

Review

Augmented Reality, Mixed Reality, and Hybrid Approach in Healthcare Simulation: A Systematic Review

Rosanna Maria Viglialoro ^{1,*}, Sara Condino ^{1,2,*} , Giuseppe Turini ^{1,3} , Marina Carbone ^{1,2} ,
Vincenzo Ferrari ^{1,2} and Marco Gesi ^{4,5}

¹ EndoCAS Center, Department of Translational Research and of New Surgical and Medical Technologies, University of Pisa, 56124 Pisa, Italy; gturini@kettering.edu (G.T.); marina.carbone@endocas.unipi.it (M.C.); vincenzo.ferrari@endocas.unipi.it (V.F.)

² Dipartimento di Ingegneria dell'Informazione, University of Pisa, 56122 Pisa, Italy

³ Computer Science Department, Kettering University, Flint, MI 48504, USA

⁴ Department of Translational Research and of New Surgical and Medical Technologies, University of Pisa, 56126 Pisa, Italy; marco.gesi@med.unipi.it

⁵ Center for Rehabilitative Medicine "Sport and Anatomy", University of Pisa, 56121 Pisa, Italy

* Correspondence: rosanna.viglialoro@endocas.org (R.M.V.); sara.condino@endocas.unipi.it (S.C.); Tel.: +39-050-995-689 (R.M.V. & S.C.)

Abstract: Simulation-based medical training is considered an effective tool to acquire/refine technical skills, mitigating the ethical issues of Halsted's model. This review aims at evaluating the literature on medical simulation techniques based on augmented reality (AR), mixed reality (MR), and hybrid approaches. The research identified 23 articles that meet the inclusion criteria: 43% combine two approaches (MR and hybrid), 22% combine all three, 26% employ only the hybrid approach, and 9% apply only the MR approach. Among the studies reviewed, 22% use commercial simulators, whereas 78% describe custom-made simulators. Each simulator is classified according to its target clinical application: training of surgical tasks (e.g., specific tasks for training in neurosurgery, abdominal surgery, orthopedic surgery, dental surgery, otorhinolaryngological surgery, or also generic tasks such as palpation) and education in medicine (e.g., anatomy learning). Additionally, the review assesses the complexity, reusability, and realism of the physical replicas, as well as the portability of the simulators. Finally, we describe whether and how the simulators have been validated. The review highlights that most of the studies do not have a significant sample size and that they include only a feasibility assessment and preliminary validation; thus, further research is needed to validate existing simulators and to verify whether improvements in performance on a simulated scenario translate into improved performance on real patients.

Keywords: healthcare simulation; augmented reality; mixed reality; hybrid; medical training



Citation: Viglialoro, R.M.; Condino, S.; Turini, G.; Carbone, M.; Ferrari, V.; Gesi, M. Augmented Reality, Mixed Reality, and Hybrid Approach in Healthcare Simulation: A Systematic Review. *Appl. Sci.* **2021**, *11*, 2338. <https://doi.org/10.3390/app11052338>

Academic Editor: Youngho Lee

Received: 9 February 2021

Accepted: 4 March 2021

Published: 6 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Until the 20th century, the apprenticeship model, focused on the educational philosophy of "see one, do one, teach one", was the standard teaching methodology in medical education. This model, developed by Dr. William Halsted in 1890, is based on progressive responsibility culminating in almost-independence [1–3]. In other words, the trainee directly observes a procedure performed by the expert supervisor several times, then (once the apprentice is considered ready) he/she executes the same procedure by imitating the supervisor's skills; possible mistakes are prevented or fixed immediately by the supervisor to protect the patient. This model undoubtedly has strengths thanks to the trainee's early immersion in the clinical environment which allows him/her to acquire practical and applied knowledge. However, it is inefficient because it is characterized by long hours of work with poorly defined goals and random experiences depending on the flow of patients in the operating theatre [4].

Over the last decades, the rapid introduction of new techniques and surgical approaches, such as the minimally invasive and robotic procedures, combined with the new legislative restrictions on surgeons' working hours have worsened the issues of the apprenticeship model [5–7]. Thus, based on these considerations, the surgical community was forced to reconsider training strategies [8].

Simulation-based training, relying on “see one, simulate many deliberately, do one” principle, has been proposed as an excellent adjunct method to traditional medical education [9,10]. Simulation is defined as “a technique to replace or amplify real experiences with guided experiences, often immersed in nature, that evoke or replicate substantial aspects of the real world in a fully interactive way” [11]. Healthcare simulation provides the opportunity to develop the knowledge, the skills, and the attitude of medical professionals, without any risk to the patient. The trainee can explore, repeat the surgical practice in a setting that “fosters permission to fail”, recognize the mistakes and correct them, and try different approaches without jeopardizing patient safety. Thus, simulation-based medical training can be a platform for learning to mitigate ethical tensions and acquire/refine technical skills [12–15]. The main simulation techniques in healthcare include virtual reality (VR) simulation, physical simulation, and hybrid (virtual-physical) simulation.

VR simulation is based on the principles of immersion, interaction, and user involvement at different degrees: learners are immersed in a highly realistic virtual clinical environment (e.g., an operating room or an intensive care unit) and interact in real-time with the environment through virtual interactions/interfaces (e.g., virtual menus, point-and-click, etc.). Two interesting features of VR simulation are the objective evaluation of the trainee's performance, and the flexibility in selecting the virtual scenario. The latter allows the simulation to be adapted to the trainee's learning needs, enabling the learner to gain new cognitive skills, or develop and refine technical skills (e.g., instrumental navigation, clipping, slicing, etc.). However, two major limitations of VR-based simulation are the moderate sense of haptic feedback during interactions and the high costs making it prohibitively expensive for many institutions [16–18].

On the contrary, physical simulation is based on the reproduction of clinical scenarios with various degrees of realism. Low-fidelity scenarios include simple synthetic models (such as tubes or rings) made of plastic, rubber, and latex that are used to train basic clinical skills (such as knot-tying and suturing). High-fidelity scenarios include human organs or mannequins that can be used to perform the physical examination and to complete more complex procedures than just basic clinical tasks, and, unlike VR simulation, they provide the trainee realistic haptic feedback during all tool–tissue interactions [19,20]. However, the main negative aspects of this type of simulation are the need to replace the model in case of a destructive task increasing training costs and the lack of the objective evaluation of trainee's performance.

Hybrid simulators combine VR elements and physical models of the anatomy which can have an active or passive role in the simulation. Hybrid systems retain the natural haptic feedback of the physical simulation and the performance evaluation tools typical of VR simulation.

We present a systematic review of simulation techniques based on augmented reality (AR), mixed reality (MR), and hybrid approaches in the context of healthcare, to investigate the challenges and trends in this discipline. Firstly, we define AR, MR, and hybrid approaches (Section 2.1), then we examine the simulation techniques in terms of the implementation of their virtual and physical components and the accuracy validity of the adopted techniques. Finally, we evaluate the clinical validity of the simulators and their frequency of use.

2. Materials and Methods

2.1. Simulation Approach

In the literature, there is no univocal definition of the terms hybrid simulator, AR simulator, and mixed reality simulator. In performing this review, we have categorized the simulators using the following definitions:

- Augmented reality (AR) simulator: an interactive simulator in which the real-world environment is enhanced by computer-generated content perceived by the user using different senses. In these simulators, the specifically designed physical component could be absent (using only pre-existing elements of the real-world environment, such as the ground, a wall etc.), passive (not actively participating in the simulation), or active (providing/enabling specific functionalities in the simulation).
- Hybrid simulator: an interactive simulator in which the system integrates both a virtual and a physical module. In these simulators the physical parts are always present, but they could play either a passive or active role in the simulation.
- Mixed reality (MR) simulator: an interactive simulator in which real content (physical objects) and virtual information (computer-generated content) are merged so that they can interact with each other in real time. In these simulators the physical parts can interact with the virtual content (and/or vice versa).

Therefore, as an example, a simulation system could have been classified as AR and hybrid but not as MR if the system integrated both virtual content and physical parts but did not enable any virtual physical interaction.

The results of the classification according to these definitions are given in Section 3.10, which reports the associate statistical data collected by answering statistical questions SQ2 (“How many AR simulators are there?”), SQ3 (“How many MR simulators are there?”), and SQ4 (“How many hybrid simulators are there?”).

2.2. Literature Search

The literature search was conducted using the following nine electronic databases: Scopus, Google Scholar, PubMed, ProQuest, ScienceDirect, Wiley Online Library, IEEE Xplore, Taylor & Francis Online and SAGE. The searches were limited to studies published between February 2008 and April 2020 inclusive. The review was conducted by four reviewers and the searches in all the databases above returned 262 results.

We have used the following research terms:

- (Augmented Reality OR AR) AND (Simulation OR Simulator) AND (Healthcare OR Medicine OR Surgery OR Surgical)
- (Mixed Reality OR MR) AND (Simulation OR Simulator) AND (Healthcare OR Medicine OR Surgery OR Surgical)
- (Hybrid) AND (Simulation OR Simulator) AND (Healthcare OR Medicine OR Surgery OR Surgical)

The search in the online digital libraries was conducted in April 2020.

2.3. Study Selection

The selection process started with 262 studies collected from online digital libraries. We defined three questions to select and include relevant studies:

- Q1: Is the study relevant to healthcare simulation for improving the medical technical and/or non-technical skills?
- Q2: Are the simulation techniques based on AR, MR, and/or hybrid approach?
- Q3: Does the study concern the development of an ad-hoc simulator or the evaluation of a commercial simulator?

The selection process has been divided into four phases:

1. Removal of duplicates from nine different databases. After removing them, 98 studies remained.

2. Removal of editorials (1), reviews (2), book chapters (1), conference abstracts (8), thesis (9), reports (1), and reflections (4). After removing them, 72 studies remained for the next phase.
3. Removal of studies after reading the title and abstract. The removed articles do not resolve question Q1. After removing them, 35 studies remained.
4. Removal of studies after reading the full text, since some papers are still dubious after step 3. The removed articles do not resolve questions Q2 and Q3. A total of 23 studies remained relevant for our review.

Figure 1 shows the flow chart for the selection of studies according to PRISMA statements [21].

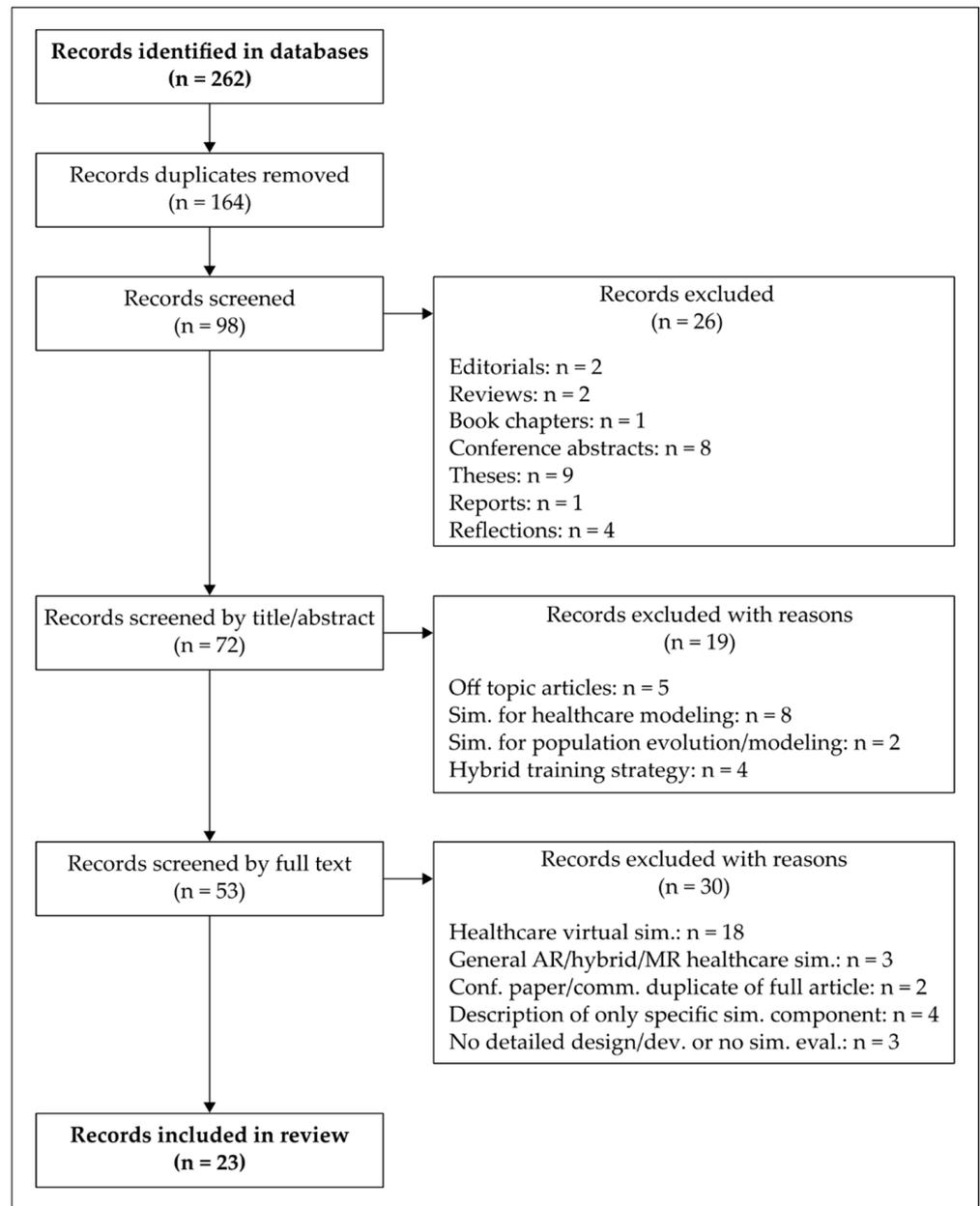


Figure 1. Flow chart illustrating the selection of studies according to PRISMA statements.

2.4. Research Questions

To guide the entire methodology and to help define the purpose of our systematic review, we have defined 23 research questions that have been classified into three categories: General Question (GQ), Focus Question (FQ), and Statistical Question (SQ). All these research questions (GQs, FQs, and SQs) are listed in Table 1.

Table 1. Research questions.

Type/Code	Research Questions	
General Questions	GQ1	Which is the target clinical area?
	GQ2	Which skills are addressed by simulator? (technical skills, non-technical skills)
	GQ3	Does the simulator integrate haptic feedback?
	GQ4	What kind of sensors are used?
	GQ5	Is the simulator patient-specific? (patient-specific, not-patient-specific)
	GQ6	Are the simulator components reusable?
	GQ7	Is a clinical performance evaluation performed?
	GQ8	Does the simulator allow the selection of different scenarios based on trainee's needs?
	GQ9	How portable is the simulator? (very portable, portable, not portable)
Focus Questions	FQ1	How invasive is the simulated surgical approach? (non-inv., minimally-inv., inv.)
	FQ2	Which mode is used to convey haptic feedback?
	FQ3	Which tracking approach is used to implement AR? (marker-based or marker-less)
	FQ4	What visualization type is used? (monitor, hand-held display, HMD, projection)
	FQ5	What kind of artificial intelligence technique is implemented?
	FQ6	Which types of phantom are used? (commercial, custom-made)
	FQ7	Which simulator manufacturing technique is used?
	FQ8	Which metric is used for performance evaluation?
	FQ9	Which evaluation method is performed to validate the simulator?
	FQ10	Is a statistical analysis performed?
Statistical Questions	SQ1	How many commercial simulators are used?
	SQ2	How many AR simulators are there?
	SQ3	How many MR simulators are there?
	SQ4	How many hybrid simulators are there?

GQs concern the target clinical area (GQ1), the skills addressed by the simulator (GQ2), the presence of haptic feedback (GQ3), integrated sensor types (GQ4), the presence of patient-specific anatomy (GQ5), the reusability of simulator components (GQ6), the evaluation of clinical performance (GQ7), the possibility to select different scenarios based on the trainee's needs (GQ8), and simulator portability level (GQ9).

FQs refer to specific questions to help answering GQs providing more details on how invasive the simulated surgical approach is (FQ1), the mode used to convey haptic feedback (FQ2), the tracking approach adopted (FQ3), the visualization techniques used (FQ4), the artificial intelligence technique implemented (FQ5), the phantom type used (FQ6), the manufacturing techniques adopted (FQ7), the performance evaluation metric used (FQ8), the evaluation method performed (FQ9), and the presence of statistical analysis (FQ10).

SQs concern the current trends in the choice of simulation approaches in the healthcare sector (SQ2, SQ3, and SQ4) and the types of simulators (commercial or custom-made) used to implement a specific approach or to compare two different approaches (SQ1).

3. Results

A total of 23 studies remained relevant for our review. Table 2 contains basic information on the selected studies: the authors and year of publication and reference number (column Reference) and the publication title (column Title).

Table 2. Selected studies.

Reference	Title of Selected Study
Coelho 2020 [22]	Augmented reality and physical hybrid model simulation for preoperative planning of metopic craniosynostosis surgery.
Condino 2018 [23]	How to build a patient-specific hybrid simulator for orthopedic open surgery: benefits and limits of mixed-reality using the Microsoft HoloLens.
Condino 2011 [24]	How to build patient-specific synthetic abdominal anatomies. An innovative approach from physical toward hybrid surgical simulators.
Feifer 2011 [25]	Randomized controlled trial of virtual reality and hybrid simulation for robotic surgical training.
Ferrari 2016 [26]	Augmented reality visualization of deformable tubular structures for surgical simulation.
Fuerst 2014 [27]	A novel augmented reality simulator for minimally invasive spine surgery.
Fushima 2016 [28]	Mixed-reality simulation for orthognathic surgery.
Halic 2010 [29]	Mixed reality simulation of rasping procedure in artificial cervical disc replacement (ACDR) surgery.
Huang 2018 [30]	The use of augmented reality glasses in central line simulation: “see one, simulate many, do one competently, and teach everyone”
Jain 2019 [31]	Virtual reality based hybrid simulation for functional endoscopic sinus surgery.
Keebler 2014 [32]	Building a simulated medical augmented reality training system.
Lahanas 2015 [33]	A novel augmented reality simulator for skills assessment in minimal invasive surgery.
Lee 2013 [34]	Augmented reality intravenous injection simulator based 3D medical imaging for veterinary medicine.
Lin 2012 [35]	The development of optical see-through display based on augmented reality for oral implant surgery simulation.
Loukas 2013 [36]	An integrated approach to endoscopic instrument tracking for augmented reality applications in surgical simulation training.
Nomura 2015 [37]	Laparoscopic skill improvement after virtual reality simulator training in medical students as assessed by augmented reality simulator.
Onishi 2013 [38]	AR dental surgical simulator using haptic feedback.
Parkes 2009 [39]	A mixed reality simulator for feline abdominal palpation training in veterinary medicine.
Tai 2009 [40]	Augmented-reality-driven medical simulation platform for percutaneous nephrolithotomy with cybersecurity awareness.
Thomas 2014 [41]	The validity and reliability of a hybrid reality simulator for wire navigation in orthopedic surgery.
Tsujita 2018 [42]	Development of a surgical simulator for training retraction of tissue with an encountered-type haptic interface using MR fluid.
Van Duren 2018 [43]	Augmented reality fluoroscopy simulation of the guide-wire. insertion in DHS surgery: a proof of concept study.
Viglialoro 2018 [44]	Augmented reality to improve surgical simulation. Lessons learned towards the design of a hybrid laparoscopic simulator for cholecystectomy.

3.1. Target Clinical Area (GQ1, FQ1)

Simulation-based education is widely incorporated as a means of effective training to learn technical and non-technical skills in almost all areas of healthcare. In the past, it has been mainly used for training medical professionals to reduce errors during surgery. Today, there is a growing trend to use simulation both as a tool for objective skill assessment and as an anatomy learning tool to circumvent the drawbacks associated with conventional anatomy training based on cadaver dissection.

With regard to question GQ1 (“Which is the target clinical area?”), 22 of the 23 studies concern simulators for training of surgical tasks/procedures (such as neurosurgery [22,42], abdominal surgery [24,26,40,44], orthopedic surgery [23,27,29,41,43], dental surgery [28,35,38],

otorhinolaryngological surgery [31], laparoscopic basic tasks [25,33,36], abdominal surgery and basic tasks [37], palpation [39], intravenous injection [34], and central venous catheter insertion [30]); eight of them use the simulation also as an assessment tool [25,33,34,37,40,41,43]. One study reports the use of simulation only for didactic purposes to learn anatomy [32].

Surgical Approaches (FQ1)

A surgical procedure can be classified according to the degree of invasiveness in non-invasive surgery, minimally invasive surgery (MIS), or invasive surgery (i.e., open surgery). The non-invasive surgical approach is a conservative treatment that does not require any incision in the skin; such procedures range from simple observation to surgical specialties such as radiosurgery. The MIS approach refers to surgical procedures using small incisions. This approach allows the patient to recover faster with less pain than the open surgery approach characterized by large incisions.

In answering question FQ1 (“What surgical approach is implemented in sim.? (non inv., minimally inv., inv.)”), a total of 15 studies concern the simulation of MIS procedures [24–27,29–31,33,34,36,37,40,41,43,44], six concern the simulation of invasive approaches [22,23,28,35,38,42], and only one concerns the simulation of a non-invasive approach [39].

3.2. Technical and Non-Technical Skills in Surgery (GQ2)

Mastery of both technical and non-technical skills is required to perform a safe procedure. The former are defined as psychomotor actions or related mental faculties acquired through practice and learning, pertaining to a particular craft or profession. The latter are defined as the social (teamwork, leadership, communication), cognitive (situation awareness, decision-making), and personal resource skills (stress, fatigue, and stress management) that complement technical skills, contributing to safe and effective performance [45]. As reported by [46], the technical skills are no longer enough for the delivery of a modern and safe surgical practice; indeed, more frequently the risk to the patient is from failure of non-technical skills [45,46].

Regarding question GQ2 (“Which skills are addressed by simulator? (technical skills, non-technical skills)”), in the largest part of the analyzed studies [22–31,33,34,37,38,40,41,44], the simulators address both the technical and non-technical skills. Only in five studies [35,36,39,42,43] do the simulators address only the technical skills. Indeed, many studies show that simulation-based education is a powerful tool to teach both technical and non-technical skills to individual surgeons and surgical teams without risking patient safety.

3.3. Haptic Feedback (GQ3, FQ2)

Haptic feedback plays a key role in medical simulations because it increases the simulation fidelity with beneficial effects on training. In particular, haptic feedback is an important add-on for virtual reality simulators because it improves the perception of virtual objects giving the user the illusion of touch. Many VR simulators recreate the sense of touch through haptic devices that apply forces, vibrations, or motions to the user during their interaction with the virtual environment.

In particular, haptic technology can be used to simulate the tactile properties of tissues (i.e., the stiffness of a tissue is essential during a palpation procedure) and the manipulation of surgical instrumentation [47].

With regard to question GQ3 (“Does the simulator integrate haptic feedback?”), 22 [22–31,33–44] of the 23 studies integrate haptic feedback. The only system that does not offer such feedback is the simulator presented in [32] designed for the AR visualization of anatomical models for educational purposes.

In more detail, answering question FQ2 (“Which mode is used to convey haptic feedback? (actual interaction with physical components, haptic interface)”), the largest part

of the studies analyzed [22–24,26,27,29–31,33,34,36,38,41–44] offer the trainee the possibility to use actual surgical instruments and/or manipulate/palpate physical anatomical replicas.

Two of the reviewed simulators [39,40] incorporate a haptic software module and deliver the simulated haptic experience via commercial haptic interfaces: Phantom Premium devices [39] and Phantom Omni devices [40] (Sensable Technologies, Woburn, MA). The main difference between the two studies is that haptic interfaces are used in [39] to interact with virtual models of organs superimposed on a physical model (a cat toy), whereas in [40] they are used to interact with a purely virtual representation of anatomy to mimic the insertion of a percutaneous needle and to mimic hand palpation.

3.4. Implementation of the Virtual Component of the Simulators (FQ2–5)

In the following paragraphs we report technical details on the implementation of the virtual component of the simulator, including the tracking approach adopted for deriving the spatiotemporal relationship between the real and virtual worlds (FQ3), the display technologies to provide the user with computer-generated information (FQ4), and the implementation of artificial intelligence (AI) techniques (FQ5).

3.4.1. Tracking Approach (FQ3)

The knowledge of the spatiotemporal relationship between the real and virtual worlds is a key aspect in the development of a MR simulator for allowing a proper interaction of virtual elements (e.g., virtual models of the anatomy) with real objects (e.g., surgical tools). Moreover, accurate and fast estimation of the viewing pose relative to the real objects is a crucial challenge for a proper alignment of the virtual content to the real-world in AR/MR application.

Tracking methods commonly employed comprise the following approaches:

1. Vision-based approaches, that can be further categorized into two mutually non-exclusive methods: marker-based and marker-less (i.e., location-based position).

A marker is a distinguishable artificial element that a computer system can detect using image segmentation, pattern recognition, and computer vision techniques. Marker-based methods are fast; however, inherent drawbacks of these approaches lie in the fact that marker detection is very sensitive to marker occlusion and poor ambient lighting (that can make the markers unrecognizable). As for the latter issue, infrared (IR) retro-reflective markers can be used to improve the reliability of tracking, reducing the effects of ambient illumination.

2. Other sensor-based approaches (apart from vision sensors) including electromagnetic tracking, acoustic tracking, and inertial tracking sensors.
3. Hybrid techniques that combine marker-based with marker-less approaches or vision-based and sensor-based techniques.

Our analysis shows that none of the literature simulators employ a marker-less method alone, instead there are several systems based on the use of markers. Explored marker-based solutions mainly use planar printable markers such as Vuforia Image Targets (images that Vuforia Engine can detect and track) [48], square black and white patterns [32–35,42], and colored strips [41]. Only two systems [29,43] employ non-planar markers: retroreflective IR markers tracked by the Vicon system in [29] and two colored (green and yellow) markers in [43].

As for the sensor-based approach, our literature search shows that the most common sensors used for the development of AR/MR simulators are electromagnetic (EM) coils [24,26,31,41,44] that do not require line-of-sight and can be used to track hidden anatomical structures as described in [24,26,44].

Finally, some simulators adopt a hybrid approach. For example, the minimally invasive laparoscopic system described in [33] employs a planar marker to track the box-trainer; an electromagnetic sensor, a rotary encoder, and IR sensors are instead used to provide the laparoscopic tool kinematics (3D pose of the tool in six DOF (degrees of freedom), shaft

rotation, and opening angle of the tooltip). The simulator in [34] employs a gyro sensor coupled with a planar marker for an accurate tracking of the instrument (a syringe). Another hybrid approach is presented in [41] for estimating the pose of the surgical instrument with respect to the camera: first, an algorithm creates an adaptive model of a color strip attached to the distal part of the tool, then another program tracks the endoscopic shaft, using a combined Hough–Kalman approach.

3.4.2. Visualization Modality (FQ4)

Available display technologies to provide the user with computer-generated information include: 2D monitors, hand-held displays (mobile phones and tablets), head-mounted displays (HMDs), and spatial projection-based AR displays.

Most of the revised simulators [25–28,33,34,36,38,39,41,43,44] use a traditional 2D display. In these simulators (except for the commercial LapSim[®] and ProMIS[®] systems), the virtual information is not designed to interact in real-time with the real content (physical objects). Moreover, most of these simulators are for minimally invasive procedures not performed under direct vision. Indeed, [25,26,36,44] are designed for laparoscopy, for endoscopic surgery [33], and for wire navigation in minimally invasive procedures [41,43]. Finally, in [27] the monitor is used to show simulated fluoroscopic images to guide spine surgery. Among the other simulation systems for procedures involving the direct visualization of the anatomy, a particular solution is presented in [38]. This system uses a traditional monitor in a tilted position coupled with a half mirror placed horizontally between the user's head and his/her hands to visualize the virtual information superimposed onto the anatomical physical replica.

Six papers [23,29,30,34,35,40] report the use of a HMDs. These intrinsically provide the user with an egocentric viewpoint and allow handsfree work; for these reasons, HMDs are deemed the most ergonomic solution for applications including manual tasks performed by the user under direct vision, similar to what happens in open surgery. The devices selected by the AR/MR simulators evaluated are optical see-through HMDs (OST HMDs): these displays offer an instantaneous and unobstructed full resolution view of the real environment allowing a naturalistic experience. In more detail, the simulators in [23,40] are based on Microsoft HoloLens (1st generation), [30] is based on AiRScouter WD-200B, and [35] presents the design of an innovative custom-made OST HMD. Finally, HMDs are also used in [29] (NVIS nVisor ST50) and [31] (HTC VICE MIS) for virtual reality immersive experience.

Only two simulators [22,32] employ a hand-held display: these systems are designed for AR visualization of VR anatomical models for surgical planning [22] or training of medical and anatomy students [32]. None of the simulators employ a spatial projection-based AR display.

3.4.3. Artificial Intelligence Techniques Integrated in the Simulation (FQ5)

The use of AI algorithms is raising in the development of surgical simulators as the AI potential is huge for automatic performance evaluation, metrics extraction, simulation level adaptation to the trainee performance, and realistic force feedback implementation.

Actually, in the literature, AI methods are studied but not yet robustly integrated. Indeed, among all the studies analyzed, only four report the use of AI methods [25,36,37,40]. In [25] and [37], the authors use the commercial ProMIS hybrid laparoscopic simulator in a training program. ProMIS integrates a machine learning algorithm for the tracking of the laparoscopic instrumentation. In particular, the instruments are optically recognized through color-contrasted adhesive labels that are affixed to their distal tips and through a proprietary formula that combines time and path length the simulator manages to record and evaluate the economy of motion, and then generates a numeric score for the execution time. The same approach is used in [36] with the scope to estimate the 3D pose of the surgical instrument with respect to the camera and follow its movements. In this paper the adaptive algorithm developed through the color tracking method is compared with

a second algorithm that tracks the endoscopic shaft using a combined Hough–Kalman approach. Here the final aim is to achieve a robust interaction with the virtual world and to improve the realism of the rendering when the virtual scene is occluded by the instrument.

In [40], AI is used to improve the realism of the force feedback provided in the AR simulator. In particular, the authors use a cyber security algorithm to anonymously and safely record actual force feedback from real surgical interventions, then they analyze the data recorded to provide precise muscle memory acquisition during the training procedure.

3.5. Implementation of the Physical Component of the Simulator (GQ4, FQ6–7)

The physical components of the simulators are based mostly on phantoms (plastic structures) simulating parts of the patient body or a full human body. The more complex phantoms are equipped with sensors and computer software that allow an objective performance assessment and the implementation of advanced functionalities such as guidance information to the trainee, simulation of physiological functions, etc. The main advantage of using physical models is that they provide a realistic training environment owing to the real haptic feedback provided and the possibility to use actual surgical tools.

3.5.1. Materials and Fabrication Techniques (FQ6–7)

Simulators can consist of commercial or custom-made physical components. The analysis performed to answer question FQ6 (“Which phantoms are used? (commercial, custom-made)”) reveals that most of the simulators analyzed are custom-made; indeed, a total of 13 studies employ custom-made phantoms, and only five studies employ commercial phantoms.

Below, the manufacturing techniques for both of these types of simulators are reported to answer question FQ7 (“Which simulator manufacturing technique is used?”). As for the simulators with custom-made anatomical replicas, only three studies [29,35,38] do not provide details on the manufactured process, while the others provide a description of the manufacturing method useful to reproduce the simulator.

The study in [27] presents a peculiar method for fabricating the anatomical replicas. The phantom includes artificial vertebrae and soft tissue. The synthetic vertebrae are manufactured starting from frozen and formalin-fixed thoracolumbar spines; they are defrosted and dissected, then embedded using a fast curing plastic. In particular, polyurethane foam recipe is used for the inner cancellous core of vertebrae while a resin is used covering a layer of cortical bone. Finally, different mixtures of silicone rubber are used to manufacture the human’s erector spinae and the skin, while a thin plastic foil is used to imitate a muscle fascia.

In a large part of the studies [22–24,26,28,29,31,34,44], the authors instead developed phantoms extracting the 3D anatomical models from CT images of real patients. In [22–24,26,29,31,34,44], the virtual anatomical model is turned into a tangible physical replica by using 3D printing technologies and casting fabrication processes. For example, in [34], the authors fabricate a silicone model using a casting technique based on a 3D printed forelimb prototype of a beagle dog.

The remaining studies explored the use of 3D printing with the resin and acrylonitrile butadiene styrene (ABS). The resin is used to mimic the skull bone in [22] and to manufacture anatomical structures and surgical tools in [31]. In addition, in [22] the authors improve the model according to a handmade process with a platinum-cure silicone (Dragon Skin; Smooth-On, Inc, Macungie, PA, USA), mixed with some additives in order to mimic human tissue properties such as textures, consistencies, and mechanical resistance. Instead, ABS is employed to 3D print a hip models in [23] because it adequately approximates the mechanical behavior of the bony natural tissue. After that, the hip model is embedded in a soft synthetic foam and covered with a skin-like material which allows an accurate simulation of palpation and incision. In addition, in [24,26,44] the ABS material is used to 3D print molds for anatomical replicas. In [24], the hybrid simulator includes a physical commercial trunk phantom with replicas of the liver, gallbladder, pancreas, and stomach.

The physical environment is enriched with coherent virtual models of the entire abdomen. The replicas are obtained using casting processes of silicone materials and pigment powders. In addition, the stomach and liver are EM sensorized to quantify deformations caused by surgical action. In [26,44], the hybrid simulator developed by Viglialoro et al. includes both patient-specific models (e.g., liver, gallbladder, pancreas, abdominal aorta, esophagus–stomach–duodenum) and non-patient-specific synthetic organs designed in a CAD environment (e.g., arterial tree, biliary tree, and connective tissue). The strategy used to manufacture patient-specific replicas is the same as that used in [24]. The manufacturing process of the arterial tree and biliary tree involves the use of nitinol tubes joined together by tin wires and covered by a thin silicone layer. In addition, EM sensors are inserted inside the nitinol tubes to implement an AR solution allowing the real-time visualization of the Calot’s triangle. Finally, the synthetic tissue is produced in the form of thin sheets in gelatinous material (Pysyllogel Fiber powder).

The research in [28] developed a hybrid simulator that integrates a 3D virtual dento-skeletal model with the real dental cast model using the reference splint. However, the authors do not provide details on the manufactured process used to create the dental cast model.

Only five studies employ commercial phantoms. Among these, four are anthropomorphic [30,37,41,43]. In [41,43], both studies employ a plastic foam femur surrogate (Sawbones AG, Karlihof, Switzerland) to fabricate a hybrid simulator for wire navigation in orthopedic surgery. The research reported in [30] completes a commercial physical mannequin for internal jugular vein central line insertion with AR glasses. Finally, in [37] the authors customize the ProMIS simulator with a gallbladder model (Limbs & Things, Bristol, UKd) to perform a cholecystectomy procedure.

In [40], the authors report the use of commercial non-anthropomorphic phantoms. In detail, they employ two Phantom Omni (a six DOF haptic device) with one stylus that mimics the percutaneous needle during the simulation of percutaneous nephrolithotomy.

Finally, the research in [39] reports the use of toy as the main physical component of a simulator developed for training of veterinarians: two commercial Phantom Premium 1.5 haptic devices are positioned on either side of a toy cat with virtual representations of the chest and abdominal organs superimposed on the physical model.

3.5.2. Sensors Types (GQ4)

In addition to the use of position sensors for the tracking of tools and anatomical structures, in [40,42] force sensors are used to measure the puncture force during percutaneous renal access [40] and the retracting force of soft biological tissues [42].

3.6. Patient-Specific Simulation (GQ5)

As reported by Ryu et al. [49], the patient-specific simulation is very useful because it provides an accurate representation of intraoperative conditions related to the patient anatomy and the target surgical technique, and it allows trainees to try different approaches that can translate into training for complication avoidance.

Among all the studies analyzed, 13 report the use of a patient-specific simulator. In two studies [28,40], the 3D models are purely virtual, whereas in nine studies [22–24,27,29,31,34,35,38] virtual models are combined with physical replicas to create a more complex environment; finally, in only two studies [26,44] the patient-specific models are purely physical. All 3D models are obtained starting from the segmentation of the CT images of real patients.

3.7. Reusability of Simulator Components (GQ6)

Two key challenges to the reduction of the training costs are the reusability of the medical simulator components and the minimization of spare parts cost. With regard to question GQ6 (“Are the simulator components reusable?”), in the largest part of the studies analyzed [22–25,27–35,38–41,43] the entire simulator is reusable. This is due to the fact that the authors perform only non-destructive tasks such as palpation, basic laparoscopic

tasks (i.e., instrument navigation, peg transfer, clipping of virtual objects, etc.), insertion of tools (i.e., wire navigation, intravenous injection, etc.), and surgical planning. In particular, the authors use materials that are extremely durable over time such as silicone rubbers, polyurethane, and plastic materials.

In studies [26,44], the authors have designed solutions to make all anatomical components reusable such as the liver, gallbladder, pancreas, abdominal aorta, esophagus–stomach–duodenum, arterial tree, and biliary tree, except the connective tissue which must be dissected during each training session. Finally, only in four studies [22,29,36,37] was the entire simulator substituted after each training session.

3.8. General Features of Simulation Systems (GQ7–9, FQ7–10)

The key feature of simulation-based medical education is the knowledge of the results of the trainee performance during a learning experience because that leads to effective learning. Other important simulation features include the possibility to create medical scenarios of progressive difficulty and the portability of the medical simulators [50]. Each feature will be explained in the next three subsections.

3.8.1. Performance Evaluation Metrics (GQ7, FQ7–8)

The assessment of clinical competence is one of the most difficult and important tasks in medical education because it provides feedback to trainees about their clinical skills, supports self-paced learning, assures the public, and provides evidence toward the certification of achievement of clinical competencies.

There are several methods of assessment, each with its own strengths and limitations. As reported by Epstein in [51], the main strategies include written examinations, direct observation or video review, assessments by supervising clinician, clinical simulation, and multisource (“360-degree”) assessments and portfolios. The appropriateness of each method depends on the goal that is addressed (i.e., measuring performance or skill acquisition, etc.) and on the level of the learner’s education.

Ryall et al. [52] present a systematic review of the simulation as a clinical assessment tool. They suggest that the simulation can be a valid and reliable method to assess the technical skills and to determine the skill level and the capability to practice safely. Indeed, the simulation can both differentiate performance between experts and novices on given tasks and also identify poor performers.

With regard to question GQ7 (“Is a clinical performance evaluation performed?”), a total of eight studies assess clinical performance. In answering question FQ8 (“Which metric is used for performance evaluation?”), two methods of performance evaluation are used [53]: the first technique is based on the human rater’s score performance using checklists on scoring rubrics, and it is used in [30]; instead, the second method is based on the automated scoring of measurements integrated into the simulator itself, and it is used in the remaining studies [25,33,34,37,40,41,43].

The results in answering question FQ7 (“What kind of artificial intelligence technique is implemented?”) are that one of six studies used an automated scoring based on artificial neural networks [40].

3.8.2. Implementation of Different Levels of Complexity (GQ8)

To design an effective learning experience in simulation-based medical education, an important factor is the possibility to offer a wide range of task difficulty levels. An appropriate level of training allows the trainee to increase the mastery of skills; the learners have opportunities to engage in the practice of medical skills, starting from basic techniques (novice levels) and proceeding to train at progressively higher difficulties (expert levels) [50]. Five of the 23 articles address question GQ8 (“Does the simulator allow selection of different scenarios based on trainee’s needs?”).

In [33], the authors propose an AR laparoscopic simulator including a box-trainer, a camera, and a set of laparoscopic tools equipped with custom-made sensors. Such

a system allows the trainee to interact with various VR training elements. To this, the authors implement three different training tasks: instrument navigation, to improve the perception of depth of field; peg transfer, for hand-eye coordination skills; and clipping, for bimanual operation.

In [40], the authors developed an AR simulator (SimPCNL) for percutaneous nephrolithotomy and successively compared it with the commercial virtual simulator, PERC Mentor (Symbionix, Cleveland, OH). Both simulators allow the trainees to practice basic tasks for percutaneous access procedures performed under real-time fluoroscopy on a variety of virtual patients.

Two studies [25,37] concern two laparoscopic commercial simulators: the LapSim® virtual simulator (Surgical Science Inc, Minneapolis, MN, USA) and the ProMIS hybrid simulator (Haptica, Dublin, Ireland) that offer different training scenarios with different levels of complexity. No outcome has been reported on this aspect of simulation. The former simulator (the LapSim) is a high-fidelity simulator available in two versions: with and without haptic feedback. This modular system consists of laparoscopic instruments as interfaces connected to a computer. Modules include basic laparoscopic skills, ranging from navigation to more advanced skills (e.g., coordination, grasping, cutting, clip applying, suturing, etc.) and multiple surgical procedures (e.g., cholecystectomy, appendectomy, laparoscopic gynecology, etc.). The latter system (the ProMIS) combines a laparoscopic mannequin connected to a laptop with integration also into a virtual environment. It uses real surgical instruments which are tracked during the tasks to provide an accurate and objective assessment of the user's performance. The system includes both basic laparoscopic tasks and entire surgical procedures, such as appendectomy, colectomy, cholecystectomy, etc. Different physical models (such as suturing pads) can be inserted into the mannequin; for example, in [25] the authors used the MISTELS (McGill Inanimate System for Training and Evaluation of Laparoscopic Skills) task set, whereas in [37] the authors used an object-positioning module and a gallbladder model during the training sessions.

In [44], the authors developed an AR simulator for laparoscopic cholecystectomy. The key feature of this system is the capability to create physical and virtual scenarios with different degrees of complexity, allowing the trainee to acquire both the dexterity necessary for good practice and the decision-making skills. Due to the morphological and topological variations that occur naturally in human hepatobiliary anatomy, the authors predisposed physical models of different anatomical variations of the arterial and biliary trees and implemented an easy connection/disconnection coupling to facilitate any substitution on demand. This gives the tutor the possibility to choose the anatomical variations according to the trainee's level of experience.

3.8.3. Portability (GQ9)

The portability of a simulator is a major factor in spreading the use of medical simulators in educational and training settings. In answering question GQ9 ("How portable is the simulator? (very portable, portable, not portable)"), we have defined three levels of portability:

- A very-portable simulator is a commercial or custom system designed to be held in the hands and/or on the head or that can be easily carried. Seven studies [22,23,25,30,32,35,37] report the use of very portable simulators.
- A portable simulator is a system designed to have a simple assembly/disassembly of the setup and (eventually) an easy calibration procedure. Thirteen studies [24,26–28,31,33,34,36,39,41,43,44] report the use of portable simulators.
- A non-portable simulator is a system that requires a dedicated room and/or has a complex assembly/disassembly of the setup and (eventually) a difficult calibration procedure. Two studies [29,42] report the use of non-portable simulators.

In [40], the authors use two simulators: one commercial, identified as very portable; and one custom, identified as portable.

3.9. Evaluation of Simulators (FQ9–10)

To guarantee proper simulator-based training it is important to verify whether improvements in performance on medical simulators translate into improved performance on real patients. In [53], the authors affirm that the evaluation of medical simulators has to satisfy two main criteria that are “validity” and “reliability”.

Overall, the “validity” is the degree to which a method measures something. The main types of validity are face, content, construct, concurrent, and predictive validity. In the medical simulation context:

- Face validity refers to simulator realism, and it is assessed by experts by means of questionnaires or surveys;
- Content validity measures the appropriateness and usefulness of the simulator as a training tool, and it is typically assessed by experts with checklists;
- Construct validity determines the ability of simulator to differentiate between expert and novices;
- Concurrent validity indicates the correspondence between trainees’ performance tested on a simulator and on a gold standard method or against another, previously validated, simulator;
- Predictive validity denotes the ability of the simulator to predict future performance in real scenarios [53,54].

“Reliability” refers to the consistency of measurements. However, there are two other less important criteria that are “fairness” and “usability”: the first is an aspect of validity, and it refers to absence of any bias (test free of bias, lack of favoritism toward test takers, accessibility, and validity of score interpretations), while the second concerns implementation costs, time required, ease of administration, and comprehensibility of the results for the users [53,54].

To answer question FQ9 (“Which evaluation method is performed to validate the simulator? (technical, validity, technical + validity)”), we have defined three evaluation stages: only technical evaluation, only validity evaluation, both validity and technical evaluations.

The technical evaluation concerns accuracy, reliability, and efficiency, and only nine studies out of 23 report a technical evaluation [22,27,28,31,35,36,38,42,43]. In [35,38], the authors evaluate the displays in terms of registration accuracy [35,38] and frames-per-second (FPS) during the visualization of surface and volume data [35].

Measurements of force are performed in [27,42]. Fuerst et al. [27] evaluate the physical components of the simulator measuring the average force during transpedicular insertions performed on six vertebral models to perform a comparison with human specimens. Tsujita et al. [42] instead evaluate whether the developed haptic interface and the mechanical components satisfy the specifications in terms of force requirements.

In [23,28,36,43], the authors evaluate the accuracy of simulators that include an AR component. Among them, Condino et al. in [23] evaluate the accuracy of a simulator developed for Microsoft HoloLens. The accuracy was evaluated in terms of the perceived position of AR targets; moreover, the authors evaluate the workload and usability of the HoloLens for their application considering visual and audio perception, interaction, and ergonomics issues.

In [31], the authors evaluate the reliability and efficiency of the simulator. The authors evaluate the accuracy of the registration between real and virtual scenarios using the Euclidean distance between the 3D point where the resident hit the equipment in the simulation model and the actual 3D point in the real world, and the robustness, checking if the registration error depended on landmark points. Additionally, the authors demonstrated the efficiency of the system.

Validity evaluation: Eleven studies out of 23 report face/content/construct/concurrent/predictive validity tests [22,25,29,30,32–34,37,39,40]. Among them, three studies [29,32,39] perform a non-detailed preliminary evaluation of simulators, and the results obtained should be confirmed in larger studies. The sample sizes are five (expert physicians), five (undergraduate students), and seven (veterinary students), respectively. In [29], the au-

thors evaluate the effectiveness of the system to teach anatomy and to grasp the basics of a specific orthopedic surgery. In [32], both content validity and usability evaluation are performed in terms of how easy to use the simulator is. In [39], the authors perform only face validity.

In [40], the authors report face, content, construct, skills improvement validity, and criterion validity (often divided into concurrent and predictive validity). They compare the developed simulator and the PERC Mentor.

Simulator's face and content validity are evaluated by 38 expert surgeons in [22] and by 40 subjects in [34]. In both of these studies, a questionnaire on the realism and the usefulness of simulators is used, and in [22] the authors also address questions on the value of AR for medical training. Overall, the opinions of the users are very positive in all studies. Finally, in [34], proficiency level is also assessed.

Concurrent validity is reported in four studies [25,30,37,40]. In [25], the authors evaluate the performance of 20 medical students in robotic surgery sessions before and after training using ProMIS and the LapSim. The results show that the use of ProMIS and LapSim simulators in conjunction with each other can improve robotic console performance. In [37], laparoscopic skills are assessed in ProMIS tasks before and after LapSim training to clarify whether this training improves operation skills: the results confirm an improvement. In [30], the authors compare physical simulation and AR simulation of central venous catheters mannequin. The results show a significant difference in the adherence level between the AR group and the non-AR group, owing to the real-time feedback the AR group received as they performed the procedure. Additionally, tests of usability, workload, and ergonomics of AR glasses are performed.

Lahanas et al. [33] report both face and construct validity of their simulator. Face validity is evaluated by 20 users (10 novices and 10 expert surgeons) using a questionnaire about the realism of the VR objects and the interaction between the instruments and the VR objects, and the difficulty of the task and the lack of force feedback during tool-object interaction. Construct validity is evaluated in three tasks between two experience groups. The results demonstrate highly significant differences in all performance metrics.

Validity and technical evaluation: Four studies out of 23 report face and content validity and technical evaluation [24,26,41,44].

With regard to face and content validity, all studies use a questionnaire on the realism and the usefulness of the simulators; in [26,44], the questions also address the value of AR for medical training. Overall, the opinions of the users are very positive. Additionally, in [26,44] the authors evaluate the robustness of the simulator hardware. The users enrolled in all studies are five expert surgeons in [26], 13 clinicians in [24], 10 expert surgeons in [44], six novices, and four expert surgeons in [41].

With regard to the technical evaluation, in [26,44] the accuracy of AR visualization is evaluated as adequate for training purposes. In [24], the authors perform three tests to measure eventual damage of EM sensors during embedding steps, the correspondence between planned and actual sensor positions, and the correspondence between real and virtual scenarios; all the results were coherent.

In [41], the authors assess both the simulator reliability measuring the precision of the tracking and the simulator face, content, and construct validity. All users agree on the realism of the simulator and on the fact that practice on the simulator would improve their intraoperative wire navigation performance. About the content validity assessment, the authors compare the desired task characteristics from the design specification checklist with the features of the assembled simulator. Additionally, construct validity results confirm the ability of a simulator to differentiate performance between experts and novices.

With regard to question FQ10 ("Is a statistical analysis performed?"), 16 studies out of 23 conducted a statistical analysis [22–28,30,31,33,34,37,40,41,43,44]. One criticism of these studies is that often the number of participants was small: only five studies involved more than 20 subjects [22,30,34,40,43], so further studies are needed to generalize the findings obtained.

3.10. Trends of Simulation Techniques in Healthcare (SQ1–4)

This review also reports the trends of simulation techniques in healthcare and the types (commercial or custom-made) of simulators used to integrate the chosen approach or to compare two or more different approaches.

The statistical data collected in this section are obtained by the answering the statistical questions SQ1 (“How many commercial simulators are used?”), SQ2 (“How many MR simulators?”), SQ3 (“How many AR simulators?”), and SQ4 (“How many hybrid simulators?”). Figure 2 shows the percentage of studies using commercial and custom-made simulators: it can be seen that 22% used commercial simulators, and 78% custom-made simulators.

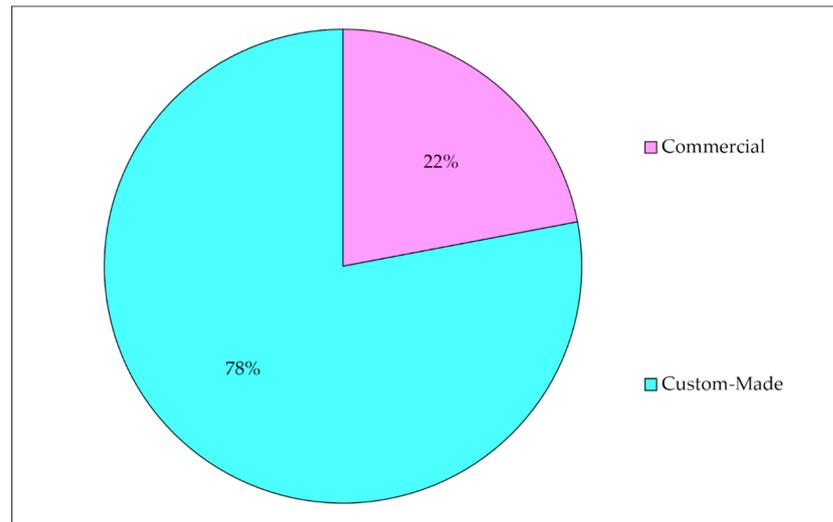


Figure 2. The pie chart shows the percentages of studies using commercial (22%) and custom-made (78%) simulators.

Figure 3 shows the distribution of the approaches used in the implementation of medical simulators: 43% of selected studies combine two approaches (MR and hybrid), 22% combine all three approaches, 26% employed only the hybrid approach, and 9% applied only the MR approach.

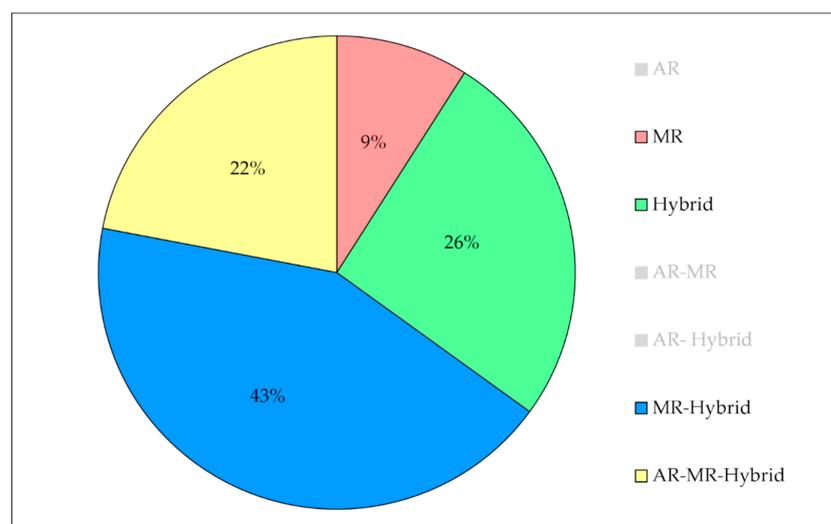


Figure 3. The pie chart shows the distribution of analyzed papers in terms of the approaches used in the implementation of medical simulators. The categories grayed out (augmented reality (AR), AR-mixed reality (MR), and AR-Hybrid) are associated with a percentage of 0%.

4. Discussion

This systematic review investigates the impact of AR, MR, and hybrid approaches on medical simulation and reveals that most of the selected studies (43%) combine MR and hybrid approaches.

Most studies use simulation predominantly as a medical/surgical training tool addressing both technical and non-technical skills. Eight of them use the simulator also as an automatic assessment tool for the evaluation of the user performance. However, given the promising role of simulation in objective skill assessment, we believe that more research is needed to integrate such functionalities in AR, MR, and hybrid simulators.

Furthermore, the review examines how the virtual and physical components are implemented. As for the virtual component, we analyzed the following core elements: the tracking approach for deriving the spatiotemporal relationship between the real and virtual worlds, the display technologies, and the implementation of software haptic feedback and AI techniques.

Tracking methods commonly employed for simulators include the use of EM sensors, vision-based approaches (mainly with planar printable markers), and hybrid methods. The latter use marker-based procedures combined with marker-less or sensor-based methods.

Concerning the visualization modality, most of the revised simulators use a traditional 2D display; six studies report the use of HMDs, and only two a hand-held display.

As for haptic feedback, it is mostly obtained through actual interaction with a physical replica of the anatomy. However, two simulators integrate a commercial haptic interface. More specifically, in one study, interaction with the physical environment is enriched by virtual haptic feedback generated by means of commercial haptic interfaces used to interact with internal organs not reproduced in the mannequin.

Actually, in the literature, AI algorithms are studied but not yet robustly integrated. Indeed, among all the studies analyzed, only four report the use of AI methods. The use of AI techniques is increasing in the development of surgical simulators, as the AI potential is huge for automatic performance evaluation, metrics extraction, simulation level adaptation to the trainee performance, and realistic force feedback implementation. Thus, we believe that this issue deserves future attention for the development of high performance AR, MR, and hybrid simulators.

Among all the studies analyzed, 78% use custom-made simulators. The manufacturing process for most custom-made phantoms starts with extracting 3D models from CT images of real patients. The 3D model is turned into a tangible physical replica using 3D printing technologies and casting fabrication processes. Printing materials such as resin and ABS are used to reproduce rigid parts (i.e., bones), and they adequately approximate the mechanical behavior of the natural tissue. Silicone mixtures and polyurethane materials are used for the manufacturing of soft parts to mimic human tissue properties. Among the revised custom-made phantoms, two include both patient-specific and non-patient specific synthetic organs. In addition, some authors equip the phantoms with EM sensors to implement AR applications. Overall, the manufacturing materials used are extremely durable over time allowing the reusability of the phantom (where is it possible) and thus a reduction in training costs. Given the importance of physical interaction in the skill acquisition for surgery, it is essential in the future both to study deeply the most suitable materials to mimic soft tissues and bones and to validate the realism of the interaction between physical models and surgical instrumentation.

Other aspects investigated by this review include the complexity of the medical scenario, the integration of methods for evaluating clinical performance, and the level of portability of the simulator. Despite the importance of the first two aspects on learning effectiveness, only a few studies have addressed these issues. Regarding portability (a significant factor in the widespread use of medical simulators in educational and training settings), the review shows that most simulators are simple to assemble/disassemble, and some can be easily transported.

Finally, concerning the evaluation of simulators, most articles conduct either a technical or a validity evaluation or, more rarely, both. However, a common limitation of the reviewed studies is the small number of participants (only five studies involved more than 20 subjects [22,30,34,40,43]) recruited to test simulator. More specifically, with regard to the validity assessment, the number of experts enrolled to validate simulators is often small (only two studies involved more than 10 experts [22,30]). Thus, although preliminary results are encouraging, further research is needed to validate existing AR, MR, and hybrid simulators for surgical training and to test whether improvements in performance on a simulated scenario translate into better performance on real patients.

At the time of the literature search, there was no systematic review covering the topic of simulation techniques based on AR, MR, and hybrid approaches in healthcare. Only one review in English [55] was found on augmented reality-based simulators in laparoscopic surgery. More specifically, the authors described five commercial AR simulators in terms of the features of simulators (modules and tested skills, recorded parameters, and feedback), an overview of measurements (need for observer and instructions), the assessment methods of performance, the most important aspects, shortcomings, validity, and costs of the simulators. However, our review provides a wider range of studies on medical simulation techniques not limited to AR but including also MR and hybrid approaches, providing a complete analysis of virtual and physical components of commercial and custom-made simulators and identifying current trends in the choice of simulation approaches.

Author Contributions: Conceptualization, M.G. and V.F.; methodology, R.M.V., S.C., G.T. and M.C.; investigation, R.M.V., S.C., G.T. and M.C.; data curation, R.M.V., S.C., G.T. and M.C.; writing—original draft preparation, R.M.V., S.C., G.T. and M.C.; writing—review and editing, G.T. and S.C.; supervision M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Italian Ministry of Education and Research (MIUR) in the framework of the CrossLab project (Departments of Excellence) of the University of Pisa, Laboratory of Augmented Reality.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kerr, B.; O’Leary, J. The training of the surgeon: Dr. Halsted’s greatest legacy. *Am. Surg.* **1999**, *65*, 1101–1102. [[PubMed](#)]
2. Scott, D.J.; Cendan, J.C.; Pugh, C.M.; Minter, R.M.; Dunnington, G.L.; Kozar, R.A. The changing face of surgical education: Simulation as the new paradigm. *J. Surg. Res.* **2008**, *147*, 189–193. [[CrossRef](#)]
3. Tan, S.S.Y.; Sarker, S.K. Simulation in surgery: A review. *Scott. Med. J.* **2011**, *56*, 104–109. [[CrossRef](#)]
4. Meier, A.H.; Rawn, C.L.; Krummel, T.M. Virtual reality: Surgical application—challenge for the new millennium. *J. Am. Coll. Surg.* **2001**, *192*, 372–384. [[CrossRef](#)]
5. Hutter, M.M.; Kellogg, K.C.; Ferguson, C.M.; Abbott, W.M.; Warshaw, A.L. The impact of the 80-hour resident workweek on surgical residents and attending surgeons. *Ann. Surg.* **2006**, *243*, 864–875. [[CrossRef](#)] [[PubMed](#)]
6. Sutherland, L.M.; Middleton, P.F.; Anthony, A.; Hamdorf, J.; Cregan, P.; Scott, D.; Maddern, G.J. Surgical simulation: A systematic review. *Ann. Surg.* **2006**, *243*, 291–300. [[CrossRef](#)]
7. The Southern Surgeons Club; Moore, M.J.; Bennett, C.L. The learning curve for laparoscopic cholecystectomy. *Am. J. Surg.* **1995**, *170*, 55–59. [[CrossRef](#)]
8. Scott, D.J. Patient Safety, Competency, and the Future of Surgical Simulation. *Simul. Healthc.* **2006**, *1*, 164–170. [[CrossRef](#)]
9. Stefanidis, D.; Sevdalis, N.; Paige, J.; Zevin, B.; Aggarwal, R.; Grantcharov, T.; Jones, D.B. Simulation in surgery: What’s needed next? *Ann. Surg.* **2015**, *261*, 846–853. [[CrossRef](#)]
10. Vozenilek, J.; Huff, J.S.; Reznick, M.; Gordon, J.A. See One, Do One, Teach One: Advanced Technology in Medical Education. *Acad. Emerg. Med.* **2004**, *11*, 1149–1154. [[CrossRef](#)]
11. Bradley, P. The history of simulation in medical education and possible future directions. *Med. Educ.* **2006**, *40*, 254–262. [[CrossRef](#)] [[PubMed](#)]
12. Gaba, D.M. The future vision of simulation in healthcare. *Simul. Healthc.* **2007**, *2*, 126–135. [[CrossRef](#)] [[PubMed](#)]
13. Lateef, F. Simulation-based learning: Just like the real thing. *J. Emerg. Trauma. Shock.* **2010**, *3*, 348–352. [[CrossRef](#)]

14. Ziv, A.; Wolpe, P.R.; Small, S.D.; Glick, S. Simulation-based medical education: An ethical imperative. *Acad. Med.* **2003**, *78*, 783–788. [[CrossRef](#)] [[PubMed](#)]
15. Ziv, A.; Small, S.D.; Wolpe, P.R. Patient safety and simulation-based medical education. *Med. Teach.* **2000**, *22*, 489–495. [[CrossRef](#)] [[PubMed](#)]
16. de Visser, H.; Watson, M.O.; Salvado, O.; Passenger, J.D. Progress in virtual reality simulators for surgical training and certification. *Med. J. Aust.* **2011**, *194*, S38–S40. [[CrossRef](#)] [[PubMed](#)]
17. Gurusamy, K.; Aggarwal, R.; Palanivelu, L.; Davidson, B. Systematic review of randomized controlled trials on the effectiveness of virtual reality training for laparoscopic surgery. *Br. J. Surg.* **2008**, *95*, 1088–1097. [[CrossRef](#)]
18. Loukas, C. Surgical Simulation Training Systems: Box Trainers, Virtual Reality and Augmented Reality Simulators. *Int. J. Robot. Autom.* **2016**, *1*, 1–9. [[CrossRef](#)]
19. Badash, I.; Burt, K.; Solorzano, C.A.; Carey, J.N. Innovations in surgery simulation: A review of past, current and future techniques. *Ann. Transl. Med.* **2016**, *4*, 453. [[CrossRef](#)]
20. Dunkin, B.; Adrales, G.L.; Apelgren, K.; Mellinger, J.D. Surgical simulation: A current review. *Surg. Endosc.* **2007**, *21*, 357–366. [[CrossRef](#)] [[PubMed](#)]
21. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; The, P.G. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med.* **2009**, *6*, e1000097. [[CrossRef](#)] [[PubMed](#)]
22. Coelho, G.; Rabelo, N.N.; Vieira, E.; Mendes, K.; Zagatto, G.; Santos de Oliveira, R.; Raposo-Amaral, C.E.; Yoshida, M.; de Souza, M.R.; Fagundes, C.F.; et al. Augmented reality and physical hybrid model simulation for preoperative planning of metopic craniosynostosis surgery. *Neurosurg. Focus* **2020**, *48*, E19. [[CrossRef](#)]
23. Condino, S.; Turini, G.; Parchi, P.D.; Vigliani, R.M.; Piolanti, N.; Gesi, M.; Ferrari, M.; Ferrari, V. How to Build a Patient-Specific Hybrid Simulator for Orthopaedic Open Surgery: Benefits and Limits of Mixed-Reality Using the Microsoft HoloLens. *J. Healthc. Eng.* **2018**, *2018*, 5435097. [[CrossRef](#)] [[PubMed](#)]
24. Condino, S.; Carbone, M.; Ferrari, V.; Faggioni, L.; Peri, A.; Ferrari, M.; Mosca, F. How to build patient-specific synthetic abdominal anatomies. An innovative approach from physical toward hybrid surgical simulators. *Int. J. Med. Robot. Comput. Assist. Surg.* **2011**, *7*, 202–213. [[CrossRef](#)]
25. Feifer, A.; Al-Ammari, A.; Kovac, E.; Delisle, J.; Carrier, S.; Anidjar, M. Randomized controlled trial of virtual reality and hybrid simulation for robotic surgical training. *BJU Int.* **2011**, *108*, 1652–1656, discussion 1657. [[CrossRef](#)] [[PubMed](#)]
26. Ferrari, V.; Vigliani, R.M.; Nicoli, P.; Cutolo, F.; Condino, S.; Carbone, M.; Sesto, M.; Ferrari, M. Augmented reality visualization of deformable tubular structures for surgical simulation. *Int. J. Med. Robot.* **2016**, *12*, 231–240. [[CrossRef](#)]
27. Fuerst, D.; Hollensteiner, M.; Schrempf, A. A novel augmented reality simulator for minimally invasive spine surgery. In Proceedings of the 2014 Summer Simulation Multiconference, Monterey, CA, USA, 6–10 July 2014; p. 28.
28. Fushima, K.; Kobayashi, M. Mixed-reality simulation for orthognathic surgery. *Maxillofac. Plast Reconstr. Surg.* **2016**, *38*, 13. [[CrossRef](#)] [[PubMed](#)]
29. Halic, T.; Kockara, S.; Bayrak, C.; Rowe, R. Mixed reality simulation of rasping procedure in artificial cervical disc replacement (ACDR) surgery. *BMC Bioinform.* **2010**, *11* (Suppl. 6), S11. [[CrossRef](#)]
30. Huang, C.Y.; Thomas, J.B.; Alismail, A.; Cohen, A.; Almutairi, W.; Daher, N.S.; Terry, M.H.; Tan, L.D. The use of augmented reality glasses in central line simulation: “see one, simulate many, do one competently, and teach everyone”. *Adv. Med. Educ. Pract.* **2018**, *9*, 357–363. [[CrossRef](#)]
31. Jain, S.; Lee, S.; Barber, S.R.; Chang, E.H.; Son, Y.-J. Virtual reality based hybrid simulation for functional endoscopic sinus surgery. *IIEE Trans. Healthc. Syst. Eng.* **2020**, *10*, 127–141. [[CrossRef](#)]
32. Keebler, J.R.; Lazzara, E.H.; Patzer, B. Building a Simulated Medical Augmented Reality Training System. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2014**, *58*, 1169–1173. [[CrossRef](#)]
33. Lahanas, V.; Loukas, C.; Smailis, N.; Georgiou, E. A novel augmented reality simulator for skills assessment in minimal invasive surgery. *Surg. Endosc.* **2015**, *29*, 2224–2234. [[CrossRef](#)] [[PubMed](#)]
34. Lee, S.; Lee, J.; Lee, A.; Park, N.; Lee, S.; Song, S.; Seo, A.; Lee, H.; Kim, J.I.; Eom, K. Augmented reality intravenous injection simulator based 3D medical imaging for veterinary medicine. *Vet. J.* **2013**, *196*, 197–202. [[CrossRef](#)]
35. Lin, Y.-K.; Tsai, K.-L.; Yau, H.-T. The Development of Optical See-through Display Based on Augmented Reality for Oral Implant Surgery Simulation. *Comput. Aided Des. Appl.* **2012**, *9*, 111–120. [[CrossRef](#)]
36. Loukas, C.; Lahanas, V.; Georgiou, E. An integrated approach to endoscopic instrument tracking for augmented reality applications in surgical simulation training. *Int. J. Med. Robot.* **2013**, *9*, e34–e51. [[CrossRef](#)]
37. Nomura, T.; Mamada, Y.; Nakamura, Y.; Matsutani, T.; Hagiwara, N.; Fujita, I.; Mizuguchi, Y.; Fujikura, T.; Miyashita, M.; Uchida, E. Laparoscopic skill improvement after virtual reality simulator training in medical students as assessed by augmented reality simulator. *Asian J. Endosc. Surg.* **2015**, *8*, 408–412. [[CrossRef](#)]
38. Onishi, K.; Mizushino, K.; Ikemoto, H.; Noborio, H. AR Dental Surgical Simulator Using Haptic Feedback. In Proceedings of the International Conference on Human-Computer Interaction HCI 2013, Las Vegas, NV, USA, 21–26 July 2013; Springer: Berlin/Heidelberg, Germany, 2013; pp. 202–205.
39. Parkes, R.; Forrest, N.; Baillie, S. A mixed reality simulator for feline abdominal palpation training in veterinary medicine. *Stud. Health Technol. Inform.* **2009**, *142*, 244–246. [[PubMed](#)]

40. Yonghang, T.; Wei, L.; Zhou, H.; Peng, J.; Li, Q.; Li, F.; Zhang, J.; Shi, J. Augmented-reality-driven medical simulation platform for percutaneous nephrolithotomy with cybersecurity awareness. *Int. J. Distrib. Sens. Netw.* **2019**, *15*. [[CrossRef](#)]
41. Thomas, G.W.; Johns, B.D.; Kho, J.Y.; Anderson, D.D. The Validity and Reliability of a Hybrid Reality Simulator for Wire Navigation in Orthopedic Surgery. *IEEE Trans. Hum. Mach. Syst.* **2015**, *45*, 119–125. [[CrossRef](#)]
42. Tsujita, T.; Sase, K.; Chen, X.; Tomita, M.; Konno, A.; Nakayama, M.; Nakagawa, A.; Abe, K.; Uchiyama, M. Development of a Surgical Simulator for Training Retraction of Tissue with an Encountered-Type Haptic Interface Using MR Fluid. In Proceedings of the 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO 2018), Kuala Lumpur, Malaysia, 12–15 December 2018; pp. 898–903.
43. van Duren, B.H.; Sugand, K.; Wescott, R.; Carrington, R.; Hart, A. Augmented reality fluoroscopy simulation of the guide-wire insertion in DHS surgery: A proof of concept study. *Med. Eng. Phys.* **2018**, *55*, 52–59. [[CrossRef](#)]
44. Viglialoro, R.M.; Esposito, N.; Condino, S.; Cutolo, F.; Guadagni, S.; Gesi, M.; Ferrari, M.; Ferrari, V. Augmented Reality to Improve Surgical Simulation: Lessons Learned Towards the Design of a Hybrid Laparoscopic Simulator for Cholecystectomy. *IEEE Trans. Biomed. Eng.* **2019**, *66*, 2091–2104. [[CrossRef](#)]
45. Lynch, A. Simulation-based acquisition of non-technical skills to improve patient safety. *Semin. Pediatric Surg.* **2020**, *29*, 150906. [[CrossRef](#)]
46. Agha, R.A.; Fowler, A.J.; Sevdalis, N. The role of non-technical skills in surgery. *Ann. Med. Surg.* **2015**, *4*, 422–427. [[CrossRef](#)]
47. Escobar-Castillejos, D.; Noguez, J.; Neri, L.; Magana, A.; Benes, B. A Review of Simulators with Haptic Devices for Medical Training. *J. Med. Syst.* **2016**, *40*, 104. [[CrossRef](#)] [[PubMed](#)]
48. Turini, G.; Condino, S.; Parchi, P.D.; Viglialoro, R.M.; Piolanti, N.; Gesi, M.; Ferrari, M.; Ferrari, V. A Microsoft HoloLens Mixed Reality Surgical Simulator for Patient-Specific Hip Arthroplasty Training. In Proceedings of the International Conference on Augmented Reality, Virtual Reality and Computer Graphics (AVR 2018), Otranto, Italy, 24–27 June 2018; Springer: Cham, Switzerland, 2018; pp. 201–210.
49. Ryu, W.H.A.; Dharampal, N.; Mostafa, A.E.; Sharlin, E.; Kopp, G.; Jacobs, W.B.; Hurlbert, R.J.; Chan, S.; Sutherland, G.R. Systematic Review of Patient-Specific Surgical Simulation: Toward Advancing Medical Education. *J. Surg. Educ.* **2017**, *74*, 1028–1038. [[CrossRef](#)]
50. Issenberg, B.; McGaghie, W.; Petrusa, E.; Gordon, D.; Scalese, R. Features and uses of high-fidelity medical simulations that lead to effective learning: A BEME systematic review. *Med. Teach.* **2005**, *27*, 10–28. [[CrossRef](#)] [[PubMed](#)]
51. Cox, M.; Irby, D.; Epstein, R. Assessment in Medical Education. *N. Engl. J. Med.* **2007**, *22*, 13–16.
52. Ryall, T.; Judd, B.K.; Gordon, C.J. Simulation-based assessments in health professional education: A systematic review. *J. Multidiscip. Healthc.* **2016**, *9*, 69–82. [[PubMed](#)]
53. Bewley, W.; Oneil, H. Evaluation of Medical Simulations. *Mil. Med.* **2013**, *178*, 64–75. [[CrossRef](#)]
54. Moglia, A.; Ferrari, V.; Morelli, L.; Ferrari, M.; Mosca, F.; Cuschieri, A. A Systematic Review of Virtual Reality Simulators for Robot-assisted Surgery. *Eur. Urol.* **2016**, *69*, 1065–1080. [[CrossRef](#)]
55. Botden, S.M.; Jakimowicz, J.J. What is going on in augmented reality simulation in laparoscopic surgery? *Surg. Endosc.* **2009**, *23*, 1693–1700. [[CrossRef](#)] [[PubMed](#)]