Repetition of phase 2 with slight variations. In this section, we summarize the procedures and classical errors involved in each suturing phase (P from here onwards). An extended description of each phase and error can be found in the Supplementary Materials, in the form of tables. Here, we include an abridge of these tables compiling a few procedures and errors, selected because they represent the main phases and cover procedures and errors of every kind (See Table 1). Next, we describe the different phases:

FIRST PHASE (P1). The first phase, "Preparing the materials," is a preparatory phase happening at the beginning of the suturing process, and every time a stitch is completed. Hence, after phase 4, "Cutting the knot," phase 5 repeats this preparatory phase. This phase only included procedures related to preparing and handling surgery equipment, such as the needle holder, the forceps, and the needle (i.e. ID1, ID2, and ID3 respectively in Table 3 in Supplementary Material). Some of these procedures are particularly tricky for how they stipulate a different holding of a tool that resembles common daily ones. E.g. the forceps resemble tweezers, except that the former requires handling the joint tip of the tool outside the hand (crease between thumb and index finger), instead of inside the hand like the tweezers. Similarly, the needle holder resembles scissors, except the fingers that need to be inserted in the finger rings are different. These errors, which are very common among students (i.e. surgeons in training), and relatively hard for them to pick up, impact the way the students can later handle and manoeuvre, which might lead to additional difficulties and dexterity problems in the suturing procedure.

SECOND PHASE (P2). The procedures in this phase, "First stitch," mostly require paying attention to relative positions of key elements such as insertion points, stitch, wound, and thread. P2 focuses on issues of length, distance and position of stitches with relation to several aspects, such as needle insertion points (e.g. IDs 7-9 in Table 4 in Supplementary Material), and the wound (e.g. IDs 4, 8-9 in Table 4 in Supplementary Material); as well as issues of length and distance of the thread (e.g. IDs 5-6 in Table 4 in Supplementary Material).

PHASE 3 (P3). The procedures within phase 3, "Knot," starts with an equipment handling ("EQUI") task, ID10 in Table 5 in Supplementary Material, and the relatively complex sequence ("SEQ") described in ID11 about tying the first knot in a "straight way," which involves the sub-procedures specified in IDs 12-14 (Table 5 in Supplementary Material). This is a particularly challenging step in the suturing process leading to frequent mistakes (also specified in IDs 12-14). The complexity lies in the fact that a "straight knot" can be visually distinguished by the trained eye, but not by novices. Similarly, unlike novices, trained surgeons have interiorised the key movement specified in ID14 (Table 5 in Supplementary Material), involving alternating the positions of the hands, which makes for a "straight knot:" If the right hand (with the needle holder) picks the loose end of the thread below the wound, this hand (and that end of the thread) should finish above the wound; the opposite hand, in turn, makes the opposite movement: starting from above and moving below the wound. The complexity of this sequence required a special kind of visualisation, explained in detail in the following subsection. Then, the mistakes in procedures 15-17 (Table 5 in Supplementary Material) relate to the count, type, and kind of stitches ("C&T/M"), e.g. the number of knots (ID15), whether knots should be double or simple (ID16), and how to do the consecutive knots (ID17). The last procedure in this third phase (ID18, in Table 5 in Supplementary Material) refers to the duration of this whole phase ("TIME"), which should take about 20-30 seconds. In suture training, practice takes time; and practice shortens the time it takes surgeons to tie the knot.

PHASE 4 (P4), 5 (P5), and 6 (P6). The procedures in phase 4, "Cutting the knot," are two, one related to the handling of equipment

Table 1: Examples of steps and errors in Phases 1, 2 and 3. The first column in the table, "ID," refers to the identification number (ID) of a particular procedure. The second column, "Ph," refers to the phase of the stitching process. The third, "Code," refers to a code assigned to the procedure among the following: "EQUI" refers to the handling of equipment or materials; "L&D&P" refers to length, distance, and position (of e.g. stitches, thread, from different points, e.g. the wound); "C&T/M" refers to count, type, and mode (of stitches); "TIME" refers to the approximate time that should be spent; and "SEQ" refers to the sequence of actions within a particular action. These codes emerged bottom-up in a posterior classification of procedures and errors in the video analysis. The fourth column, "Name," names the particular procedure at focus. The fifth column "Procedure," describes the procedure, emphasizing key aspects that might present particular difficulties. The sixth column, "Errors," describes the typical errors made by future surgeons.

ID	Ph	Code	Name	Procedure	Errors
1	1	EQUI	Handling the needle holder	The needle holder should be handled using the first (thumb) and fourth (ring) fingers. (Figure 2, Right)	Use of first and second or third fingers instead.
2	1	EQUI	Handling the forceps	The first (thumb), second (index) and third (middle) fingers are needed to use the forceps. The joint tip of the forceps should stay outside the hand, leaning in the crease between those fingers. (Figure 2, Left)	The joint tip of the forceps stays inside the hand.
3	1	EQUI	Handling the needle	The needle should be picked up and manipulated with the needle holder.	The needle is picked or manipulated with the hand
5	2	L&D&P	Thread length	After the needle goes through both sides of the wound, the end of the remaining thread should not be much longer than 3cm. (Figure 3, Left)	After the needle goes through both sides of the wound, the end of the remaining thread is longer than 3cm.
7	2	L&D&P	Stitches alignment	The two points where the needle is inserted should be aligned. (Figure 3, Right)	Misalignment of the two points where the needle is inserted
11	3	SEQ	Tying the first knot	The hand holding the forceps should take the thread (at about 10cm); the needle holder on the other hand should go on top (or below) of the thread and turn twice (double knot) around the thread, creating a double loop. Then, the loose end of the thread should be picked up and pulled through the loop, drawing it tight and down to the wound.	Several errors, specified below.
15	3	C&T/M	Knots number	Three knots need to be made.	More or less than three knots are made.
18	3	TIME	Time	The knot should be done within approximately 20-30 seconds time.	Taking much longer than 30 sec to tie the knot.

("EQUI"; ID19 in Table 6 in Supplementary Material), i.e. the forceps and the needle holder, which should be picked with one hand while the other cuts the threads; and another one related to the length of the thread ("L&D&P") that should remain after cutting the knot (ID20 in Table 6 in Supplementary Material). After cutting the knot comes phase 5, which involves preparing the materials like in phase 1 (with the same procedures and errors). Then, phase 6 follows, when giving the second stitch. This involves the same procedures and errors as phase 2, adding one more procedure (ID23), which relates to the distance ("L&D&P") between the new stitch and the previous one.

A posterior analysis of articulation of the procedures and errors, focused on classifying them onto different kinds allowed us to characterize the suturing process with 32 procedures:

- 9 related to equipment handling ("EQUI"): 2 in phase 1 (P1); 1 in p2; 2 in P3; 1 in P4; 2 in P5; and 1 in P6.
- 15 related to length, distance and position ("L&D&P"): 6 in phase 2 (P2); 1 in P3; 1 in P4; and 7 in P6.
- 3 related to count, type and kind of stitches ("C&T/M"), all in P3.
- 1 related to "TIME" in P3.
- 4 involving sequences of actions: 1 in P2; 2 in P3; and 1 in P6; one of which (in P2) also related to "L&D&P."

Interestingly, only four procedures were sequences of actions, two of which involved pulling the thread in a particular way (ID6 in P2 and 6), which is a relatively simple task requiring paying attention to when it is conducted (after the needle goes through the two insertion points; see Table 4 in Supplementary Material). The remaining two refer to the one complex sequence action in suturing, which involves tying the first knot (ID11), and in particular, lowering this knot in a "straight way" (ID14), which is an aspect of ID11. Most of the procedures involve paying attention to issues related to length, distance, and position ("L&D&P," 15), and equipment handling ("EQUI," 9). Still, doing all of these procedures right is not straightforward, as stated by our expert surgeons, and they take quite some practice and time to get right.

4.2 Design Features and Design Concept Proposal

The hybrid bodystorming session yielded design proposals for each of the procedures and errors future surgeons could make. A posterior analysis of the resulting proposals revealed patterns regarding what to display, where, how, and when. In this section, we first describe and illustrate some of the resulting design proposals. For the sake of space, we describe summaries of the resulting design proposals per phase and illustrate only certain proposals, which we chose because they show at least one per phase and some of those characterizing patterns. All the descriptions of the resulting designs can be found in the Supplementary Material. Then, we reflect on the outcomes of our posterior analysis of the resulting designs, deriving relevant design features and requirements for future designs in the domain of XR experiences for suture training.

4.2.1 Summary of Design Proposals.

PHASE 1 (P1). All the procedures in P1 involved handling equipment and the emerging design concepts were similar. They visualised an error warning on the student's hands (red X mark and red arrow pointing to key student's hand area at focus) and displayed the correction/guidance on the ghosts' hands, signalled with an outward arrow originating at the student's hand and ending on the ghost's hands. Depending on the error, the guidance appeared at the same time (ID3), or right afterwards (ID1, ID2—See Figure 2). Also, some of the emerging guidance/corrections on the ghost's hands displayed both the error and the correction (ID) or just the correction (ID3).

PHASE 2 (P2). All the procedures in P2 relate to length, distance, and position ("L&D&P") of stitches, producing also similar visualizations. Since the stitches on the kit are the reference, the kit is used to signal error warnings and to explain the error. Right after, a ghost kit appears guiding the student on the correct performance. In all cases (ID4-ID9), a dimension line appears, in red and on the kit signalling problematic distances or positions; and in green and on the ghost's kit to signal correct ones. In most cases (ID4, ID5, ID7, ID8, ID9), a dot marks particular points (e.g. needle insertion point (ID4, ID7, ID8, ID9), and wound (ID4, ID8, ID9)). Dotted lines also appear to highlight alignment (ID4, ID5, ID7, ID8, ID9). Both dots and lines appear in red or green to signalling incorrect/or correct points respectively. In a few cases (ID4, ID5, ID9), text or a symbol (=, or ≠ in ID8) is also displayed to specify measures, e.g. "<3mm" in ID5. As in P1, here a red X mark and red arrow signal incorrect procedure; while a green check symbol signals correct procedure. In a case (ID6), an element is highlighted (the thread) in red or green to signal incorrect or correct procedure respectively.

PHASE 3 (P3). The two procedures (ID10, ID12) in P3 handling equipment ("EQUI") refer to dropping equipment on the work area and are displayed similarly: signalling an error warning sign on the student's kit with a red X mark and a red arrow pointing to the interest area; and displaying at the same time, the ghost hand properly handling the equipment, highlighted in green. Then there's a procedure related to the distance of the thread ("L&D&P"; ID13) that results in a similar visualisation of those in a similar category in P2, except this time the student's or the ghost's kit & hand are displayed. This is because the distance in ID13 refers to the distance from the wound (kit) to the point of the tread picked up by the surgeon's hand, hence both, hand and kit are necessary to guide performance.

Then, this phase contains two complex sequences ("SEQ") of actions related to tying the first knot (ID11, ID14), which involves different visualizations depending on the specific error made. The most interesting one is ID14, referring to the procedure of "lowering the knot straight." This triggers an error warning on the student's kit, and a visualisation on the ghost hands & kit of three alternating sequences of actions: first, the correct procedure showing three key moments of the sequence of actions to lower the knot straight; then three key moments of the sequence of actions to that result in lowering the knot twisted; and again the sequence of the correct procedure. The former and latter are signalled with a green arrow and green checkmark pointing at the interest area; the second is signalled with a red cross and arrow pointing to the interest area. This is the only emerging design proposal that shows an animation of sorts, and the repetition of the correct procedure twice. Last, procedures ID15-17 relate to the count, and type of stitches ("C&T/M"), and all showcase the error warning (red cross and arrow) and visualisation of the error on the student's kit; and signal (green checkmark and arrow) and showcase the right procedure on the ghost's kit. All include some text clarifying the type or count referred to (e.g. "double knot," ID16).

PHASE 4 (P4). This phase contains two procedures, one related to handling equipment while cutting the knot ("EQUI," ID19), and one related to the length of the remaining thread of the knot ("L&D&P"; ID20). The former refers to not dropping the equipment (forceps and needle holder). It visualizes an error warning on the student's kit (red X mark and arrow pointing to the instruments dropped) and, at the same time, a green checkmark and arrow pointing to the ghost's hands handling the instruments, highlighted in green. The latter displays an error warning (red cross and arrow) to the remaining thread, which is either shorter or larger than needed and is highlighted in the student's kit; and signalling (green checkmark and arrow) and guidance on the ghost kit, highlighting the remaining thread. Coloured dimension lines and texts reading the lengths (>,<, or 0,5cm can be read in each of the kits).

PHASE 5 (P5) and PHASE 6 (P6). These phases repeat phases 1 and 2 respectively. P6 adds a procedure (ID23) related to the distance between consecutive stitches, ("L&D&P"), which is visualized on

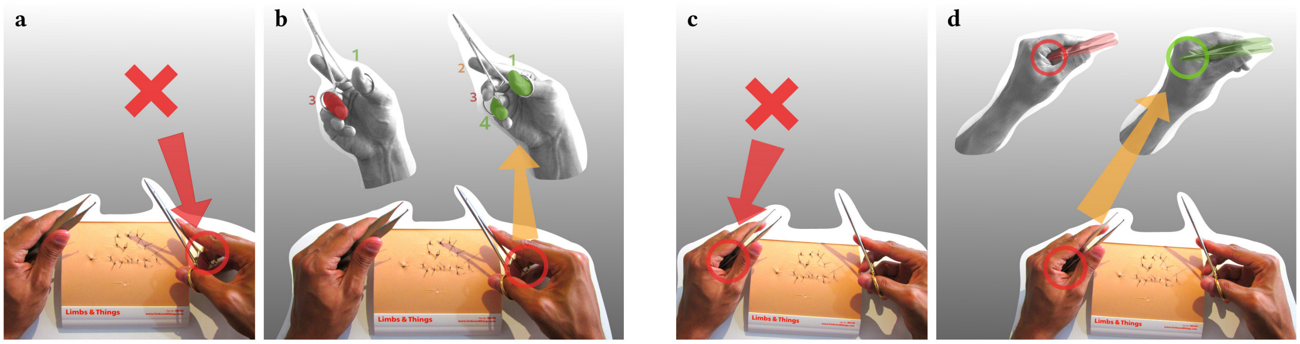

Figure 2: Left (a, b), Correction sequence for wrong holding of the needle holder (ID1). Right (c, d), Correction sequence for the wrong holding of the forceps (ID2).

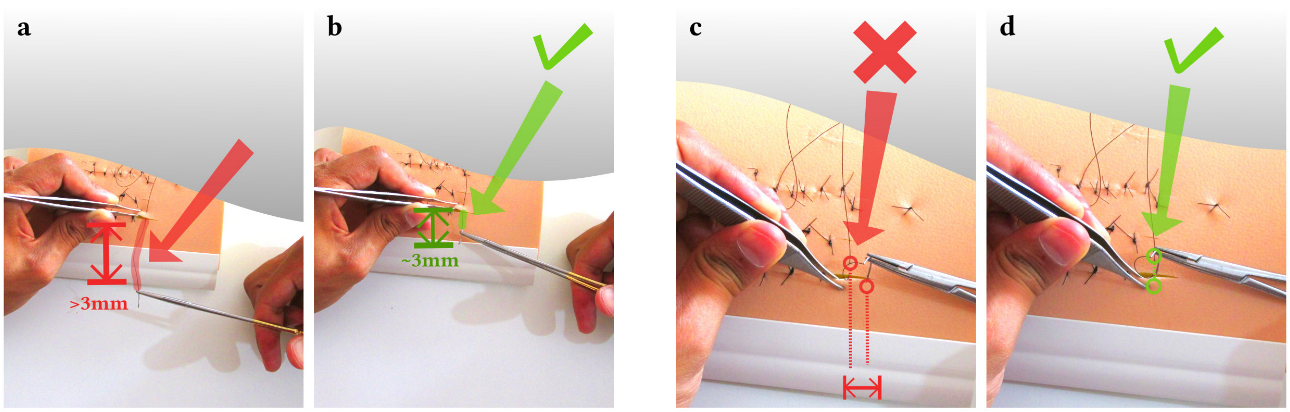

Figure 3: Left (a, b), Correction sequence for an inappropriate length of thread after needle insertion (ID5), Right (c, d), Correction sequence for misalignment in the placement of the needle (ID7)

the student's and ghost's kit in a similar way than "L&D&P" errors earlier related to distance of stitches (ID4, and ID8): on the student's kit and the ghost's kit, with a red error warning and arrow on the former; and a green checkmark and arrow on the latter, and with coloured dots marking insertion points, coloured dotted line joining the consecutive insertion point dots, coloured dimension lines, and text reading the separation in mm o the stitches. The colours used, are red for errors, and green for guidance.

4.2.2 Insights and Emerging Patterns. Here, we reflect on the resulting visualization proposals, under emerging salient categories. Within them, we derive recommendations for future designs resulting from patterns in the proposals.

WHAT. Here, we briefly describe key design elements involved in the visualization. **Colours** to convey information: All instances made use of two primary colours, red and green, to signal incorrect and correct performance respectively. The choice of colours is motivated culturally. In many western countries, the red colour indicates "stop," "danger," or "caution," (see e.g. traffic conventions); in contrast the green indicates "safe" and "right way." However, we are aware that this is not valid worldwide and that it brings accessibility issues for many with colour blindness. Hence all instances also displayed **symbols** in the form of a red cross to indicate incorrect, and a green *checkmark* to indicate correct performance. Then, many instances (14) **highlighted key areas of interest** by circling an area (ID1-ID3, ID13; three "EQUI" and one "L&D&P" procedure), and by colouring (red/green again) an area of interest, which was: finger (ID1, an "EQUI" procedure), forceps (ID2, ID3, ID10; all "EQUI" procedures), the thread (ID5, ID6, ID13, ID20; all "L&D&P" procedures), the needle (ID12, an "EQUI" procedure), the knot (ID16, ID17; both "C&T/M" procedures), and the forceps and needleholder (ID19, an "EQUI" procedure). While the colour might bring the issues above mentioned, they could also be displayed by heightening the shade of the object highlighted, e.g. a finger (ID1). Then, six instances (ID4, ID5, ID7, ID8, ID9, ID23) used coloured (red/green) **dots and dotted lines**. These were all "L&D&P" procedures.

Symbols were prominently used (14 instances): in the form of dimension lines (ID4, ID5, ID7, ID8, ID9, ID13, ID20, ID23; all "L&D&P" procedures); mathematical expressions like '>', '<' (ID4, ID5, ID9, ID18, ID20, ID23; all except one "L&D&P" procedures, and one "TIME" procedure), and equal kinds of symbols (= ≠ ~(ID4, ID5, ID8, ID9, ID20, ID23; all "L&D&P" procedures); knots (ID14, ID15, ID16; two "C&T/M" and one "SEQ" procedures); loops (ID17, a

"C&T/M" procedure); and a clock symbol (ID18, a "TIME" procedure). Finally, **text** was also used together with symbols to further specify measurements (ID4, ID5, ID9, ID20, ID23; all "L&D&P" procedures); frequency (ID15, ID17; both "C&T/M" procedures); kinds type (e.g. simple/double knot, ID16, a "C&T/M" procedure); and time (ID18, a "TIME" procedure).

Derived design implications. Make clear what is the wrong and right performance. Consider using colours with traditional cultural significance, but bear in mind accessibility issues (e.g. people with colourblindness) as well as cultural differences in colour meaning. Including a symbol and/or brief text label (e.g. 'YES' 'NO') can help address these issues. Length, distance and position ("L&D&P") procedures may benefit from dimension lines signalling distances; mathematical expressions, and text specifying measurements. Then, count, type, and mode instances ("C&T/M") may benefit from symbols specifying the type of element (e.g. twisted vs. straight knot, ID14) with accompanying text, and/or text specifying frequency. Consider circling areas of interest that might require closer attention, which can be typically the case for "EQUI" and "L&D&P" procedures. Consider highlighting areas of interest, such as equipment and fingers for "EQUI" procedures, the thread for "L&D&P" procedures, and the knot for "C&T/M" procedures.

TIMING. Here, we refer to aspects related to when certain visualizations appeared in the resulting design proposals. Most instances (13: ID1, ID2, ID4-ID9, ID13, ID14, ID17, ID20, ID23) involved displaying more than one visualization that was displayed sequentially: first, the warning sign (and sometimes the error), which then disappears appearing another visualization displaying the correct procedure. At times, the error is also displayed in the second step (ID1, ID2), this is when the guidance benefits from a side-to-side visualization. Then, a few (6) of the resulting design proposals (ID3, ID10, ID12, ID15, ID16, ID19) involved visualizations that appeared *at the same time*. This relates to the complexity of the error and the amount of (visual) explanation and guidance needed to understand the error and/or the correction. For example, ID14 is a complex sequence procedure ("SEQ") that resulted in the display of three sequences of actions (wrong, right, wrong procedure) (See descriptions of procedure in section 4.1 and of visualization in section 4.2.1; Also see full descriptions in Table 5 in Supplementary Material). Another example: ID2, and equipment handling error ("EQUI") involving the needle holder in phase 1 (See left (a,b) in Fig. 2) unfolds in a sequence: first an error warning is signalled on the student's hands; then on the ghost hands, a detailed visualization explaining what is wrong and how it should be done appear. A similar situation happens for ID1 another P1 equipment handling error involving the forceps (See right (a,b) in Fig. 2). In contrast, ID3 shows another equipment handling error in P1 involving trying to pick up the needle with the fingers while it should be done with the needle holder displays at the same time the error and the guidance (See Figure 4 in Supplementary Material). This error is likely to be understood in a glimpse. The shown forceps in the guidance acts *symbolically* to tell the student the following message: "Pick up the needle with the forceps" rather than directly with their hands. Similar cases are ID10 (dropping the forceps on the working area), ID12 (dropping the needle in the work area), and ID19 (holding the forceps and needle holder). While these are all equipment handling procedures, there are also two counts, type, and mode of stitches ("C&T/M") procedures (ID15, and ID16), a "TIME" procedure (ID17), all of which can be considered featuring symbolic warnings and guidance, and add simple text. E.g. ID17 ("TIME") displays above the student's hands a blinking red time symbol and a message in red: ">30 sec." Then, ID15, and ID16, which are about knots (ID15, the number of knots performed; ID16, whether they are simple or double), feature knot symbols and simple text (e.g. ID15 may display "x3" to indicate three knots should be done). These kinds of errors and guidance contrast with the above-explained ID1, and ID2 that are displayed on a sequence, in that the guidance of ID1, and ID2 requires a more nuanced perception and understanding of what exactly is wrong and how to correct it.

Derived design implications. Corrections for procedures and errors that are likely easier to perceive and understand can be displayed at the same time as their guidance. More complex errors and/or correct procedures require a sequence of visualizations first showing the error, and then the guidance. In suturing, the former were typically "EQUI" and "C&T/M" procedures displaying symbolic warnings and guidance; and the latter were "EQUI" and "SEQ" that required careful attention to pick up seemingly subtle differences (e.g. where in the hand to position the tip of the forceps, ID2).

WHERE. Here, we address where the error is signalled and explained, as well as the guidance. While we expected to obtain many instances involving the **student's hand & the ghost's hands**, only three featured those (ID1-ID3): all "EQUI" types (handling the forceps, the needle handle, and the needle with needle handle instead of hands). They signalled the error in the student's hand (red cross and an arrow pointing to a red circled area in the student's hand), and showed both the error and the guidance in the ghost's hands (See e.g. Fig 2). This was an aspect discussed among the team, which initially explored the option of displaying error and error explanations on the student's hand and using the ghost's only for guidance. However, we could not rely on the "real" student's hand to remain static, in the same position during the error explanation, which could complicate the design considerably[4]. Hence, only a brief error warning sign (red cross, arrow, and circled area) would refer to the "real" student's hand. Also, in some cases, the team (expert and designers) found particularly useful a certain hand position (e.g. supine in ID2, see left (a, b) in Fig. 2) to show the error and the guidance. Since designers have more control over the ghost hands (e.g. position), we decided to use them for both error explanation and guidance.

Most instances (10) involved using the **student's kit & the ghost's kit** (ID4-ID9, ID15, ID16, ID20, ID23). Almost all (8) of these cases are related to either length, distance, and position ("L&D&P"); the rest (2), are to count, type, and kind of stitches ("C&T/M"). This is not surprising since both kinds of procedures and errors require references to key elements (insertion points in the skin, wound, thread, etc.) tied to the kit. This made us consider using augmented reality (AR) to display error and guidance explanations, which seemed the best option precisely because of how important the reference to the kit was. Unlike in the student's hand & the ghost's

[4]Note that, in general, we tended to lean towards the more simple solution that would likely address the problem in an effective way.

hands case, here we expected the kit to be relatively static and could rely on relatively simply overlaying the relevant information on the kit.

Then, a few instances (5) required both the kit or working space, as well as the hands in some form. Three (ID10, ID12, ID19) required the **student's kit (work area) & the ghost's hand**. All of those were "EQUI" types of procedures and errors that involved dropping equipment in the work area, hence the need to display that area. AR would be used to overlay a red cross and arrow pointing to the equipment of interest. *At the same time*, the ghost's hands would appear showing the correct handling of the equipment. As mentioned earlier (timing), here the handling of the equipment would be rather *symbolic*, and act as a reminder to hold onto the equipment. Then, one instance (ID13) required the **student's kit + the student's hand & ghost's kit and the ghost's hand**. This was a "L&D&P" procedure/error related to the length of the thread from the extraction point of the needle (placed on the kit) and the point where the surgeon would pick up the thread (their fingers) in preparation for a knot. Hence, the need emerged to show both the kit and hand ("real" and ghost's). Another instance (ID14) required the **student's kit & the ghost's hand and the ghost's kit**. This is the error of not lowering the knot straight. The student's kit is used just to display a warning with a red cross and arrow pointing to the knot, and a red symbol of a twisted knot next to the knot in the kit. Here, both the ghost's hands and kit are necessary to display the sequences of actions showing the correct and the incorrect movement of the hands, i.e. the initial, middle, and last position of the hands while doing this. Since this position and movement take as reference the loose thread end, which is placed on the kit, both the ghost's kit and the ghost's hands need to appear.

The last instance, ID17, is just a "TIME" error, which does not require referring to the hands or kit specifically. A blinking red time symbol and some text (">30 sec") appear above and next to the student's hands, indicating a procedure is taking too long.

Derived design implications. Errors uniquely related to the hands of the student should display the student's hands as well as the ghost's hands. Consider using the "real" student's hands to display a warning error, and the ghost's hands to illustrate the error and produce guidance. Errors related to the kit and elements tied to the kit should display the student's kit & the ghost's kit. AR should be used to overlay visual pointers to the error on "the real kit."

5 CONCLUDING REMARKS

In this paper, we have focused our design inquiry on understanding needs emerging from the lived experience of an expert surgeon and teacher with the technology prior to (and in parallel to) co-creating possible future design concepts. This is the result of our Research through Design approach, which lies in stark contrast to existing approaches and trends in technology-oriented works, which tend to impose arbitrary designs and simplified assessment metrics [33, 39].

In particular, through our hybrid bodystorming, we were able to build a rich and nuanced description of the phases and procedures of the suturing process, propose design solutions, and derive and reflect on emerging patterns. Our hybrid embodied design method is timely, supported by very recent technological advances, such as XR headsets with high-resolution passthrough capabilities. Such headsets can become an embodied tool that supports participants to think and act, communicate and share ideas, and in general, fuel their creative process. Like in prior works [44], passthrough capabilities allow blending in elements from the physical and virtual worlds. This supported the creative process of the social hybrid game of Astaire [44]. In our case, it supported reflection and articulation. In particular, it helped the surgeon better articulate key procedures, errors, and assessment metrics within the suturing process. While problems and requirements, and even potential solutions already emerged during the early stages of our inquiry (i.e. in debriefing sessions with the experts using classical research methods like interviews), the hybrid embodied design activity was essential to complete (i.e. making more comprehensive) our list of problems to address. It also helped refine, prioritize, further classify, and even better articulate these problems thanks to their materialization through physical enactment.

For example, during early debriefs, assessment metrics were very unclear, in part because verbally articulating movement errors and corrections is extremely hard [17, 28, 35]. This was particularly clear with a complex sequence of actions in the "Knott" phase (P3), i.e. ID14 "lower the knot straight." The surgeon could easily identify which knot was straight (after the fact, and also foresee it when it was being done), and also had a sense of the importance of certain hand movements ("like walking," he explained); Yet, until physical enactment (several strategies that helped in concrete were: repetition, slowing down the pace, and gestural amplification), it was not clear exactly what made a knot "straight," which movements and positions were key and under which conditions.

This resonates with the concepts of "tacit knowledge" and "thinking through doing" described by Polanyi and Schön [28, 35]. Insights into the sociology of knowledge explain why our knowledge and skills are tacit and resist verbal articulation [17]. We are highly more able to focus on skill performance and not on constituent gestures, movements, and physical actions. In Polanyi's words: "We are attending from these elementary movements to the achievement of their joint purpose, and hence are usually unable to specify these elementary acts" [28].

Another reason behind the difficulty in articulating problems well and finding interesting solutions lies in the fact that the surgeon was not initially aware of the capabilities of the technology despite being a technology enthusiast. To this, the design team suggested several possibilities considering the technology capabilities: e.g. angle of the hands, speed, etc., which the expert deemed interesting. However, most of them faded away to the light of actual errors and corrections that would be meaningful for the students, which only emerged through physical enactment.

The hybrid bodystorming was not only key to elucidating problems, but also to jointly envisioning meaningful solutions, from pedagogical, design, and technical perspectives.

From a design perspective, exploring and testing different visualizations (i.e. physical props in our bodystorming basket that the surgeon could thanks to the passthrough capabilities of the headset) helped the expert and us understand if and how what was in front of us would work and imagine other new visualizations. This relates to Schön's idea that sketches, in this case, embodied sketches [21] (i.e. sketches done with the body instead of with pen and paper), "talk-back" to the practitioner, supporting understanding (i.e. helping practitioners to read and "see as") and evoking new ideas (i.e. helping practitioners to "see that") [4, 35]. This adds to the body of works using embodied design methods in XR (e.g. [41, 44]), which support live embodied improvisation and design. Both in [44] and this work, this involved using objects in the physical space (props) to imagine them in VR and XR. In [41], they used objects in the physical world but also operated technical controls operated by a "puppeteer," which triggered real-time events in VR. Different than in [44], the props in our bodystorming basket were carefully curated, including wooden hand models, which were mostly used in the hybrid design session to represent the ghost's hands; and physical visualization prompts, which would be used to explore error signalling and potential corrections and guidance for future design (See Figure 1). In turn, in Astaire [44], more general props were brought to the hybrid bodystorming session, such as play toys, cardboard boxes, and everyday gadgets bought in bazaars. This was in part due to the fact that the bodystorming in [44] was very open and explorative, and meant to open the design space widely; while in our work, we were targetting from the onset particular kinds of designs, i.e. displaying procedures, and errors. Hence this narrower scope benefited from specific physical props to the intended future design.

From a design and technical perspective, the hybrid bodystorming process also helped us designers to shift design directions and choose the right technology for our design. We were opting for a VR experience at the beginning, however, using a device (Meta Quest Pro) that seamlessly allowed us to experiment with both virtual reality and mixed-reality setups helped us realize the importance of AR (and hence XR): real-world haptic feedback is key for surgical training, and we (surgeon and designers) found extremely interesting the possibility of using virtual elements to provide visual cues that are spatially-aligned with the physical elements currently used in their educational practices (e.g. suturing kit). Hence, this will set our immediate future design direction. It would also be interesting for future work, to explore VR setups and investigate in a deeper way the effect of the physical feedback on learning as well as explore to what extent the learning process of rich and complex body movements would be supported by using movement reduction strategies, similar to marking in dance, instead of the full movement for rehearsal [14].

The hybrid bodystorming process also allowed us to uncover peculiar needs for future technology, such as AI-powered technology for the automatic detection of errors and assessment of the stitching process, e.g. choosing the right tool, performing the stitches, the type of knots, or understanding the position of the hand and fingers relative to the tools. We will address this in future work as well. During the co-design process, we gathered practical knowledge on how best to collect data on the hands' joints. It will be valuable for creating a dataset for training machine learning models to automatically assess performance and provide error cues for gesture training in XR environments [12].

Then, our hybrid embodied design process helped better articulate procedures, and errors, and visualize how these errors materialize and can be addressed by practitioners. Our expert surgeons expressed the usefulness of these outcomes from a pedagogical perspective, and we expect this work to be already of use in surgical training classes.

We acknowledge the limitations of not including novice learners at this step of the design process, which precluded us from gathering insights into their obstacles, needs and desires. As a future step, we aim to involve medical students in the evaluation of the proposed solution and learn about the difficulty they encounter in their training (which could be not only skill-related but also cognitive, physiological or psychological), drawing from theoretical models of embodied and multisensory learning [1, 37].

We conclude with a method call for new innovative hybrid embodied design methods, which will be key to supporting the emergence of meaningful XR designs. We hope this work is part of the beginning of a new wave of innovative design methods featuring hybrid setups, which will leverage participants' embodied exploration (situated, physical and social) of current needs, and innovative future possibilities.

ACKNOWLEDGMENTS

This research was partially supported by the Madrid Government (Comunidad de Madrid) under the Multiannual Agreement with UC3M in the line of Research Funds for Beatriz Galindo Fellowships (MovIntPlayLab-CM-UC3M 2021/00050/001) in the context of the V PRICIT (Regional Programme of Research and Technological Innovation), by the EU Erasmus+ project Method Cards for Movement-based Interaction Design (2020-1-DK01-KA203-075164), and by the Instituto de Salud Carlos III (ISCIII) (PI22/00862).

REFERENCES

[1] Julian Abich, Jason Parker, Jennifer S. Murphy, and Morgan Eudy. 2021. A review of the evidence for training effectiveness with virtual reality technology. *Virtual Reality* 25, 4 (Dec. 2021), 919–933. https://doi.org/10.1007/s10055-020-00498-8
[2] Medhat Alaker, Greg R. Wynn, and Tan Arulampalam. 2016. Virtual reality training in laparoscopic surgery: A systematic review & meta-analysis. *International Journal of Surgery* 29 (May 2016), 85–94. https://doi.org/10.1016/j.ijsu.2016.03.034
[3] Luluah Albarrak, Oussama Metatla, and Anne Roudaut. 2023. Using Virtual Reality and Co-Design to Study the Design of Large-scale Shape-Changing Interfaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3544548.3581144
[4] Bill Buxton. 2007. *Sketching User Experiences: Getting the Design Right and the Right Design.* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
[5] Janis A. Cannon-Bowers, Clint Bowers, and Katelyn Procci. 2010. Optimizing Learning in Surgical Simulations: Guidelines from the Science of Learning and Human Performance. *Surgical Clinics of North America* 90, 3 (June 2010), 583–603. https://doi.org/10.1016/j.suc.2010.02.006
[6] A. Chellali, W. Ahn, G. Sankaranarayanan, J. T. Flinn, S. D. Schwaitzberg, D. B. Jones, Suvranu De, and C. G. L. Cao. 2015. Preliminary evaluation of the pattern cutting and the ligating loop virtual laparoscopic trainers. *Surgical Endoscopy* 29, 4 (April 2015), 815–821. https://doi.org/10.1007/s00464-014-3764-7
[7] Amanda B. Cooper and Elizabeth J. Tisdell. 2020. Embodied aspects of learning to be a surgeon. *Medical Teacher* 42, 5 (May 2020), 515–522. https://doi.org/10.1080/0142159X.2019.1708289
[8] Alexandra C. Cope, Stella Mavroveli, Jeff Bezemer, George B. Hanna, and Roger Kneebone. 2015. Making Meaning From Sensory Cues: A Qualitative Investigation of Postgraduate Learning in the Operating Room. *Academic Medicine* 90, 8 (Aug. 2015), 1125–1131. https://doi.org/10.1097/ACM.0000000000000740
[9] Eléonore Ferrier-Barbut, Ignacio Avellino, Geoffroy Canlorbe, Marie-Aude Vitrani, and Vanda Luengo. 2023. Learning With Pedagogical Models: Videos As Adjuncts to Apprenticeship for Surgical Training. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (April 2023), 1–40. https://doi.org/10.1145/3579615
[10] Payal Guha, Jason Lawson, Iona Minty, James Kinross, and Guy Martin. 2023. Can mixed reality technologies teach surgical skills better than traditional methods?

A prospective randomised feasibility study. *BMC Medical Education* 23, 1 (March 2023), 144. https://doi.org/10.1186/s12909-023-04122-6
[11] Christian Heath, John Hindmarsh, and Paul Luff. 2010. *Video in Qualitative Research: Analysing Social Interaction in Everyday Life* (1 edition ed.). SAGE Publications Ltd, Los Angeles.
[12] Florian Jeanne, Yann Soullard, and Indira Thouvenin. 2016. What is wrong with your gesture? An error-based assistance for gesture training in virtual environments. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 247–248.
[13] Hans-Christian Jetter, Roman Rädle, Tiare Feuchtner, Christoph Anthes, Judith Friedl, and Clemens Nylandsted Klokmose. 2020. "In VR, everything is possible!": Sketching and Simulating Spatially-Aware Interactive Spaces in Virtual Reality. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–16. https://doi.org/10.1145/3313831.3376652
[14] David Kirsh. 2013. Embodied Cognition and the Magical Future of Interaction Design. *ACM Trans. Comput.-Hum. Interact.* 20, 1, Article 3 (apr 2013), 30 pages. https://doi.org/10.1145/2442106.2442109
[15] Hillary Lia, Gregory Paulin, Nelson Yi, Hassan Haq, Steve Emmanuel, Kristian Ludig, Zsuzsanna Keri, Andras Lasso, Gabor Fichtinger, Caitlin T. Yeo, and Jessica Andrews. 2018. HoloLens in suturing training. In *Medical Imaging 2018: Image-Guided Procedures, Robotic Interventions, and Modeling*, Robert J. Webster and Baowei Fei (Eds.). SPIE, Houston, United States, 69. https://doi.org/10.1117/12.2293934
[16] Alexander Hvidbjerg Kjær Lund, Amalie Finnemannn Sørensen, Lars Elbæk, and Maximus D. Kaos. 2021. Insights from design processes used in developing exergames. In *15th European Conference on Game Based Learning, ECGBL 2021*. Dechema eV, 490–498.
[17] Claudia Mareis. 2012. The Epistemology of the Unspoken: On the Concept of Tacit Knowledge in Contemporary Design Research. *Design Issues* 28, 2 (2012), 61–71.
[18] Joshua McVeigh-Schultz, Max Kreminski, Keshav Prasad, Perry Hoberman, and Scott S. Fisher. 2018. Immersive Design Fiction: Using VR to Prototype Speculative Interfaces and Interaction Rituals within a Virtual Storyworld. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. Association for Computing Machinery, New York, NY, USA, 817–829. https://doi.org/10.1145/3196709.3196793
[19] Adnan Munawar, Zhaoshuo Li, Punit Kunjam, Nimesh Nagururu, Andy S. Ding, Peter Kazanzides, Thomas Looi, Francis X. Creighton, Russell H. Taylor, and Mathias Unberath. 2022. Virtual reality for synergistic surgical training and data generation. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization* 10, 4 (July 2022), 366–374. https://doi.org/10.1080/21681163.2021.1999331
[20] Elena Márquez Segura, Laia Turmo Vidal, and Asreen Rostami. 2016. Bodystorming for movement-based interaction design. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments* 12, 2 (2016), 193–251. https://doi.org/10.17011/ht/urn.201611174655
[21] Elena Márquez Segura, Laia Turmo Vidal, Asreen Rostami, and Annika Waern. 2016. Embodied Sketching. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 6014–6027. https://doi.org/10.1145/2858036.2858486
[22] Rebekah A. Naylor, Lisa A. Hollett, R. James Valentine, Ian C. Mitchell, Monet W. Bowling, A. Moe Ma, Sean P. Dineen, Brandon R. Bruns, and Daniel J. Scott. 2009. Can medical students achieve skills proficiency through simulation training? *The American Journal of Surgery* 198, 2 (Aug. 2009), 277–282. https://doi.org/10.1016/j.amjsurg.2008.11.036
[23] Jakob Nielsen. 1993. *Usability Engineering* (1st edition ed.). Morgan Kaufmann, Boston.
[24] Antti Oulasvirta, Esko Kurvinen, and Tomi Kankainen. 2003. Understanding contexts by being there: case studies in bodystorming. *Personal and Ubiquitous Computing* 7, 2 (July 2003), 125–134. https://doi.org/10.1007/s00779-003-0238-7
[25] Robert V. O'Toole, Robert R. Playter, Thomas M. Krummel, William C. Blank, Nancy H. Cornelius, Webb R. Roberts, Whitney J. Bell, and Marc Raibert. 1999. Measuring and developing suturing technique with a virtual reality surgical simulator11All authors, except Krummel, have an affiliation with Boston Dynamics Inc., an organization that has a financial interest in the editorial content discussed in the manuscript. *Journal of the American College of Surgeons* 189, 1 (July 1999), 114–127. https://doi.org/10.1016/S1072-7515(99)00076-9
[26] Jocelyn Parong and Richard E. Mayer. 2021. Cognitive and affective processes for learning science in immersive virtual reality. *Journal of Computer Assisted Learning* 37, 1 (Feb. 2021), 226–241. https://doi.org/10.1111/jcal.12482
[27] Panayiotis E. Pelargos, Daniel T. Nagasawa, Carlito Lagman, Stephen Tenn, Joanna V. Demos, Seung J. Lee, Timothy T. Bui, Natalie E. Barnette, Nikhilesh S. Bhatt, Nolan Ung, Ausaf Bari, Neil A. Martin, and Isaac Yang. 2017. Utilizing virtual and augmented reality for educational and clinical enhancements in neurosurgery. *Journal of Clinical Neuroscience* 35 (Jan. 2017), 1–4. https://doi.org/10.1016/j.jocn.2016.09.002
[28] M Polanyi. 1966. The Tacit Dimension (Edition 2009 with a new forward by Amartya Sen).
[29] Rachel Prentice. 2007. Drilling Surgeons: The Social Lessons of Embodied Surgical Learning. *Science, Technology, & Human Values* 32, 5 (Sept. 2007), 534–553. https://doi.org/10.1177/0895904805303201
[30] Rachel Prentice. 2013. *Bodies information: an ethnography of anatomy and surgery education.* Duke University Press, Durham, NC.
[31] Unnikrishnan Radhakrishnan, Francesco Chinello, and Konstantinos Koumaditis. 2023. Investigating the effectiveness of immersive VR skill training and its link to physiological arousal. *Virtual Reality* 27, 2 (June 2023), 1091–1115. https://doi.org/10.1007/s10055-022-00699-3
[32] Anke V. Reinschluessel, Thomas Muender, Daniela Salzmann, Tanja Döring, Rainer Malaka, and Dirk Weyhe. 2022. Virtual Reality for Surgical Planning – Evaluation Based on Two Liver Tumor Resections. *Frontiers in Surgery* 9 (Feb. 2022), 821060. https://doi.org/10.3389/fsurg.2022.821060
[33] Ana Rojo, Laura Raya, and Alberto Sanchez. 2021. A Novel Mixed Reality Solution Based on Learning Environment for Sutures in Minor Surgery. *Applied Sciences* 11, 5 (March 2021), 2335. https://doi.org/10.3390/app11052335
[34] Dennis Schleicher, Peter Jones, and Oksana Kachur. 2010. Bodystorming as embodied designing. *Interactions* 17, 6 (Nov. 2010), 47–51. https://doi.org/10.1145/1865245.1865256
[35] Donald A. Schön. 1983. *The reflective practitioner: how professionals think in action.* Ashgate, Farnham.
[36] Adalberto L. Simeone, Robbe Cools, Stan Depuydt, João Maria Gomes, Piet Goris, Joseph Grocott, Augusto Esteves, and Kathrin Gerling. 2022. Immersive Speculative Enactments: Bringing Future Scenarios and Technology to Life Using Virtual Reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 17, 20 pages. https://doi.org/10.1145/3491102.3517492
[37] Alexander Skulmowski and Günter Daniel Rey. 2018. Embodied learning: introducing a taxonomy based on bodily engagement and task integration. *Cognitive research: principles and implications* 3, 1 (2018), 1–10.
[38] Erik Stolterman. 2008. The Nature of Design Practice and Implications for Interaction Design Research. *International Journal of Design* 2, 1 (2008), 55–65.
[39] Yuka Tashiro, Shio Miyafuji, Dong-Hyun Hwang, Satoshi Kiyofuji, Taichi Kin, Takeo Igarashi, and Hideki Koike. 2023. GAuze-MIcrosuture-FICATION: Gamification in Microsuture training with real-time feedback. In *Augmented Humans Conference*. ACM, Glasgow United Kingdom, 15–26. https://doi.org/10.1145/3582700.3582704
[40] José Manuel Vega-Cebrián, Elena Márquez Segura, Laia Turmo Vidal, Omar Valdiviezo-Hernández, Annika Waern, Robby Van Delden, Joris Weijdom, Lars Elbæk, Rasmus Vestergaard Andersen, Søren Stigkær Lekbo, and Ana Tajadura-Jiménez. 2023. Design Resources in Movement-based Design Methods: a Practice-based Characterization. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. Association for Computing Machinery, New York, NY, USA, 871–888. https://doi.org/10.1145/3563657.3596036
[41] Joris Weijdom. 2022. Performative prototyping in collaborative mixed reality environments: an embodied design method for ideation and development in virtual reality.. In *Sixteenth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '22)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3490149.3501316
[42] Danielle Wilde, Anna Vallgårda, and Oscar Tomico. 2017. Embodied Design Ideation Methods: Analysing the Power of Estrangement. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5158–5170. https://doi.org/10.1145/3025453.3025873
[43] George J. Xeroulis, Jason Park, Carol-Anne Moulton, Richard K. Reznick, Vicki LeBlanc, and Adam Dubrowski. 2007. Teaching suturing and knot-tying skills to medical students: A randomized controlled study comparing computer-based video instruction and (concurrent and summary) expert feedback. *Surgery* 141, 4 (April 2007), 442–449. https://doi.org/10.1016/j.surg.2006.09.012
[44] Zhuoming Zhou, Elena Márquez Segura, Jared Duval, Michael John, and Katherine Isbister. 2019. Astaire: A Collaborative Mixed Reality Dance Game for Collocated Players. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '19)*. ACM, New York, NY, USA, 5–18. https://doi.org/10.1145/3311350.3347152 event-place: Barcelona, Spain.
[45] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research Through Design As a Method for Interaction Design Research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*. ACM, New York, NY, USA, 493–502. https://doi.org/10.1145/1240624.1240704
[46] John Zimmerman, Erik Stolterman, and Jodi Forlizzi. 2010. An Analysis and Critique of Research Through Design: Towards a Formalization of a Research Approach. In *Proceedings of the 8th ACM Conference on Designing Interactive Systems (DIS '10)*. ACM, New York, NY, USA, 310–319. https://doi.org/10.1145/1858171.1858228
[47] Peter Zimmermann. 2008. *Virtual Reality Aided Design. A survey of the use of VR in automotive industry.* Springer Netherlands, Dordrecht, 277–296. https://doi.org/10.1007/978-1-4020-8200-9_13