Article

TrainAR: A Scalable Interaction Concept and Didactic Framework for Procedural Trainings Using Handheld Augmented Reality

Jonas Blattgerste [1,*], Kristina Luksch [2], Carmen Lewa [3] and Thies Pfeiffer [1]

1. Faculty of Technology, University of Applied Sciences Emden/Leer, Constantiaplatz 4, 26723 Emden, Germany; thies.pfeiffer@hs-emden-leer.de
2. Institut für Angewandte Gesundheitsforschung—Midwifery & Reproductive Health, Hochschule für Gesundheit Bochum, Gesundheitscampus 6-8, 44801 Bochum, Germany; kristina.luksch@hs-gesundheit.de
3. Zentrum für Medizinische Lehre, Ruhr-Universität Bochum, Universitätsstraße 150, 44801 Bochum, Germany; carmen.lewa@ruhr-uni-bochum.de
* Correspondence: jonas.blattgerste@hs-emden-leer.de

Abstract: The potential of Augmented Reality (AR) for educational and training purposes is well known. While large-scale deployments of head-mounted AR headsets remain challenging due to technical limitations and cost factors, advances in mobile devices and tracking solutions introduce handheld AR devices as a powerful, broadly available alternative, yet with some restrictions. One of the current limitations of AR training applications on handheld AR devices is that most offer rather static experiences, only providing descriptive knowledge with little interactivity. Holistic concepts for the coverage of procedural knowledge are largely missing. The contribution of this paper is twofold. We propose a scalabe interaction concept for handheld AR devices with an accompanied didactic framework for procedural training tasks called TrainAR. Then, we implement TrainAR for a training scenario in academics for the context of midwifery and explain the educational theories behind our framework and how to apply it for procedural training tasks. We evaluate and subsequently improve the concept based on three formative usability studies (n = 24), where explicitness, redundant feedback mechanisms and onboarding were identified as major success factors. Finally, we conclude by discussing derived implications for improvements and ongoing and future work.

Keywords: mobile; handheld; augmented reality; mixed reality; procedural; education; learning; training; framework; interaction concept; midwifery

1. Introduction

In their book "The Teaching Gap", Stigler and Hiebert wrote in 1999: "School learning will not improve markedly unless we give teachers the opportunity and support they need to advance their craft by increasing the effectiveness of the methods they use" [1]. Since then, digitization of learning provided new opportunities for teaching, e.g., by introducing asynchronous learning approaches based on e-learning techniques. This not only allows teachers to work more efficiently but also provides benefits to the learner, such as spatially independent communication, self-regulated learning, as well as access to learning anytime and anyplace [2].

1.1. Computer- and Web-Based Trainings

Computer-Based Trainings (CBT) started in the 70s to increase the learners' independence in space and time. In the 90s, the emerging internet yielded advanced Web-Based Trainings (WBT), providing new ways to intensify teacher-learner as well as learner-learner communication and to provide new forms of feedback. WBT approaches in general have been shown to be well accepted by students, reporting self efficacy as the main factor, independent of the perceived usefulness [3]. Beyond the motivational benefits though, WBTs

struggle to be more effective than traditional approaches, as a meta analysis showed [4]. This might be attributed to the challenge of creating high quality and sustainable e-learning content [5], which is underlined by the finding that different desired learning outcomes require different kinds of instructions [6].

For procedural training tasks, which primarily consist of a combination of cognitive strategies and motor skills rather than basic declarative knowledge [7], conventional CBTs or WBTs might not be sufficient. The reason becomes apparent when contextualizing the coverage of CBTs and WBTs in Bloom's Taxonomy [8]. It is hard to argue that CBTs can support the learner beyond the levels of *remembering* and *understanding*. While WBTs, with the embracing of social media and communication components, also address the fifth level, *evaluation*, they neither sufficiently support the learners in *applying* (3rd level) or *analyzing* (4th level) procedural task knowledge nor do they provide the freedom of exploration that would be necessary to reach the highest level of Bloom's Taxonomy: *creating*. Though WBTs can arguably be applied to the evaluating level of procedural task learning, it has to be noted that this mostly consists of quite time consuming, often hand-crafted, methods.

1.2. Augmented Reality-Based Trainings

What is needed in terms of technical features to fully address the third and fourth level? One technology potentially able to fill this gap is Augmented Reality. Endeavors towards Augmented Reality-based Training (ARBT) combine the benefits of WBTs with AR's biggest strength of contextualizing information in the physical world. This makes ARBTs interesting for practical training and procedural tasks [9]. Current findings indicate that applying AR as an additional "multimedia source" into existing curricula can already lead to improved retention, attention and satisfaction [10]. Furthermore, a meta analysis conducted by Ozdemir et al. [11] indicates increased academic achievement compared to traditional learning methods, increased concentration and the enabling of teachers to convey concepts faster and with more clarity through demonstration of connections between concepts and principles. Generally, systematic literature reviews also point towards a consistently positive impact of AR tools used in educational settings [12], especially through interaction, catching the learners attention and increasing motivation [13]. Notably, while significant differences can be observed for all levels of education, the largest effect size of learning benefits is observed for students of undergraduate level [11].

1.3. Acceptance & Scalability of ARBTs

Despite those apparent didactic benefits, several challenges for a realistic, scalable deployment of ARBTs into training procedures remain. For one, AR-headsets are still expensive and have a half-life period under 2 years, which renders it almost impossible to deploy larger set-ups at University level. They thus do not scale up to group sizes of today's university-level training of practical skills or vocational training. Also, the technology has limitations, such as a narrow field of view, experimental gesture-based interaction methods and unstable tracking under non-optimal conditions. In combination with a lack of media competences in teachers and students with this technology, this can lead to acceptance problems.

As success factors for AR deployment, user experience, stability, adaptability, and independent self learning capabilities have been identified [14,15]. Technology acceptance models (TAM) applied to potential AR trainings in educational contexts show that students perceive the technology as useful, easy to use [16] and teachers attitudes imply their intention to utilize AR [17]. Nonetheless, those studies measure perceived use and not actual usage [18]. While behavioural intention can influence morale, disposition and performance, perceived usefulness and perceived ease of use are not reliable indicators for practical acceptance and consequent usage [18].

Guiding educational theories tailored towards AR training are still ongoing research [15]. However, generalized concepts for AR training that teachers can directly apply into their curriculum are a primary demand from their perspective [19].

Those limitations are, in our view, the reasons AR is still not extensively used in education. While those limitations also apply to handheld AR applications, the users' familiarity with smartphones as well as recent advancements in hardware and tracking solutions (e.g., ARCore [20] for Android & ARKit [21] for iOS smartphones) make them feasible candidates as platforms for AR training applications that can be realistically implemented in educational curricula, e.g., even considering bring-your-own-device (BYOD) approaches. As a consequence, in line with the success factors identified by Dalim et al. [14] and Cheng et al. [15], for us **scalability** requires:

1. **ubiquitous availability** of devices,
2. **place and time independence** for self-regulated learning,
3. **high usability and low entry threshold** to compensate low levels of media competency,
4. **clear concepts for interaction and didactic** to maximize the support for teachers in defining new learning materials

Requirement number one is technically met by consumer smartphones of recent years, as has been detailed above. AR-based training is independent of the availability of the teacher, and thus in principle time independent. Whether it is place independent depends on the required context objects: expensive special purpose devices might only be available in laboratories or special training facilities and thus restrict spatial flexibility. To provide maximum spatial flexibility, AR applications can, however, provide alternative virtual proxy environments and context objects in cases the physical ones are not accessible.

For the last two requirements, comprehensive interaction concepts, including feedback mechanisms and didactic contextualisation are still largely missing. This is especially true for more complex training scenarios, such as procedural training tasks. Systematic literature reviews reveal that most handheld-based AR training scenarios are rather static, only displaying non-procedural information bits with very little to no interaction [22,23]. They also often only focus on small learning scopes, mostly covering only an isolated topic without long-term focus or feasible scopes of deployment beyond what was necessary for evaluation [13]. While exemplary, scenario-specific AR training applications already elicit the mentioned didactic benefits, there is a need for generalized concepts that work beyond isolated topics for targeted evaluation studies, in particular addressing the challenges of scalability and long-term deployments.

Therefore the contribution of this paper is twofold: First, we propose Train*AR*, a scalable interaction concept for handheld Augmented Reality devices in combination with a didactic contextualization framework. The combination is intended to supplement the training of procedural training tasks. Second, we report on an iterative evaluation of an implementation of Train*AR* for Android & iOS in the practical training of academic midwifery.

2. Related Work

While holistic interaction concepts and frameworks for the application of AR for the training of procedural tasks are largely missing and even specific procedural trainings on handheld AR devices are sparse, there is some notable research on procedural trainings utilizing head-mounted and projection-based Augmented Reality approaches for specific procedural scenarios but also interaction concepts for handheld augmented reality.

2.1. Procedural Augmented Reality Trainings

To develop procedural AR trainings, the action sequences, already existing conventional instructions and corresponding 3D data have to be either developed or combined and conferred for AR usage. Müller [24] explored this challenge in the context of manual procedural tasks in the context of maintenance. Here, they identified the five major challenges of clarity, consistency, visibility, orientation and information linking for the transformation of conventional to AR based training tasks. Chidambaram et al. [25] proposed a head-monuted AR-based training tool that can record procedural expert movements during tool usage e.g., for maintenance tasks, store them on the device itself and replay them as training instructions for novice users. Additionally, while projection-based AR

approaches are mostly utilized as fixed installations with permanent usage intentions, Büttner et al. [26] explored the appropriation of projection-based approaches for the means of procedural task training and showed that, while a systematic mislearning of procedures was prevented through the immediate, behaviouristic feedback of the AR system, pure AR trainings did not reach the personal training in terms of speed or recall for manual assembly tasks.

While a large portion of procedural task training is contextualized in manual assembly, maintenance and industrial settings, there is also notable previous work in other settings. For example, Singh et al. [27] utilized procedural AR task trainings for people with cognitive disabilities like autism. While the cognitively impaired students preferred in-person training, AR was still outperforming traditional desktop-based learning approaches. In the context of physics, Hruntova et al. [28] explored the usage of AR as a means of conveying experimental procedures in higher technical education settings. Similarly, Solmaz et al. [29] utilized procedural AR trainings to teach students liquid-soap synthesis processes to deepen students understanding of procedures and underlying chemical processes. Finally, Wang et al. [30] develop a training platform using the Microsoft Hololens, enabling remote procedural medical trainings, combining procedural task training with simultaneous AR task visualization and verbal & non-verbal communication from mentors remotely, without being in the same physical location.

2.2. Handheld Augmented Reality Interaction Concepts

After three decades of research on AR, interaction concepts still remain an area of active interest. With advancements in hardware, software and tracking algorithms, more advanced and increasingly intuitive interaction concepts are possible and explored in the literature. Especially in the area of handheld AR applications, major progress was made throughout recent years. Generally, those endeavours can be split in three major areas of AR interaction research: Tangible interaction techniques, e.g., utilizing tracked markers or tangible objects for interactions, Hand recognition approaches that utilize computer vision algorithms to enable manual interaction with virtual objects and "traditional" interaction metaphors that use on-screen interactions on the device to interact with augmented objects [12,31,32].

Tangible interaction techniques provide an inherent synergy effect with AR and some studies suggest they are perceived as more enjoyable compared to traditional interaction techniques [33]. This is especially true for procedural interactions as demonstrated by Billinghurst et al. [34], where they defined and explored the tangible interaction metaphor and proposed guidelines for tangible interfaces based on interaction prototypes. However, they are not only harder to implement for handheld AR devices but also require additional materials such as markers or tangible objects for the concepts to work, making them hard to scale compared to traditional interaction approaches like on-screen touch or hand-gesture-based approaches. Additionally, recent studies indicate that in terms of learning outcome, tangible interactions are not significantly increasing retention or transfer of knowledge compared to purely virtual interaction approaches [35].

With advances in tracking solutions, today even monocular RGB cameras are able to estimate hand positions in three-dimensional space, making gesture and hand recognition possible for consumer grade smartphones (in theory). Qian et al. [36] explored the design space of gesture interaction in handheld augmented reality, especially identifying the major usability issues perception, manipulation and behavioural understanding as the biggest hurdles to overcome. Hürst et al. [37] showed that, while gesture-based approaches provided a high entertainment value, they were lower in accuracy and more time consuming for the user, ultimately limiting their usage for serious tasks. Datcu et al. [38] found that user generally preferred hand-based interaction concepts in combination with physical objects, rather than using them as a means of interaction on virtual objects. Therefore, while no additional material or hardware is needed, interaction concepts based on hand-tracking come with their own drawbacks. While for VR headsets and head-mounted AR hand-

tracking is an intuitive way of interaction, on handheld AR this concept might not only be unfamiliar for the users, as they would have to judge the result of an interaction through the camera and screen of the handheld device instead of contextualized on the physical position of the hand itself but algorithms also still are not production ready, leading to questions concerning their stability and thus realistic scalability.

Ultimately, interaction concepts inspired by traditional smartphone-based interaction metaphors have the benefit of reusing familiar metaphors, users have already learned from traditional smartphone usage. As already the novel visual elements of AR are unfamiliar, sticking to well-known paradigms should not only increase the users' confidence in their ability to use the application but also speed up the users' onboarding process before usage. However, while traditional interaction concepts are well established for mostly two-dimensional interaction, AR interactions in 3D physical space and especially the non-traditional introduction of translation and rotation of the handheld device itself introduces new challenges and research questions. Besides the mentioned work by Hürst et al. [37], that also included traditional interaction metaphors, several interesting works with this focus have been published. Mossel et al. [39] compared on-screen touch interaction and interactions through translation and rotation of the smartphone for the manipulation of virtual objects. While they found that generally both are intuitive for the canonical manipulation of positioning and rotating objects, on-screen touch interaction was preferable if scaling was involved and manipulation of virtual objects through the translation and rotation of the smartphone outperformed on-screen touch interaction in terms of ease-of-use without the "scaling" interaction. Adding to those findings, Radu et al. [40] compared crosshair aiming and touch interactions for handheld augmented reality in the context of early child learning and found that for this target group, on-screen touch interaction was preferable, as the children were significantly faster, reported higher ease-of-use and higher comfort levels compared to crosshair interactions. Grandi et al. [41] compared a variety of different manipulations like translation, rotation, scaling and combining of virtual objects through interaction concepts based on device movements, touch gesture manipulations and a hybrid solution of both. While the hybrid interaction concept outperformed all others in terms of task completion times and user perception, it only outperformed the device movement condition in terms of errors. While the touch interaction was the slowest, it also resulted in the least errors.

3. Interaction Concept

The goal of the proposed handheld AR interaction concept is to be scalable, stable and easy to use in the context of procedural training tasks by users with varying levels of media competency. Therefore, it improves upon traditional AR interaction metaphors from both the literature and common non-AR applications and combines them for non-linear procedural interaction chains, creating a more holistic interaction concept. The proposed concept is mainly targeted at currently available consumer-grade Android & iOS smartphones but, with little changes, might also be applicable to tablets.
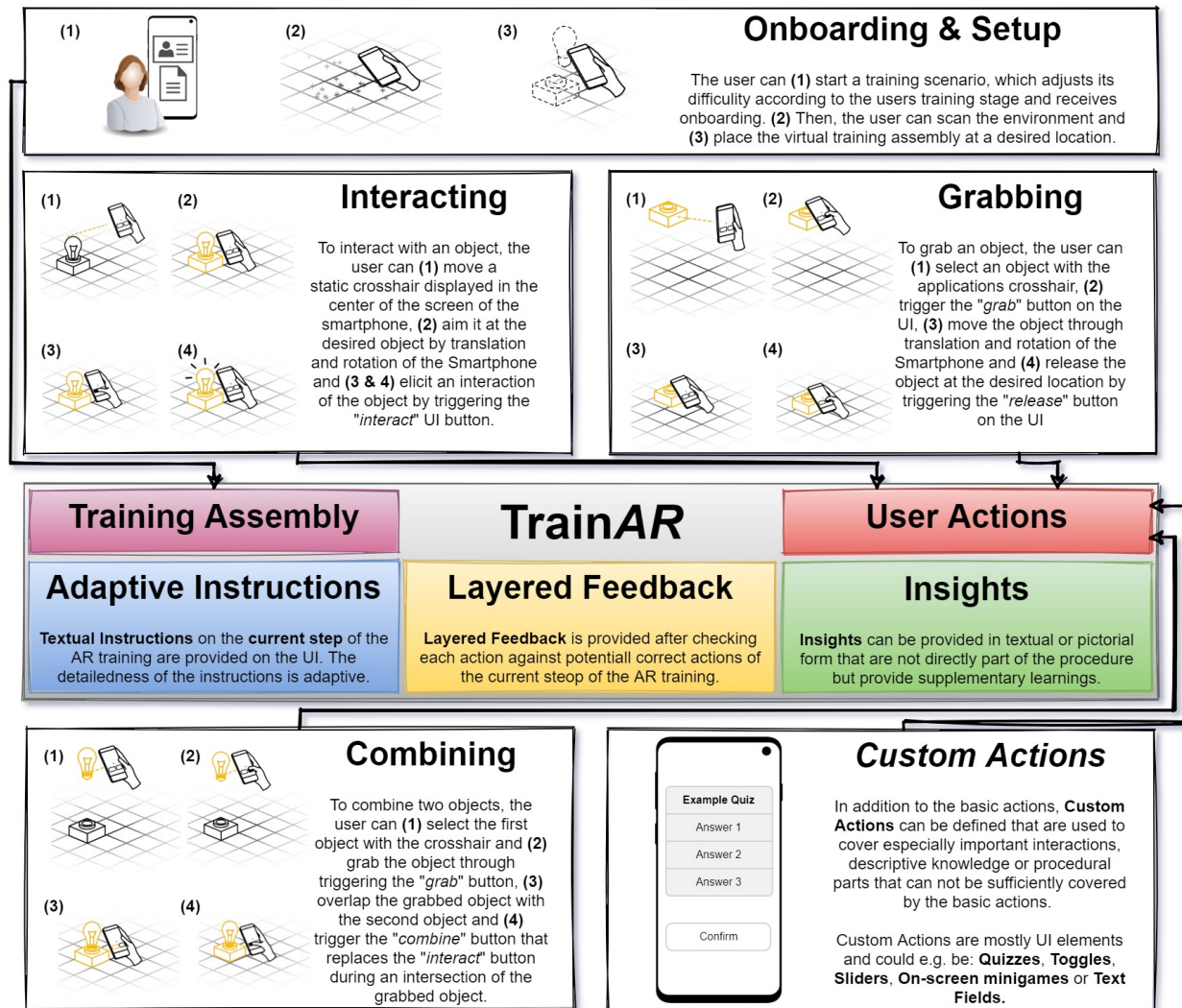
While traditional AR augments physical objects or structures with virtual computer-generated content, this interaction concept targets purely virtual procedural training through handheld AR devices. While it is true that for example in assistance scenarios a direct in-situ contextualization of instructions is beneficial [42], studies have shown that for training scenarios, tangibility has no significant effect on learning outcomes [35] but introduce limitations for the scalability and prohibit the possibility for training-at-home usage.

Therefore, the concept is deliberately purely virtual and pragmatic in both, its interaction metaphors and its design. On an abstract level, it consists of 5 interlined ideas: A **Virtual Training Assembly** representing the training setup and objects that are used for the trained tasks, **Adaptive Instructions** that are provided to the trainee, **User Actions** that are triggered by the trainee and **Layered Feedback** that provides feedback to the trainee by matching actions with instructions. Furthermore, **Insights** provide supplementary

declarative knowledge and contextual framing in relation to objects or procedures found in the training task (see Figure 1).

Two examples of AR procedural training applications derived from this interaction concept and affiliated didactic framework are developed. They are shown in Figure 2. One of them, the AR training and evaluation of preparing a tocolytic injection using the interaction concept is described in detail in Sections 4 and 5.

Onboarding & Setup

The user can **(1)** start a training scenario, which adjusts its difficulty according to the users training stage and receives onboarding. **(2)** Then, the user can scan the environment and **(3)** place the virtual training assembly at a desired location.

Interacting

To interact with an object, the user can **(1)** move a static crosshair displayed in the center of the screen of the smartphone, **(2)** aim it at the desired object by translation and rotation of the Smartphone and **(3 & 4)** elicit an interaction of the object by triggering the "*interact*" UI button.

Grabbing

To grab an object, the user can **(1)** select an object with the applications crosshair, **(2)** trigger the "*grab*" button on the UI, **(3)** move the object through translation and rotation of the Smartphone and **(4)** release the object at the desired location by triggering the "*release*" button on the UI

Training Assembly

Train*AR*

User Actions

Adaptive Instructions

Textual Instructions on the **current step** of the AR training are provided on the UI. The detailedness of the instructions is adaptive.

Layered Feedback

Layered Feedback is provided after checking each action against potentiall correct actions of the current steop of the AR training.

Insights

Insights can be provided in textual or pictorial form that are not directly part of the procedure but provide supplementary learnings.

Combining

To combine two objects, the user can **(1)** select the first object with the crosshair and **(2)** grab the object through triggering the "*grab*" button, **(3)** overlap the grabbed object with the second object and **(4)** trigger the "*combine*" button that replaces the "*interact*" button during an intersection of the grabbed object.

Custom Actions

Example Quiz
Answer 1
Answer 2
Answer 3
Confirm

In addition to the basic actions, **Custom Actions** can be defined that are used to cover especially important interactions, descriptive knowledge or procedural parts that can not be sufficiently covered by the basic actions.

Custom Actions are mostly UI elements and could e.g. be: **Quizzes**, **Toggles**, **Sliders**, **On-screen minigames** or **Text Fields**.

Figure 1. The Train*AR* interaction concept: After the **Onboarding & Setup** of the **Training Assembly**, the internal state process model provides **Adaptive Instructions** to the user of the application. Additionally it can provide **Insights** and hints if needed. The user can trigger **User Actions** with virtual objects **Interacting** with them, **Grabbing** them, **Combining** objects or trigger **Custom Actions** (e.g., quizes or minigames). The internal state process model checks those actions against an internal state catalogue, advances the procedural task and provides **Layered Feedback** to the user.

Figure 2. Train*AR* in two exemplary scenarios. The conduction of a titration experiment in the context of chemical education in project CHARMING (**left**) and the preparing a tocolytic injection in the midwifery context of project HebAR (**right**).

3.1. Virtual Training Assembly

The virtual training assembly consists of virtual 3D models of all objects relevant during the training. Besides the objects needed to complete a procedural task, this may also cover so-called distractors, objects that are needed to construct situations for decisions, as well as hidden objects in form of, e.g., trigger areas, that can be used to check if an object was placed at a specific location.

When starting a training application based on this interaction concept, users are explained the context of the training (contextual onboarding) and explained how to use the application and conduct the training (technical onboarding). They are then guided through a setup onboarding process that explains the process to establish a frame of reference by scanning for visual feature points. When completed, users can place the virtual training assembly and the training starts (see Figure 1, Setup).

3.2. Adaptive Instructions

During the training, the state process model provides continuous instructions to the trainee, detailing the next steps to complete the training tasks (see Figure 1, Adaptive Instructions). These instructions are provided through the UI, e.g., in form of text at the top of the smartphone screen. These instructions are adaptive regarding two orthogonal perspectives: Firstly, different sets of instructions can be created to support distinctive levels of difficulty, relevant to support multiple didactic contextualisation stages as well as to increase replayability. Empty sets of instructions are also supported to be used for summative training assessments or exams. Secondly, instructions can be adaptive regarding the sequence of actions chosen by the users, creating a non-linear training experience. The concept works for both, strictly linear procedures which would display instructions with specific solutions to a current step of the linear procedure but also rule-based instructions, where more than one linear path would be correct and specific actions trigger state changes and the state process model checks those against a necessary procedure list.

3.3. User Actions

To trigger actions during the training, the user of the application can use 4 basic actions provided by the interaction concept. Additionally, quizzes, sliders or toggles on the UI based on implementations of a "custom action" can be utilized to implement actions that are especially important and should be highlighted or can not be sufficiently covered by the 4 basic actions (see Figure 1).

The user can **select** and **deselect** an item by using a crosshair in the middle of the smartphone screen and aim it directly at a virtual object. The user can then **interact** with

selected objects by clicking the interaction button (see Figure 1, Interacting), triggering a state check with the state process model and corresponding feedback. Alternatively, the user can **grab** a selected object, which automatically lerps the virtual object to a position relative to the front of the smartphone while retaining a static vertical rotation towards the training assembly. This allows the user to manipulate the object's position and rotation by then releasing the object at a different location (see Figure 1, Grabbing). Positional changes can but do not necessarily have to be checked against the internal state catalogue of the state process model. Grabbed objects can also be **combined** with secondary objects by overlapping the grabbed object with the secondary one and triggering the combine button, which replaces the interact button when two objects overlap. This combining of objects is then validated against an internal state catalogue of the state process model and if allowed, the objects are combined to a single object (see Figure 1, Combining). Additionally, grabbed objects can also directly be interacted with using the interaction button.

3.4. Layered Feedback

When the user triggers actions to follow the provided instructions by the state process model, those actions when checked against an internal catalogue of potential actions, can either be ignored, correct or incorrect. Ignored actions are not processed by the state process model at all and do not elicit any feedback. This can for example be used if selection/deselection of objects or grabbing/releasing objects to move them around does not have implications in the training task the interaction concept is used in. Correct actions always trigger visual feedback, e.g., in form of a green blinking outline of the object and auditory feedback either representing the sound of the interaction itself or, if not applicable, a short sound that implies positive feedback/success. While correct interactions additionally trigger their internal event (e.g., an animation or additional visual information) and correct combinations combine the two overlapping objects, wrong interaction potentially need to elicit feedback to the user beyond simple visual and auditory error feedback. As incorrect actions can vary in severity and too much feedback could be annoying to the user, a layered feedback system is used. Here, basic interactions that are not severe only elicit the normal short error feedback comparable to the feedback of correct actions in form of a visual error symbol on the UI, a blinking red outline of the virtual object and an error sound. If an error is detrimental and should always immediately be correct or repeated errors of the same step are triggered, more intrusive feedback is given by overlaying the whole screen of the application with textual and pictorial explanations, containing hints for the user to complete the task they are struggling with.

In line with the provided instructions, feedback can be adaptive, therefore both, very behaviourist approaches are possible where wrong actions are immediately corrected but also constructivist approaches can be deployed where the user is incentivised to explore and incorrect interactions as parts of overarching procedures are not immediately prohibited but only the result is checked with the state process model.

3.5. Insights

Besides the procedure itself that is trained, there might be insights that a trainer wants to provide, that cannot be contextualised in the procedure itself or are not part of the procedural component but are rather supplementary insights or information which can not be visualized in context in physical training but could add extra learnings for the trainee. Those could for example be contextualized visualizations of declarative knowledge bits which were learned from theory or additional hints and insights from experienced trainers from practise.

4. Exemplary Implementation: Midwifery AR Training

To date, two training scenarios using the proposed interaction concept in combination with the didactic framework are in development (see Figure 2). At the time of writing this article, the implementation of a titration experiment in the context of chemical engineering

education is currently undergoing formative evaluations. However, the implementation of preparing a tocolytic injection as part of the practical training of academic midwifes in the context of project HebAR [9] is almost concluded: It is currently undergoing summative evaluation and the embedding of the AR training into an academic midwifery curricula (see Figure 2). The AR training application reported in this section describes the iteration after the conduction and subsequent improvements of all formative usability studies that are reported in Section 5.

Described in more detail in [9], midwifery education is currently transitioning towards a full academization in Germany, where midwives will soon be exclusively qualified at universities, rather than by vocational training through the dual education system. While this is an important step towards increasing the status of midwives in the medical context, it also leads to new challenges. The practical component of the training still has a high priority and this naturally leads to bottlenecks regarding available practical tutors, training space and scheduling restrictions for trainees. Preparing a tocolytic injection, an injection used for inhibiting labor contractions that is for example administered in preparation of a C-section, is a relatively basic procedural task that every midwife has to be proficient in, making it an ideal candidate for the first implementation of Train*AR*.

The implementation combines clean UI elements in healthcare-inspired color palettes with realistic high-resolution 3D models and comic/drawn stylized visualizations for conceptual contextualisation and feedback mechanisms. The clean UI elements and healthcare-inspired color palettes where chosen so they provide (mostly textual) instructions and feedback with as little distractions as possible while eliciting a sense of trust and familiarity in the user. The high-resolution 3D models where designed as realistic as possible on mobile devices to be visually recognizable as their physical counterparts. Finally, the conceptual contextualisation and feedback mechanisms like the summative assessment of the training or additional practical insights where implemented in stylized comic/drawn form to elicit some sense of play and gamification in the user while retaining the seriousness of the context and purpose of the procedure.

The procedure of preparing a tocolytic injection starts with strict hygiene procedures, preparing the workspace according to protocol and then starting the preparation of the tocolytic injection by selecting and opening all necessary material. Then, a syringe has to be connected with a needle and a carrier solution and tocolytic medication has to be drawn up in correct order and quantity. Afterwards the needle has to be disposed of according to procedure and the syringe has to be sealed using a luer lock and labelled. Afterwards, all remaining utensils have to be disposed. The virtual assembly for the training contains all objects necessary to perform this procedure and additionally several distractors, like medication that is out of date and a needle which would not be used to draw the syringe with a solution in this context.

4.1. Onboarding

When starting the AR training, the users first receive conceptual onboarding. Specifically in this case they are told that they start a shift in a midwifery ward and during a routine examination realize that the prepared tocolytic injection is expired and a new one has to be prepared (see Figure 3a) Users can then decide to receive technical onboarding, explaining how to use the application, before starting with the scenario. They can also opt-in to receive insights in form of practical know-how by an experienced midwife called "Agneta Reuter" during the training (see Figure 3b). Users then are shown 3 sets of animations and textual instructions on how to interact with virtual objects during the training (see Figure 3c–e), before they are transitioned into the AR context and instructed to scan the environment (see Figure 3f). When a sufficient amount of feature points are detected by the tracking algorithm, users can position the virtual assembly setup into the physical environment through translation and movement of the smartphone and confirm the position by an on-screen touch (see Figure 3g)
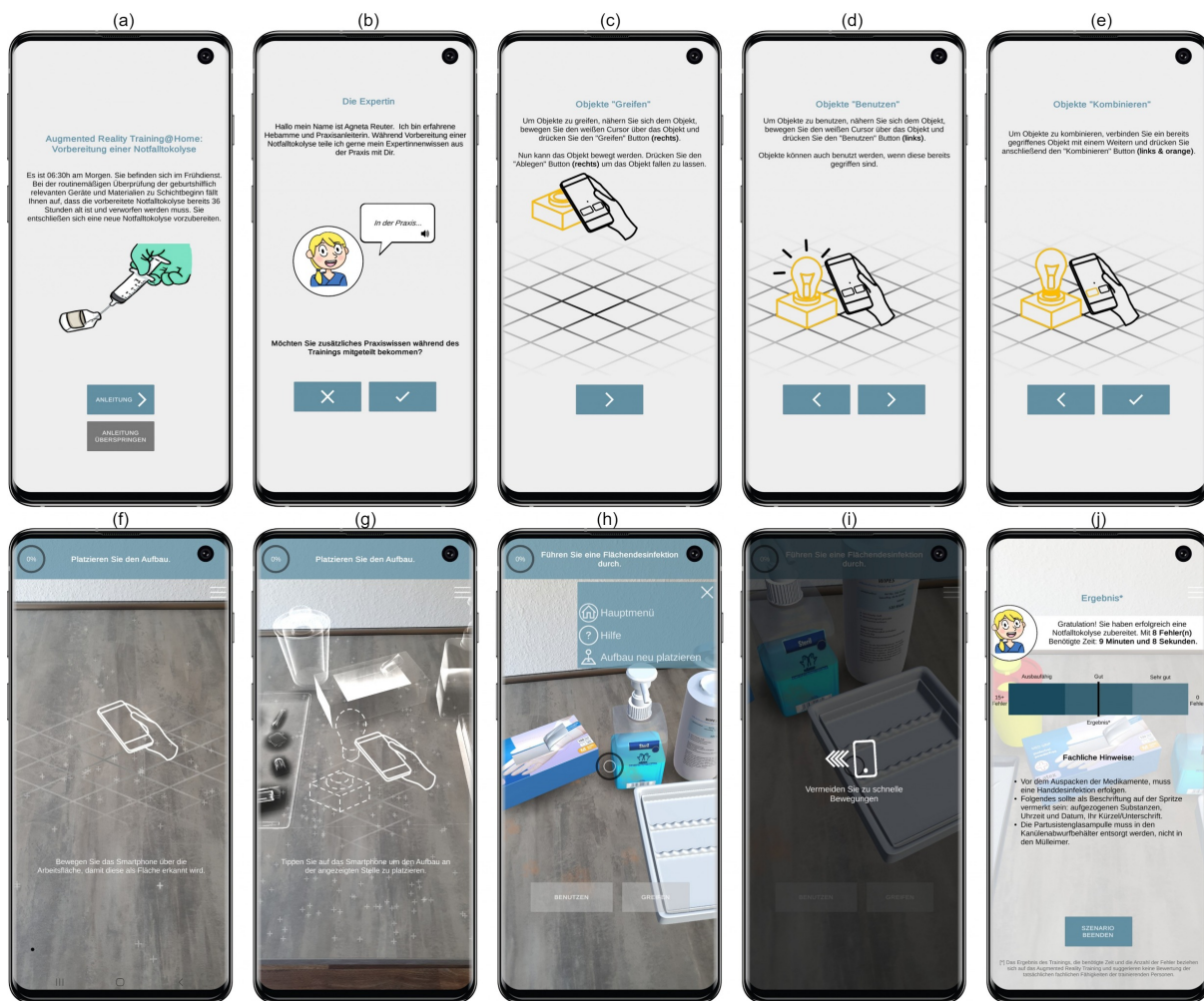
Figure 3. UI components implemented for midwifery AR training scenario. **Top**, **a–e**: The scenario and interaction onboarding. **Bottom**, **f–j**: The onboarding of placing the training assembly into the room, menus for rewatching onboarding tutorials, replacing the assembly or exiting the scenario, exemplary warning for AR tracking problems (Others warn of problems with too little illumination or insufficient feature) and the End-screen of the scenario providing contextualized performance feedback and an additional training assessment with professional feedback.

4.2. Instructions

Besides the instructions given during the onboarding (see Figure 3a–g) textual instructions are continuously provided in textual form on top of the smartphone screen. Additionally, a progression circle with a percentage number is displayed in the top left corner, showing the users' progress through the training procedure. In the current implementation 3 levels of instructions are implemented. The first one are step-by-step instructions, which guide the user through the training with explicit instructions on what to do for each step. The second one only guides the user through stages of the training, such as starting with the hygiene requirements, preparation phase and the actual preparation of the injection itself (see e.g., Figure 3h, top). The third level is to provide no initial instructions on the top UI element at all, though the progress circle, error feedback and reinforcement of correct interactions are still provided.

4.3. User Actions

The basic user actions were implemented closely following the proposed interaction concept in Section 3. A crosshair is used for the selection of objects, with visual feedback if a target object is in range (see Figure 4a,b). Selected objects have an orange outline and subtle shading to visualize the selection (see Figure 4a). Selected objects can be interacted

with, grabbed, released, and combined with other objects by pressing the corresponding buttons. Those buttons either display the name of the generic action, like "interact" or "grab", as well as object-specific derivations of that action, e.g., displaying "open" instead of "interaction" for opening packaging (see Figure 4g). If no interaction is currently possible, the buttons are greyed out (see Figure 4b). Grabbed objects are no longer outlined and shaded (see Figure 4c). If the user of the application overlaps the grabbed object with a second object, this object is outlined and shaded while the grabbed object is made transparent. Additionally, the interaction button changes to a combining button and changes its color to visualize the combining state as explicit as possible (see Figure 4d).

Figure 4. Interactions with virtual objects in the midwifery AR training scenario. (Displayed actions do not represent a sequence of events. They visualize specific functionalities described in Section 4). **Top**, **a–e**: Selection of an object with context triggered insights, no selection, a grabbed object, an object before "combining" and a scenario-specific custom action of drawing the syringe. **Bottom**, **f–j**: Positive feedback for an interaction, positive feedback with additional feedback, negative feedback for an error, an overlay for severe or repeated errors and an example of a custom action in form of a quiz.

Two custom actions were utilized in this AR training scenario. One custom action of using an UI Slider to conceptually imitate drawing up a syringe was used twice in the AR training, once for drawing up the carrier solution and then successively the medication (see Figure 4e). The other custom action was used for the labeling of the prepared injection so that users of the application do not have to type out the full label with name, date, time, carrier solution, medication and signature but the knowledge of what inscriptions are necessary can still be quizzed. (see Figure 4j).

4.4. Guidance & Feedback

Grabbing and releasing objects, combining them, or triggering their internal interaction always triggers visual feedback in form of the animated blinking outlines in either green or red on the virtual AR object itself. It also displays a success or error icon on the UI, momentarily replacing the progress bar. Additionally, all actions either play an object-specific sound, such as the ripping of packages or liquid sounds for drawing up the medication, or can play ambiguous success sounds as feedback for correct actions. Error sounds are hereby always played on incorrect actions, regardless of internal sounds present. Protruding green and red colors, not in line with the utilized color palette, were chosen to make the feedback prominent to the user (see Figure 4f,h).

Some errors in the medical context, like actions that endanger sterility or switching up the sequence of drawn up solutions, have severe implications. Subsequently, for some steps, a standard error is not sufficient and the severe layer of error feedback is provided instantly by displaying a white UI overlay, temporarily taking the users out of the scenario and focusing them on this specific feedback. This modality is also used to provide specific feedback with additional guidance if users repeatedly trigger incorrect actions, implying they need additional help (see Figure 4i).

Furthermore, some interaction, like disinfecting the hands or putting on gloves, are not exhaustively covered through basic interactions, as they would not have been implementable in a satisfactory manner and would have distracted from the core learning goals of the AR training. Therefore, they are only covered by a basic interaction on their object and an UI element informs the user that this action implicitly happened (see Figure 4g).

In the event of tracking problems, the current AR training is paused and a black screen overlay is displayed guiding users through possible steps to resolve the tracking problems, e.g., instructing them to move the handheld device more slowly, ensure sufficient light in the environment or trying to track a different surface as not enough feature points could be detected (see Figure 3i). If tracking problems persist, users can also re-position the virtual training assembly entirely, restarting the placement onboarding (see Figure 3f–h).

4.5. Professional Midwife Insights

Insights in the training scenario of preparing a tocolytic injection were implemented in form of a professional midwife called "Agneta Reuter", which provides anecdotal knowledge from practice as well as hints and contextualized advice at specific moments in the training procedure (see Figure 4a,e). When triggered, an audio file is played and a short version of the insight is displayed on the UI right under the instructions. Users can decide, if they want to use these supplementary insights at the start of the training (see Figure 3b).

4.6. Training Assessment

After the AR training is concluded, a training assessment screen is shown to users (see Figure 3j). Here, AR training specific feedback and measurements are provided, such as how fast the training was concluded and how many incorrect actions were triggered. The amount of incorrect triggered actions is also contextualized on a feedback graph to make results comparable. This graph deliberately does not use traffic light colors but rather shades of blue, to not discourage trainees, e.g., if they would be in the yellow or red in early iterations. In line with this endeavor, users are also informed that the assessment measures are AR training specific and do not imply assessment of their professional performance. Additionally, "professional notes" are provided that are displayed when specific actions were triggered which suggest that users were not following the correct procedure. e.g., this could be trying to use a carrier solution which is out of date, placing a used syringe onto the work area, or trying to throw away the medication before the syringe is labelled.

5. Formative Evaluations & Subsequent Improvements

The interaction concept was implemented for the training of preparing a tocolytic injection in the context of academic midwifery training, targeting at both Android and iOS.

For the evaluation, three formative usability studies were conducted iteratively, using the Android version on the Samsung Galaxy S10 (SM-G973F).

The focus of the first study was on gesture based interactions, as suggested by related work, and textual as well as pictorial onboarding. Other elements, such as instructions and error feedback (see Section 3) were also realized but not in focus. The second study improved upon the onboarding and introduced the training assessment at the end of the training. Additionally, it implemented an alternative, more explicit interaction concept based on buttons, subsequently referred to as the "explicit" interaction concept. The differences of the two types of interaction are visualized in Figure 5 for the combination of two objects. The third study only provided the explicit interaction concept with further improvements to the explicitness of the interaction feedback. Furthermore, improvements to the training assessment, the technical handling of AR tracking and feedback thereof, and user actions were made.

For all studies, a task-based research methodology was used, where participants were given a context in which the training task would have to be completed and were encouraged to "think aloud" during the experiment. Participants did not receive external help during the experiment. After completing the task, participants were asked to fill out a System Usability Scale (SUS) questionnaire [43], a user experience questionnaire (UEQ) [44] and a qualitative questionnaire, asking what participants liked or did not like about the application, what they had problems with and additionally gave the opportunity for further feedback or remarks.

5.1. Participants

Overall, 24 participants (16 unique participants), aged between 21 and 46, with an average age of 28.75 (SD = 6.16), took part in the studies. All participants were either midwifery or nursing students that were familiar with the preparation of a tocolytic injection. 15 out of 16 participants were female. To gather both iterative feedback across the developed versions, including increased familiarity with the application, as well as fresh feedback and "first impressions", some participants were deliberately invited to multiple studies, while others only conducted the experiment once. In the first study 6 students participated, in the second 10 participated with 5 per condition for the explicit and implicit interaction concept and in the third study 7 students participated. Across the studies, 9 participants took part in one, 5 participants in two (3 participated in 1 & 2, the other 2 in 1 & 3) and the remaining two participants took part in all 3 studies.
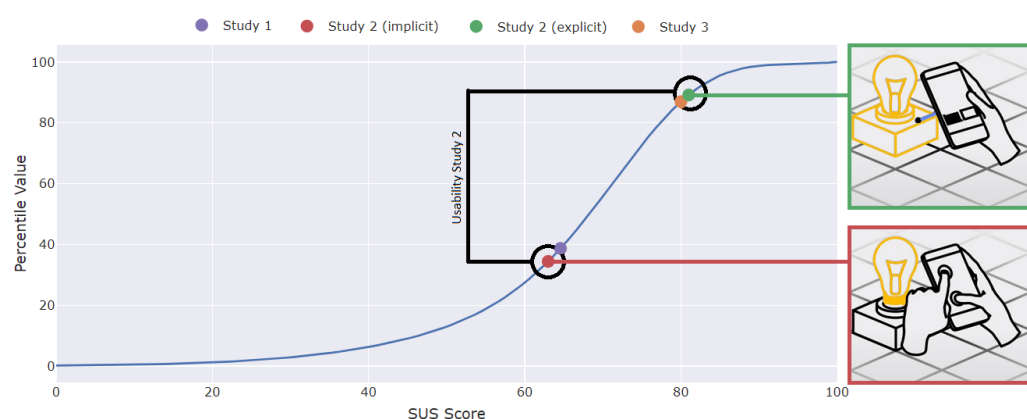
5.2. Results

5.2.1. Usability

Regarding usability, participants reported an average SUS score of 64.58 (SD = 7.81) in the first study. In the second study, participants reported an average SUS score of 63 (SD = 8.91) for the implicit interaction concept and an average SUS score of 81 (SD = 3.35) for the explicit interaction concept. This difference is highly significant according to an independent-samples *t*-test, t(8) = 4.2283, p = 0.0029. For the final study, improving upon the explicit interaction concept, participants reported an average SUS score of 80 (SD = 7.91). According to Bangor et al. [45], SUS scores of 63 and 64 would be considered "ok" and thus represent a (low) marginally acceptable usability, while SUS sores of 80 and 81 are considered to be "good" to almost "excellent" and imply acceptable usability.

It has to be noted that those SUS scores are not conclusive with the number of participants used in the separate usability studies according to Tullis et al. [46], though the sample of participants was fairly homogeneous and the observable variance is small (see Figure 6).

Notably, in line with the average SUS scores, the two participants who took part in all three usability studies both reported a SUS score of 72.5 in the first study. The participant in the "implicit" interaction concept condition in the second study reported a SUS score of

70, the participant in the "explicit" interaction concept condition a SUS score 82.5. In the third usability study, the participants reported SUS scores of 82.5 and 90 respectively.

Figure 5. The SUS results for both the explicit and implicit interaction concept from usability study 2 contextualised on the percentile curve of SUS scores according to Kortum et al. [47] (**left**) and an example of combining a grabbed object with another object through both interaction concepts (**right**).
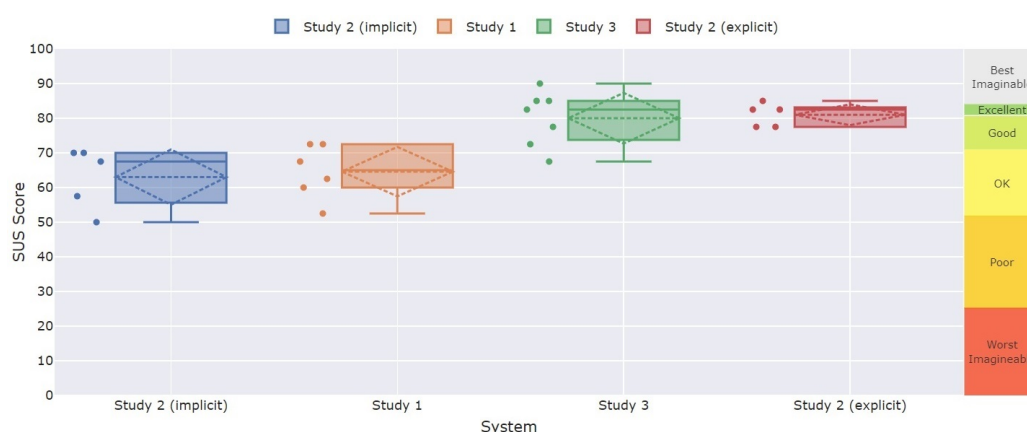

Figure 6. System Usability Scale results contextualized on the adjective interpretation scale according to Kortum et al. [47], sorted from lowest to highest average score. Datapoints are represented as dots, mean values as solid lines and average values with their standard deviation as dotted lines.

5.2.2. User Experience

For the user experience, the reported UEQ results were analyzed for the 6 measures: attractiveness, perspicuity, efficiency, dependability, stimulation and novelty using the UEQ benchmark, which contextualized the measured scale means in relation to a benchmark data set of over 450 UEQ studies [48]. In the first study, participants reported an average attractiveness score of 1.25. This would be considered "Above average" in the UEQ benchmark. Regarding perspicuity, they reported a score of −0.13, which would be considered "Bad". The average efficiency score of 0.75 indicated a "Below Average" perceived efficiency of the tool and a dependability score of 0.96 a "Below Average" dependence. The stimulation score of 2.04 and a novelty score of 2.38 would both be considered "Excellent" compared to existing values of the benchmark data set. In the second study, participants reported an average attractiveness score of 2.40 (Excellent) for the explicit interaction concept, 1.20 (Above average) for the implicit interaction concept and an average perspicuity score of 1.55 (Above Average) and perspicuity 1.10 (Below Average) respectively. In terms of efficiency, participants reported an average score of 1.65 (Good) for the explicit interaction concept and an average score of 1.25 (Above Average) for the implicit interaction concept. For the dependability, participants reported an average score of 1.80 (Excellent) for the

explicit and 0.82 (Below Average) for the implicit interaction concept. Participants reported an average stimulation score of 2.50 (Excellent) for the explicit interaction concept and an average stimulation score of 1.70 (Good) for the implicit interaction concept. For both conditions, an Excellent average novelty score was reported with 2.35 for the explicit interaction concept and 2.10 for the implicit interaction concept. In the third study, participants reported an average attractiveness score of 1.93 (Excellent), perspicuity score of 1.36 (Above Average), efficiency score of 1.21 (Above Average), dependability score of 1.14 (Below Average), stimulation score of 2.04 (Excellent) and a novelty score 2.39 (Excellent) (see Figure 7).
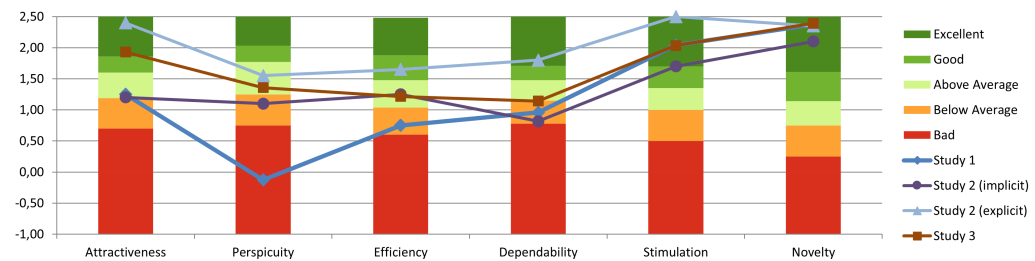
Figure 7. User Experience Questionnaire scores of all 3 usability studies in relation to the UEQ benchmark data set.

5.2.3. Qualitative Feedback

Qualitative feedback provided through qualitative questionnaires, observations, verbally during the experiment or implicitly provided through the "think aloud" methodology, were transcribed, prepared and inductively coded according to Linnenberg et al. [49]. The qualitative questionnaires consisted of 4 questions: What participants liked about the application, what they did not like, what they had problems with during the training and what additional feedback or remarks they wanted to provide. While the combined qualitative feedback was fully utilized for the design-based research process and iterative improvements to the application and Train*AR*, they are only reported in very condensed form here and filtered for feedback targeting the interaction concept.

Across studies, participants noted that they liked the *"comprehensible" "step-by-step instructions"*, the continuous feedback provided after actions and the verifiable progress of the training task. They noted that they liked the color scheme and clean design, especially the *"details"* and *"realistic graphics"*, underlining the fact that the virtual objects are *"recognizable"* as their physical counterparts. Additionally, they perceived the application as a *"promising new type of learning"* and enjoyed the gamification aspects of training in AR. Some participants also noted across studies that they sometimes had problems with the tracking and that the virtual assembly sometimes shifted out of place or was temporarily not visible, though this feedback decreased in later studies. Participants also noted that text was sometimes too small for them to read.

In the first study, participants noted that the provided onboarding based on textual instructions and pictures was not sufficient and should be repeatable. They perceived the interaction with objects as *"cumbersome"*, especially for the feedback regarding the process of combining two objects, with all participants providing qualitative feedback indicating they struggled with this interaction. Moreover, some participants struggled to understand the spatial component and distances of objects.

For the implicit interaction concept in the second study, participants who also participated in the first usability study provided feedback indicating that the interaction concept and especially the onboarding somewhat improved. In contrast, the qualitative feedback provided by new participants indicated similar perceptions to the first usability study, still describing the interactions as *"complicated"*, *"abstract"* and *"frustrating"* and especially again noting combining objects as an obstacle. Some suggested a *"trial"* scenario where the interaction could be tested. For the explicit interaction concept, all participants who took

part in the first usability study provided feedback that the interaction *"drastically improved"* and that the usage was *"less frustrating"* as it provided *"more feedback"*. This sentiment was shared by the participants who used the application for the first time, describing the instruction handling of the application as *"clear"*. In both conditions, participants noted that they liked the training assessment at the end of the scenario, though noting that they believe that high error counts might be discouraging for some users.

In the third study, participants especially liked the improved training assessment, now also explicitly stating what kind of professional errors were made. Participants who only conducted the first study or the implicit interaction condition in the second study provided feedback similar to the explicit condition in the second study, indicating *"improved"* handling and onboarding. Also, some participants stated that they think the application is somewhat *"strict"* in regards to what procedures would be correct.

5.3. Subsequent Improvements

Beside many midwifery context-specific adjustments and changes regarding the state flow of the training across the three formative usability studies, the most important implications and subsequent improvements to the Train*AR* interaction concept are as follows:

In the first study, an interaction concept based on on-screen gestures for all basic actions described in Section 3 was developed, e.g., using a short press for an interaction, a long press for grabbing & releasing objects and a combined long press with a short press while overlapping two virtual objects for combining objects. Contrary to the results suggested by the literature reported in Section 2.2 and our expectations, at least in the context of academic midwifes, those prior findings could not be replicated. Even with the improved onboarding based on textual instructions combined with explanatory animations in the second study, participants struggled to effectively utilize the interaction concept. While the perceived perspicuity did drastically improve in the second study, most likely due to the improved onboarding, it was still below average and lower than the perceived perspicuity of the explicit interaction concept. Additionally, the overall usability of this condition in the second study did not improve compared to the first study, but the usability of the newly introduced explicit interaction concept was significantly higher. When contextualizing all three studies on a percentile curve of SUS scores gathered in a meta analysis by Kortum et al. [47], this difference becomes even more apparent, clearly visualizing two groups of usability scores for both interaction concepts across the usability studies (see Figure 5). This was further affirmed by singling out the participants who took part in all studies, the qualitative feedback by all participants indicating that they would need more onboarding or even a trial scenario using the implicit concept, before starting the actual training scenario and the repeatedly noted frustration. Neither was similar qualitative feedback reported in the questionnaires, nor observable for the explicit interaction concept during the second or the third study.

During the first usability study, it was possible to select objects from any distance. It was observable that participants did not utilize the translation and rotation of the device itself effectively, some even voicing the need for "zooming" to better red displayed text in the context. Subsequently, a maximum range at which objects would be selected was introduced and the crosshair was improved, so the two circles would converge when close to the distance at which an interaction would be possible (see Figure 4a,b). This improved the observable utilization of the device translation/rotation as part of the interaction in the subsequent studies.

Partially independent of the interaction conditions (implicit or explicit), two additional trends emerged throughout the studies. Explicitness and deliberate redundancies of interaction visualisations and feedback mechanisms improved the users' perceived attractiveness and efficiency of the AR application and was particularly reported as positive through qualitative feedback. Subsequent improvements, especially for the improvement of explicitness in the third study therefore comprised of not only outlining a selected object, but also slightly coloring it in the selection color using a shader (see Figure 4a), no longer

outlining objects when they are grabbed (see Figure 4c) and, for the state of combining, making the grabbed object transparent while outlining the object to be combined with (see Figure 4d). Additionally, the buttons used in the explicit interaction concept were only displayed when an object is selected, grabbed or in a combining state when they are usable and also depicted the specific interaction that would be triggered. The redundancy of feedback mechanisms was perceived positively, therefore correct or incorrect interactions elicit a visual feedback on the UI, visual feedback through blinking outlines in the AR context itself and auditory feedback.

In the third study, spatially contextualised speech bubbles were introduced to communicate implicitly triggered interactions that are not actually performed in the AR training, like the disinfecting of the hands or the insights provided by the professional midwife. As observations, qualitative feedback and the higher variance of reported usability scores indicated this could potentially be overwhelming for some users, those speech bubbles were subsequently also transitioned into UI elements (see Figure 3a,e,g).

6. Didactic Framework

In modern education theories, the focus is on problem-based and therefore learner-centred learning settings that enable both individual and collaborative learning [50]. The aim is to promote the development of complex technical and practical knowledge as well as professional competence. Action and work process orientation, which represent central concepts of vocational pedagogy [51], are suitable for this purpose and are well compatible with the interaction concept provided through Train*AR*. Action and work-process orientation find methodological expression in the *complete action* [52]. The acquisition of competences takes place through repeated runs of application-oriented phases: behavior of the learner/actor, feedback and evaluation of the actions with renewed goal setting [53], which corresponds to the phases of complete action: 1. informing, 2. planning, 3. deciding, 4. executing, 5. controlling, 6. evaluating [52]. For the specification, conception, and development of work process-oriented AR teaching/learning scenarios, correspondingly detailed descriptions of the work processes including necessary decisions and information flows are required. Referring to Howe et al. [54], subject-specific methods for collecting and describing information flows are developed. Based on this, authentic, complex problems are used as a starting point for work process-oriented knowledge acquisition from the above-mentioned subject areas and diverse AR learning scenarios are derived according to the competence goals. For learning and transfer effects, one of the central concepts is to create suitable occasions for reflection and to support them with learning guides. In the practical design of Train*AR*, the minimalism dimension according to Drljević et al. [55] is taken into account, so that only the necessary information is provided. This avoids stimulus overload and supports focusing on the procedural flow.

The intention is to systematically put knowledge into practice. For this, the assumption that a person is enabled to act independently and responsibly is pursued. Train*AR*'s training scenarios are therefore based on work process descriptions and competence-oriented learning objectives, where the students' learning conditions, preexisting experiences, and knowledge are considered [14].

6.1. Training Contextualisation & Structure

Train*AR* as a training application focuses on the teaching of intellectual skills and cognitive strategies, according to instructional design theory as proposed by Gagne [7]. Therefore, first verbal information and declarative knowledge is taught through traditional class-based teaching or in self study. Afterwards, the procedural knowledge, combining intellectual skills and cognitive strategy, can be trained using Train*AR*, but motor skills are not reinforced at this point. In this case, Train*AR* serves as a pre-training and motor skills will be trained with physical material in the practical training settings (e.g., SkillsLabs in the clinical setting). Before applying learned procedures in practise or as a reinforcement

of best practices and attitudes, Train*AR* can also be applied as a retention training after the physical on-site trainings (see Figure 8).
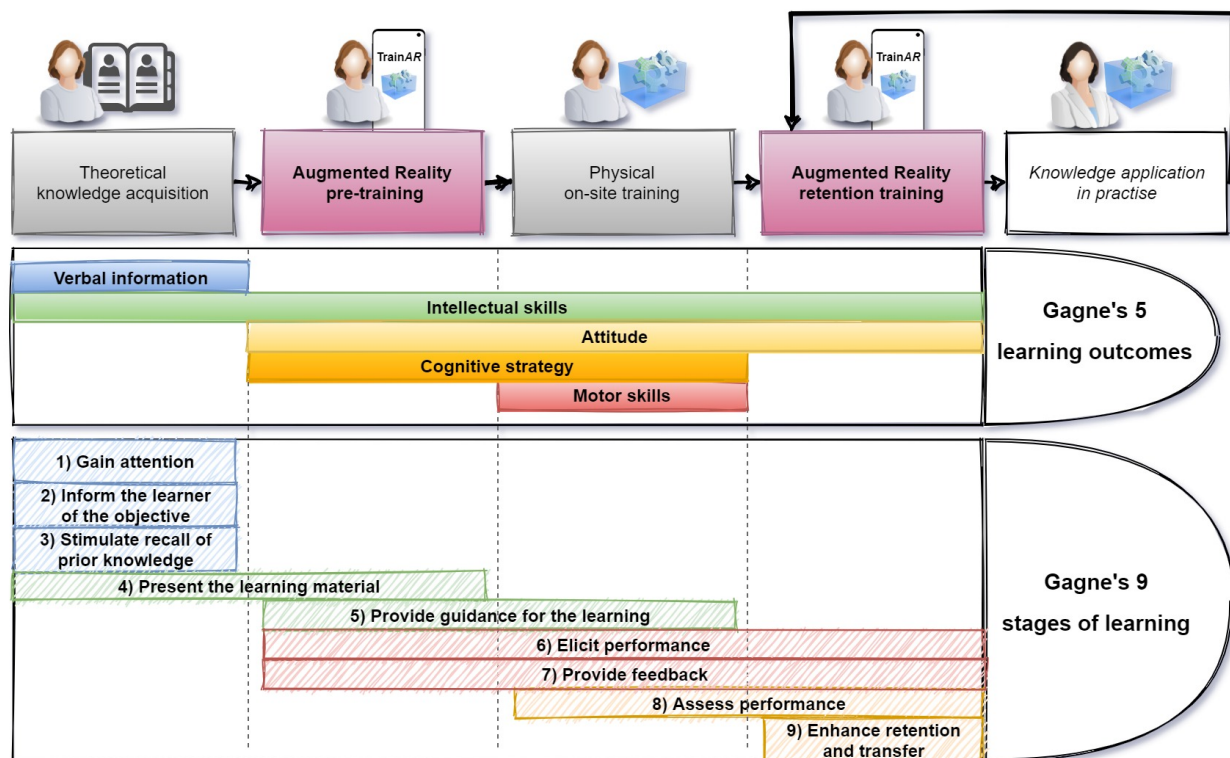
Figure 8. Possibilities for the curricular embedding of Train*AR* as a pre-traing to practical on-site training, as a retention training after the practical training or a combination of both, contextualised with instructional design theory utilizing Gagne's 5 learning outcomes and 9 events of learning [6,7].

During each Train*AR* session, the training starts with a short case description according to the principle of problem-based/learner-centred learning [50]. The trainings always run based on a specific case, therefore contextualizing the procedural knowledge taught, as described in Sections 3.1 and 4.1. The aim is to link academic theory and practical competences. During the training, expert knowledge is available in contextualized form (see Sections 3.5 and 4.5) and after completing the training, the students will receive an assessment of their training performance and professional feedback (see Section 4.6).

6.2. Integration in Curricular Teaching

Utilizing Train*AR* in the course of the curriculum, the teacher transitions from a lecturer to a tutor and (partially) gives up control and steering of the students learning activities. Instead they offer support and guidance. This is intended, among other things, to support the empowerment of the students [55]. The AR training can be used at different stages in the course of study. This is achieved through the adaptive instructions providing difficulty settings for the same training procedure (see Section 3.2). The first mode, known as *guidance*, does not require any prior experience. Above all, the intellectual skills and cognitive strategy associated with a procedural task are trained here. The students are introduced step-by-step to the procedure, following a primarily behaviouristic approach, as described in Sections 3.4 and 4.4. The second mode is the *training* mode, in which prior knowledge of the subject is required. The aim is to consolidate the process and elicit reinforcement of prior knowledge. Different courses of action can be followed. Here, the cognitivist approach is followed, taking into account the cause and effect mechanisms, including the learning process. Therefore, the focus is primarily on methods of knowledge transfer in the first place (competence transfer of procedural learning).

Expert knowledge is integrated in both modes and linked to actions or objects (see Sections 3.5 and 4.5). This knowledge is reproduced auditorily and visually. Students receive real-time feedback after each session as described in Section 4.6. In form of a point scale, the students can rank/rate their performance and feedback is also given in written form. Hence, both positive and negative aspects are highlighted according to the mastery principle, so that the students receive confirmation of their success, but also information about their mistakes or suggestions for improvement. The provision of real-time feedback has a positive influence on the motivation of the students, as it can support the comparison with the individual learning success [56]. The choice of learning environment is very open, so that the AR application can be used anywhere, e.g., at home, in the skills lab, or in the classroom, especially enabling BYOD approaches where trainees can use their own smartphones for the AR training. There is generally a need for flexibility in the educational process; The chosen flexibility dimension also makes it possible to carry out the training outside of the curricular integration, regardless of location and time, to consolidate the procedural flow. For example, before or during a practical study phase [55].

Figure 8 shows an exemplary curricular integration envisioned for the curricular integration of Train*AR*, contextualized with the five learning outcomes and nine stages of learning proposed by Gagne [6,7]. In the first step, the theoretical framework is dealt with in the context of classical forms of teaching, such as lectures and seminars. Here, the learners attention is gained, the learner is informed of the objective, the learning is contextualized in prior learning and the procedural task is presented. In this stage, primarily verbal information, therefore declarative knowledge of the procedural task, with some intellectual skills, such as broad concepts, are introduced. As a second step, the AR-supported procedural training using Train*AR* takes place as a pre-training. Therefore, the *guidance* mode offers the behaviouristic support during this training. The students have the opportunity to understand the process at their own pace but are strictly guided. Students are presented the learning material, are provided guidance and are given the opportunity to elicit the performance and receive feedback from the application. In this stage, the intellectual skills are trained in combination with their corresponding cognitive strategies. In the third step, motor skills are practiced and consolidated in practical on-site trainings (e.g., SkillsLabs in the medical setting) and the performance of the learner can be assessed. Here, students already know the entire sequence and develop a cognitive strategy to solve it, that can then be linked to the motor actions required. Finally, the AR retention training is envisioned as a *training* mode, that helps students in consolidating the sequence of their actions. The students can carry out the action more freely, compared to the pre-training, and also consider new action alternatives with AR support. Additionally, it can be used for self-directed knowledge verification, not only assessing performance but also enhancing retention and transfer.

6.3. Applying TrainAR to Procedural Training Tasks

In many vocational settings, it is important to train procedural courses of action as precisely as possible as errors in the procedure can have devastating effects. Especially in medical and health science, where standardized procedural trainings are taught regularly and their correct application is especially important, methods were developed to transform procedural knowledge from practise into controllable and verifiable training settings. Derived from those methods and applied for the exemplary implementation of Train*AR* described in Section 4, but also applicable outside of the medical scope, we propose that scenarios utilizing Train*AR* should be developed by: *Identifying & observing the procedure*, *analysing & deriving the work-process-description*, *defining the competency-based learning objectives*, and *transforming the didactic considerations towards an AR application* utilizing Train*AR*'s interaction concepts. (See Figure 9).
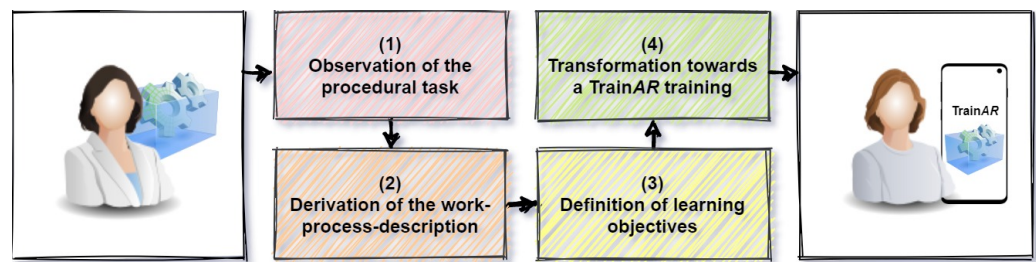
Figure 9. The transformation of a procedural action sequence into a Train*AR* training scenario.

Hereby, this procedure is envisioned as systematic and strictly sequential, condensing but still largely following the classic instructional design model by Dick et al. [57] that defined the necessary steps for the development of training instructions as a 10 step process: First, the teaching objectives have to be determined (1). Following this, teaching material and learning processes (2) as well as previous knowledge should be analyzed and determined (3). Then, criteria for learning success (4) and test items (5) have to be developed. Afterwards, the instruction strategy is defined (6), which includes the didactic method, exercises and feedback. The teaching material can then be selected and produced (7) and formative evaluations can be planned and carried out (8). Finally, the learning offer is revised (9) and summative evaluations are planned and carried out (10).

6.3.1. Identifying & Observing the Procedural Task

As a central concept of design-oriented media didactics according to Kerres [58], media sources should be utilized as a contribution towards solving an educational problem and not applied without specific cause. While new media sources fundamentally open up new opportunities and have potential for different types of learning, this is not based on an inherent effect of increased learning success. They require dedicated planning and conception in order to be able to induce benefits [58]. This includes AR training scenarios. Train*AR* scenarios should be therefore carefully identified based on their suitability for training in AR. What procedural AR trainings are suitable is dependend on the complexity and contingency of the educational field, but generally procedures that combine declarative knowledge with complex cognitive strategies are ideal. While procedures with significant amounts of motor skills are possible, as shown in Section 4, motor-learning components of the procedure itself have to be training in physical on-site trainings and can not be trained using Train*AR* autonomously (See Figure 8).

After a suitable procedural training task is identified, the training task, demonstrated by a domain expert, should be systematically observed and ideally videographed. Recording does not only allow preservation of the initial observation and expert input but also serves as a basis for the development of the work-process-description.

6.3.2. Analysing & Deriving the Work-Process-Description

When the selected procedure is observed and documented, it should be converted into a work-process-description as described in [9,59]. This should be developed towards a work-process-model, describing each possible step and action of the procedure and their interconnections. Therefore, while the work-process-description only describes the procedure as observed, the work-process-model also forces a decision about which measures have to be taken after each step. In Section 3, this is refereed to as the *state process model* from a technical perspective. In such work-process-models, a distinction is traditionally made between input, work sequences and output. Here, task instructions are the input, which are given to the trainee, including distractors and deliberate disturbances and interruptions in the course of action. The model should be derived by starting with an initially stringent, linear, idealistic action sequence and then alternative, further sequences can be added. The results are then included in the output. This means that all the necessary information from

the documented work process descriptions is in the process model and can be used for further design developments.

6.3.3. Definition of Competency-Based Learning Objectives

After the work process has been described, the definition of the competency-based learning objectives can be carried out. For this purpose, the cognitive and psychomotor learning goals are derived from both the work-process-description and the work-process-model. Those should primarily be based on taxonomy levels according to Bloom [60] and clinical competence levels according to Miller's pyramid of clinical assessment [61]. These established educational frameworks include learning objectives as well as assessment measures. Bloom's taxonomy is well established for lesson planning, design, assessment and evaluation. Bloom divided the learning levels into cognitive, psychomotor and effective areas, which are independent but mutually influence one another. In the Miller pyramid, the learning process is divided into four levels. Knowledge is the basis and routine application, especially in clinical environments, is the top priority.

To achieve this, first target group analysis should be carried out, e.g., in form of Personas [58]. This includes framing conditions such as the intended curricular integration, localization of the application and previous knowledge of the learners [57]. The previous knowledge of the learners in particular gives an important and decisive direction both in the formulation of learning objectives and in the later technical application development. The work process model should then be divided into sections and formulated in constant comparison with the prior knowledge of the learner's learning content. In order to be able to formulate learning objectives, cognitive and psychometric taxonomy levels are assigned to the learning content [8]. Here, verbs should be assigned to each taxonomy level, to formulate learning objectives precisely. Based on these taxonomy levels, the assignment to the Miller [61] pyramid levels can be made.

6.3.4. Transformation towards a Train*AR* Training

When completing the classification of the learning objectives and competence levels, the transformation towards a Train*AR* training scenario can be carried out utilizing the "mobile augmented reality education design frameworks" (MARE) [62]. The MARE-Model is a developed outcome layer that combines the Miller pyramid and the Bloom taxonomy levels. It contains these differentiated dimensions of learning and enables a transfer to AR learning activities via these classifications. The general requirements for AR learning activities describes by Zhu et al. [62] are predefined by the usage of the Train*AR* features described in Section 3. Based on those general requirements, scenario-specific AR requirements should be formulated. Depending on the taxonomy level, different approaches can be utilized: Should trainees be given an explanation for the procedure, should they carry them out independently or is a combination necessary? (See Figure 8) In addition, scenario- and location-specific AR implementation recommendations could be worked out on the basis of an AR property overview [63]. Since the scenarios are usually very complex and detailed analyzes have taken place in advance, it might be helpful to take a step back and look objectively at the combination of the state-process-model and learning objectives and go through the scenario step by step and consider which AR properties were utilized effectively.

The MARE design framework is a learning theory that serves as a guide for developing AR apps for educational purposes. Primarily aimed at educational AR apps in the medical context but arguably applicable beyond that scope, it was constructed using a conceptual framework analysis method in which Zhu et al. [62] identify interconnected key concepts. In an iterative process, they discovered three main elements: (1) Foundation, (2) Function and (3) Outcome. Learning theories form the basis (1), as they are elementary for the form of teaching content. Zhu et al. [62] selected situated-, experiential- and transformative learning theories for the foundation. The situated learning offers learners a real-life-environment of learning and interaction. Experiential learning combines experience and behavior, e.g.,

in a virtual learning environment in which feeling, thinking, observing and acting are the focus. Transformative learning involves critical reflection and transformation in meaning and perspective. The focus here is on changing problematic frames of reference. The Foundation (1) and the Outcome (3) layer support the design aim. The Outcome layer comprises learning objectives as well as expected skills of the learner and assessment of the learning. These elements are helpful in finding out which skills may be achieved utilizing MARE. For the transfer of learning objectives into AR trainings, the outcome layer offers a basis that provides orientation for implementation. This also includes Bloom's taxonomy levels, which are well known for conventional lesson planning. If there is not yet routine in the definition of learning objectives, it might be challenging to derive them. In this case, we suggest to include the outcome layer in the definition of learning objectives, as it is immediately visible which levels contain which activity, making it more practical. The Function (2) layer includes how learning can be achieved with the following levels: learner's personal paradigm, learning activities, learning environment and also learning assets [62].

Train*AR* is primarily developed with the theory of experiential learning as one of the central concepts. The learning theories and the procedure for the application of Train*AR* for a training task presented in this section do not necessarily have to be selected. Alternatively, also more constructivist planning models like the R2D2 model by Willis [64] would be conceivable as a basis for further scenario development. However, the learning and instructional design theories largely determined the design of the interaction concept and the presented application procedure provides a clear, didactically reasoned approach for the development of additional AR trainings using Train*AR*.

7. Discussion

Based on studies on the acceptance of AR-based training using TAM, which rely on intend to use or perceived usefulness, there was a high expectation that AR-based training would work for academic education. However, the clientele of midwifery students is rather specific, as they typically cover a broad range of ages, and the focus of the curriculum is far away from engineering and computer science. AR, on the other hand, is a technology so new to many people, that the actual experience is much different to what can be imagined, leading to problems with AR-specific technologies [65]. Any learning technology, however, can only prevail, if the use of the technology does not interfere with the learning. The presented work shows, that it is indeed possible to create usable AR-based trainings, if they are carefully designed and the disruption of known interaction concepts is kept at a minimum.

7.1. Scalability of TrainAR

For a general acceptance and implementation of AR-based trainings in academic and vocational training, scalable solutions, as detailed in the introduction, are an essential requirement. Only then, if hardware availability and management is not a problem and if teachers and students can focus on the content to be learned and not the mediating technology, this approach will be generally affordable and ready to scale up. And only if this technology provides added values, such as self-regulated learning, place and time independence or a reduced resource consumption in terms of rooms for practical training (e.g., laboratories or Skill Labs), human tutors or consumables, the technology will be ready to provide the proposed opportunities and support to advance the craft of teachers [1].

The presented work focuses on the first aspect, namely the scalability of AR-based procedural trainings. Follow-up work will address the second aspect, as will be detailed below. The presented framework in particular addresses the challenges of usability and approachability under varying media competences.

7.2. Towards a Usable Interaction Concept

Train*AR* was designed to address people with little to no media competences regarding AR technology. This is in particular reflected in the Virtual Training Assembly module, with its verbose onboarding and setup procedures, but also in the reduced design of the 3+1 action types (primary + custom) that are supported. The combination of Adaptive Instructions and Layered Feedback supports a high-level of self-description, which is good practice for the design of dialogues, in particular in this case, where the state and state-changes of the simulated environment, as well as the transformation of the users' interaction requests into actions applied to the simulated environment are relevant for the trainees to develop situation awareness.

The iterative evaluation presented in Section 5 shows, that Train*AR* successfully accomplishes these design goals. This is supported by an exceptional SUS score of 81 and above average perspicuity as well as efficiency scores in the UEQ. Beyond that, exceptional ratings for attractiveness, stimulation, and novelty underline the potential of AR technologies, which typically have at least short-term effects on motivation, yet long-term sustainability of these effects has still to be shown.

As a supplementary contribution, it was shown that a modern interaction concept tailored to use on-screen gestures, as suggested by related work, could not be effectively utilized in this context. The reworked interaction concept uses buttons, which significantly increased usability and observable performance of AR trainees. Furthermore, explicitness and redundancy of instructions as well as feedback modalities were identified as success factors through the evaluations in the midwifery context.

The comparatively low number of participants tested in the studies (n = 24) is due to the fact that academic midwifery has only been offered as a model university course in Germany at HSG and just recently in 2020 been opened for all universities, but courses at other universities are only starting. The research program, however, is designed to iteratively assess the developed AR trainings within each new cohort of students.

7.3. Opportunities and Challenges of the Didactic Framework

The shown didactic framework enables a differentiated development of the teaching and learning levels. The constant change of perspective from very detailed work steps and learning objectives to abstract AR learning activities allows to develop specific AR implementation recommendations. The MARE outcome layer enables an abstract AR implementation framework that is helpful in formulating general requirements for AR. In order to create concrete implementation ideas, it turned out to be expedient to go through the work process meticulously with an AR properties list. For the scenario "preparation of emergency tocolysis", learning objectives were defined, transferred to AR and tested. Basic methodological peculiarities of the AR application had to be taken into account, such as one-hand interactions and spatial restrictions due to visibility concerns caused by the small field of view of AR handheld devices. This led to repeated adjustments to the implementation of the learning objectives. Factors influencing the interaction design were determined. This includes learning objectives that develop motor skills or procedural knowledge, as well as different application locations and availability of materials, risks of injury or interactions with others.

8. Conclusions

A scalable interaction concept in combination with an accompanying didactic framework called Train*AR* is proposed. It is first described on an abstract level and then detailed and evaluated using an implemented example for academic midwifery. The didactic framework contextualizes the didactic ideas of Train*AR* with learning theory and provides guidance for the development of additional scenarios. In the evaluation, Train*AR* was generally well received and the defined criteria for scalability are met. Especially explicitness, redundant feedback mechanisms, detailed onboarding and usability with lower levels of media competencies are identified as major success factors.

Limitations & Future Work

The formative evaluations show that the requirements of scalability could be largely met with Train*AR* in the context of academic midwifery. Subsequently, several follow-up questions emerge: Do procedural AR trainings based on Train*AR* also elicit retention benefit, increased motivation and improved academic achievements compared to traditional and other AR-based approaches? How applicable is Train*AR* in new contexts with different training procedures and requirements, both in terms of the interaction metaphors but also in terms of didactic considerations? And finally, how can the process of creating such procedural AR trainings using Train*AR* realistically be authored by trainers themselves? This arguably is an important question on the path to realistic scalability from the perspective of institutions and trainers, which is not discussed in this paper.

In the context of the project HebAR [9], the application described in Section 4 is currently undergoing summative evaluation and embedding into the curricula of midwifery teaching at the faculty of "Midwifery & Reproductive Health" at the "Hochschule für Gesundheit Bochum". As part of this curricular testing of the AR training, questions on students' perception, acceptability and academic achievements are explored by providing the AR training for an intervention group and comparing them against a control group not using procedural AR trainings.

The transfer of Train*AR* to other domains, both in terms of interaction metaphors but also the didactic considerations, is currently ongoing work in cooperation with the project CHARMING [66]. This EU-funded project explores the usage of AR and virtual reality technology for chemical engineering education at the levels of pupils, students and employees. Future work will focus on subsequent insights gained from this specific context but also explore additional new cases. Other researchers, when applying Train*AR* in their settings for procedural trainings, should report on the applicability of the interaction concept and didactic considerations for their specific settings and contexts.

Finally, Train*AR* in its current form is developed using the game engine Unity as a "low level programming framework" when contextualized on the Augmented Reality Authoring Taxonomy proposed by Hampshire et al. [67]. Therefore, developers can simply utilize Train*AR*s features described in Section 3 as pre-existing components to develop their own AR training scenarios. However, programming skills are still required, e.g., to implement the process model, UI elements and custom actions. We are currently in the process of open sourcing this framework to make it usable and expandable for as many researchers as possible. On a different level, we are also developing a "low level design framework" according to Hampshire et al. [67], that enables non-technical researchers and instructors to create Tran*AR* training scenarios without any programming knowledge by providing preexisting components and abstraction layers for the process model descriptions as, e.g., described in [68].

Abbreviations

The following abbreviations are used in this manuscript:

AR	Augmented Reality
VR	Virtual Reality
ARBT	Augmented Reality-based Training
CBT	Computer-based Training
WBT	Web-based Training
TAM	Technology Acceptance Model
BYOD	Bring-your-own-device Approach
UI	User Interface
SUS	System Usability Scale
UEQ	User Experience Questionnaire
MARE	Mobile Augmented Reality Education Design Frameworks

References

1. Stigler, J.W.; Hiebert, J. *The Teaching Gap: Best Ideas from the World's Teachers for Improving Education in the Classroom*; The Free Press: New York, NY, USA, 1999.
2. Capper, J. E-learning growth and promise for the developing world. *TechKnowLogia* **2001**, *2*, 7–10.
3. Park, S.Y. An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. *J. Educ. Technol. Soc.* **2009**, *12*, 150–162.
4. Lahti, M.; Hätönen, H.; Välimäki, M. Impact of e-learning on nurses' and student nurses knowledge, skills, and satisfaction: A systematic review and meta-analysis. *Int. J. Nurs. Stud.* **2014**, *51*, 136–149. [CrossRef]
5. MacDonald, C.J.; Thompson, T.L. Structure, content, delivery, service, and outcomes: Quality e-learning in higher education. *Int. Rev. Res. Open Distrib. Learn.* **2005**, *6*. [CrossRef]
6. Gagne, R.M.; Briggs, L.J. *Principles of Instructional Design*; Holt, Rinehart & Winston: New York, NY, USA, 1974.
7. Gagne, R.M. Learning outcomes and their effects: Useful categories of human performance. *Am. Psychol.* **1984**, *39*, 377. [CrossRef]
8. Krathwohl, D.R.; Anderson, L.W. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*; Longman: London, UK, 2009.
9. Blattgerste, J.; Luksch, K.; Lewa, C.; Kunzendorf, M.; Bauer, N.H.; Bernloehr, A.; Joswig, M.; Schäfer, T.; Pfeiffer, T. Project Heb@ AR: Exploring handheld Augmented Reality training to supplement academic midwifery education. In *DELFI 2020—Die 18. Fachtagung Bildungstechnologien der Gesellschaft für Informatik eV*; Gesellschaft für Informatik e.V.: Bonn, Germany, 2020.
10. Santos, M.E.C.; Lübke, A.I.W.; Taketomi, T.; Yamamoto, G.; Rodrigo, M.M.T.; Sandor, C.; Kato, H. Augmented reality as multimedia: The case for situated vocabulary learning. *Res. Pract. Technol. Enhanc. Learn.* **2016**, *11*, 1–23. [CrossRef]
11. Ozdemir, M.; Sahin, C.; Arcagok, S.; Demir, M.K. The effect of augmented reality applications in the learning process: A meta-analysis study. *Eurasian J. Educ. Res.* **2018**, *18*, 165–186. [CrossRef]
12. da Silva, M.M.; Teixeira, J.M.X.; Cavalcante, P.S.; Teichrieb, V. Perspectives on how to evaluate augmented reality technology tools for education: A systematic review. *J. Braz. Comput. Soc.* **2019**, *25*, 1–18. [CrossRef]
13. Quintero, J.; Baldiris, S.; Rubira, R.; Cerón, J.; Velez, G. Augmented reality in educational inclusion. A systematic review on the last decade. *Front. Psychol.* **2019**, *10*, 1835. [CrossRef]
14. Dalim, C.S.C.; Kolivand, H.; Kadhim, H.; Sunar, M.S.; Billinghurst, M. Factors influencing the acceptance of augmented reality in education: A review of the literature. *J. Comput. Sci.* **2017**, *13*, 581–589. [CrossRef]
15. Cheng, K.H.; Tsai, C.C. Affordances of augmented reality in science learning: Suggestions for future research. *J. Sci. Educ. Technol.* **2013**, *22*, 449–462. [CrossRef]
16. Wang, Y.; Anne, A.; Ropp, T. Applying the technology acceptance model to understand aviation students' perceptions toward augmented reality maintenance training instruction. *Int. J. Aviat. Aeronaut. Aerosp.* **2016**, *3*, 3. [CrossRef]
17. Jang, J.; Ko, Y.; Shin, W.S.; Han, I. Augmented Reality and Virtual Reality for Learning: An Examination Using an Extended Technology Acceptance Model. *IEEE Access* **2021**, *9*, 6798–6809. [CrossRef]

18. Turner, M.; Kitchenham, B.; Brereton, P.; Charters, S.; Budgen, D. Does the technology acceptance model predict actual use? A systematic literature review. *Inf. Softw. Technol.* **2010**, *52*, 463–479. [CrossRef]

19. Tzima, S.; Styliaras, G.; Bassounas, A. Augmented reality applications in education: Teachers point of view. *Educ. Sci.* **2019**, *9*, 99. [CrossRef]

20. ARCore. Available online: https://developers.google.com/ar/discover/ (accessed on 28 May 2021).

21. ARKit 4. Available online: https://developer.apple.com/augmented-reality/arkit/ (accessed on 28 May 2021).

22. Palmarini, R.; Erkoyuncu, J.A.; Roy, R.; Torabmostaedi, H. A systematic review of augmented reality applications in maintenance. *Robot. Comput.-Integr. Manuf.* **2018**, *49*, 215–228. [CrossRef]

23. Blattgerste, J.; Renner, P.; Pfeiffer, T. Augmented reality action assistance and learning for cognitively impaired people: A systematic literature review. In Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments, Rhodes, Greece, 5–7 June 2019; pp. 270–279.

24. Müller, T. Challenges in representing information with augmented reality to support manual procedural tasks. *AIMS Electron. Electr. Eng.* **2019**, *3*, 71–97. [CrossRef]

25. Chidambaram, S.; Huang, H.; He, F.; Qian, X.; Villanueva, A.M.; Redick, T.S.; Stuerzlinger, W.; Ramani, K. ProcessAR: An Augmented Reality-Based Tool to Create In-Situ Procedural 2D/3D AR Instructions. In Proceedings of the Designing Interactive Systems Conference 2021, Virtual Event, USA, 28 June–2 July 2021. [CrossRef]

26. Büttner, S.; Prilla, M.; Röcker, C. Augmented Reality Training for Industrial Assembly Work-Are Projection-based AR Assistive Systems an Appropriate Tool for Assembly Training? In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–12.

27. Singh, K.; Shrivastava, A.; Achary, K.; Dey, A.; Sharma, O. Augmented Reality-Based Procedural Task Training Application for Less Privileged Children and Autistic Individuals. In Proceedings of the 17th International Conference on Virtual-Reality Continuum and its Applications in Industry, Brisbane, QLD, Australia, 14–16 November 2019; pp. 1–10.

28. Hruntova, T.V.; Yechkalo, Y.V.; Striuk, A.M.; Pikilnyak, A.V. Augmented reality tools in physics training at higher technical educational institutions. In Proceedings of the 1st International Workshop on Augmented Reality in Education, Kryvyi Rih, Ukraine, 2 October 2018; pp. 33–40.

29. Solmaz, S.; Alfaro, J.L.D.; Santos, P.; Van Puyvelde, P.; Van Gerven, T. A practical development of engineering simulation-assisted educational AR environments. *Educ. Chem. Eng.* **2021**, *35*, 81–93. [CrossRef]

30. Wang, S.; Parsons, M.; Stone-McLean, J.; Rogers, P.; Boyd, S.; Hoover, K.; Meruvia-Pastor, O.; Gong, M.; Smith, A. Augmented reality as a telemedicine platform for remote procedural training. *Sensors* **2017**, *17*, 2294. [CrossRef]

31. Goh, E.S.; Sunar, M.S.; Ismail, A.W. 3D object manipulation techniques in handheld mobile augmented reality interface: A review. *IEEE Access* **2019**, *7*, 40581–40601. [CrossRef]

32. Yusof, C.S.; Bai, H.; Billinghurst, M.; Sunar, M.S. A review of 3D gesture interaction for handheld augmented reality. *J. Teknol.* **2016**, *78*. [CrossRef]

33. Liarokapis, F.; Macan, L.; Malone, G.; Rebolledo-Mendez, G.; De Freitas, S. Multimodal augmented reality tangible gaming. *Vis. Comput.* **2009**, *25*, 1109–1120. [CrossRef]

34. Billinghurst, M.; Kato, H.; Poupyrev, I. Tangible augmented reality. *ACM Siggraph Asia* **2008**, *7*, 1–10.

35. Knierim, P.; Kiss, F.; Rauh, M.; Schmidt, A. Tangibility is Overrated: Comparing Learning Experiences of Physical Setups and their Virtual Equivalent in Augmented Reality. In Proceedings of the 19th International Conference on Mobile and Ubiquitous Multimedia, Essen, Germany, 22–25 November. 2020; pp. 299–305.

36. Qian, J.; Ma, J.; Li, X.; Attal, B.; Lai, H.; Tompkin, J.; Hughes, J.F.; Huang, J. Portal-ble: Intuitive free-hand manipulation in unbounded smartphone-based augmented reality. In Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, New Orleans, LA, USA, 20–23 October 2019; pp. 133–145.

37. Hürst, W.; Van Wezel, C. Gesture-based interaction via finger tracking for mobile augmented reality. *Multimed. Tools Appl.* **2013**, *62*, 233–258. [CrossRef]

38. Datcu, D.; Lukosch, S.; Brazier, F. On the usability and effectiveness of different interaction types in augmented reality. *Int. J. Hum. Comput. Interact.* **2015**, *31*, 193–209. [CrossRef]

39. Mossel, A.; Venditti, B.; Kaufmann, H. 3DTouch and HOMER-S: Intuitive manipulation techniques for one-handed handheld augmented reality. In Proceedings of the Virtual Reality International Conference: Laval Virtual, Laval, France, 20–22 March 2013; pp. 1–10.

40. Radu, I.; MacIntyre, B.; Lourenco, S. Comparing Children's Crosshair and Finger Interactions in Handheld Augmented Reality: Relationships Between Usability and Child Development. In Proceedings of the 15th International Conference on Interaction Design and Children, Manchester, UK, 21–24 June 2016; pp. 288–298.

41. Grandi, J.G.; Debarba, H.G.; Bemdt, I.; Nedel, L.; Maciel, A. Design and assessment of a collaborative 3D interaction technique for handheld augmented reality. In Proceedings of the 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Tuebingen/Reutlingen, Germany, 18–22 March 2018; pp. 49–56.

42. Blattgerste, J.; Renner, P.; Strenge, B.; Pfeiffer, T. In-situ instructions exceed side-by-side instructions in augmented reality assisted assembly. In Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference, Corfu, Greece, 26–29 June 2018; pp. 133–140.

43. Brooke, J. SUS: A quick and dirty'usability. In *Usability Evaluation in Industry*; Jordan, P.W., Thomas, B., McClelland, I.L., Weerdmeester, B., Eds.; CRC Press: Boca Raton, FL, USA, 1996.

44. Laugwitz, B.; Held, T.; Schrepp, M. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*; Springer: Berlin, Germany, 2008; pp. 63–76.

45. Bangor, A.; Kortum, P.; Miller, J. Determining what individual SUS scores mean: Adding an adjective rating scale. *J. Usability Stud.* **2009**, *4*, 114–123.

46. Tullis, T.S.; Stetson, J.N. A comparison of questionnaires for assessing website usability. In Proceedings of the Usability Professional Association Conference, Minneapolis, MN, USA, 7–11 June 2004; Volume 1, pp. 1–12.

47. Kortum, P.T.; Bangor, A. Usability ratings for everyday products measured with the system usability scale. *Int. J. Hum. Comput. Interact.* **2013**, *29*, 67–76. [CrossRef]

48. Schrepp, M.; Hinderks, A.; Thomaschewski, J. Construction of a Benchmark for the User Experience Questionnaire (UEQ). *IJIMAI* **2017**, *4*, 40–44. [CrossRef]

49. Linneberg, M.S.; Korsgaard, S. Coding qualitative data: A synthesis guiding the novice. *Qual. Res. J.* **2019**, *19*, 259–270. [CrossRef]

50. Zumbach, J. Problem-Based Learning in Österreich: Eine Bestandsaufnahme. In *Problem-Based Learning im Dialog (S. 15-25)*; Mair, M., Brezowar, G., Olswoski, G., Zumbach (Hrsg.), J., Eds.; Facultas: Wien, Austria, 2012.

51. Riedl, A.; Schelten, A. *Kompetenzentwicklung in Lernfeldern im Unterricht Gewerblich-Technischer Schulen*; Eusl: Paderborn, Germany, 2013.

52. Straka, G.A.; Macke, G. *Lern-lehr-Theoretische Didaktik*; Waxmann Verlag: Münster, Germany, 2002.

53. Kerres, M.; Bormann, M.; Vervenne, M. Didaktische konzeption von serious games: Zur verknüpfung von spiel-und lernange-boten. *MedienPädagogik Zeitschrift für Theorie und Praxis der Medienbildung* **2009**, 1–16. [CrossRef]

54. Die Crossmedia-Reihe "Praxisorientiert Ausbilden". Available online: http://www.kompetenzwerkstatt.net/lernbausteine (accessed on 16 April 2021).

55. Drljević, N.; Wong, L.H.; Botički, I. Where does my Augmented Reality Learning Experience (ARLE) belong? A student and teacher perspective to positioning ARLEs. *IEEE Trans. Learn. Technol.* **2017**, *10*, 419–435. [CrossRef]

56. Acosta, J.L.B.; Navarro, S.M.B.; Gesa, R.F.; Kinshuk, K. Framework for designing motivational augmented reality applications in vocational education and training. *Australas. J. Educ. Technol.* **2019**, *35*. [CrossRef]

57. Dick, W.; Carey, L.; Carey, J.O. *The Systematic Design of Instruction*; Scott Foresman: Glenview, IL, USA, 2005.

58. Kerres, M. *Mediendidaktik: Konzeption und Entwicklung Mediengestützter Lernangebote (4. Überarbeitete Auflage)*; Olden Bourg: München, Germany, 2013.

59. Bundesministerium des Inneren Bundesverwaltungsamt. *Handbuch für Organisationuntersuchungen und Personalbedarfsermittlung*; 2018. Available online: https://www.orghandbuch.de/OHB/DE/ohb_pdf.pdf?__blob=publicationFile&v=29 (accessed on 2 June 2021).

60. Anderson, L.W.; Bloom, B.S. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*; Longman: London, UK, 2001.

61. Miller, G.E. The assessment of clinical skills/competence/performance. *Acad. Med.* **1990**, *65*, S63–S67. [CrossRef] [PubMed]

62. Zhu, E.; Lilienthal, A.; Shluzas, L.A.; Masiello, I.; Zary, N. Design of mobile augmented reality in health care education: A theory-driven framework. *JMIR Med. Educ.* **2015**, *1*, e4443. [CrossRef]

63. Fehling, D.; Goertz, L.; Hagenhofer, T. Didaktisches Konzept des Projektes Social Augmented Learning. 2015. Available on-line: http://www.social-augmented-learning.de/wp-content/uploads/2015/04/SAL_Didaktisches_Konzept_20150409.pdf (accessed on 18 November 2015).

64. Willis, J. A recursive, reflective instructional design model based on constructivist-interpretivist theory. *Educ. Technol.* **1995**, *35*, 5–23.

65. Akçayır, M.; Akçayır, G. Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educ. Res. Rev.* **2017**, *20*, 1–11. [CrossRef]

66. Garcia Fracaro, S.; Chan, P.; Gallagher, T.; Tehreem, Y.; Toyoda, R.; Kristel, B.; Jarka, G.; Pfeiffer, T.; Slof, B.; Wachsmuth, S.; et al. Towards Design Guidelines for Virtual Reality Training for the Chemical Industry. *Educ. Chem. Eng.* **2021**. [CrossRef]

67. Hampshire, A.; Seichter, H.; Grasset, R.; Billinghurst, M. Augmented reality authoring: Generic context from programmer to designer. In Proceedings of the 18th Australia Conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments, Sydney, Australia, 20–24 November 2006; pp. 409–412.

68. Blattgerste, J.; Renner, P.; Pfeiffer, T. Authorable augmented reality instructions for assistance and training in work environments. In Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia, Pisa, Italy, 26–29 November 2019; pp. 1–11.