

# Face, Content, and Construct Validity Evaluation of Simulation Models in General Surgery Laparoscopic Training and Education: A Systematic Review.

**Roberto Javier Rueda-Esteban** (✉ [rj.rueda32@uniandes.edu.co](mailto:rj.rueda32@uniandes.edu.co))

Universidad de Los Andes Facultad de Medicina; Escola de Doctorat Universitat Autònoma de Barcelona <https://orcid.org/0000-0002-5957-6261>

**Juan Sebastian Lopez-McCormick**

Universidad de Los Andes Facultad de Medicina

**Stefania Rodriguez-Bermeo**

Universidad de Los Andes Facultad de Medicina

**Mateo Andrade**

Universidad de Los Andes

**Juan David Hernández**

Fundacion Santa Fe de Bogotá: Fundacion Santa Fe de Bogota

**Eduardo M Targarona**

Hospital de Sant Pau: Hospital de la Santa Creu i Sant Pau

---

## Research Article

**Keywords:** Validation; Face, content and construct Validity; Laparoscopic training; Education Training; Virtual reality; Augmented reality.

**Posted Date:** March 28th, 2022

**DOI:** <https://doi.org/10.21203/rs.3.rs-1436353/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

## Abstract

**Background:** Laparoscopic technical surgical skills (LTS) are considered a fundamental competence for General Surgery residents. Several simulation tools (ST) have been explored to develop these skills. Although a plethora of systematic reviews evaluate the translation of LTS developed in simulation to real surgery, there is a lack of evidence that clarifies effectiveness of different validated ST in acquisition of LTS in surgical residents. The aim of this systematic review (SR) is to summarize current published evidence on ST validation used for general surgery education and training.

**Methods:** An initial protocol was published in PROSPERO, under registration number: CRD42020205397. A SR of the literature was carried out following PRISMA guidelines. Complete published articles in English or Spanish that validate either content or construct, plus another form of validation of ST to acquire LTS in general surgery were included. Articles that used only one validation strategy or did not validate an ST were excluded.

**Results:** 1052 publications were initially identified across all searched databases. Title review identified 204 studies which were eligible for full text screening. 10 studies were included for final review. Two studies assessed both face and content, four face and construct, and four face, content and construct validity. None of the studies presented comparable outcomes due to metrics variation and scores used for the validation strategies.

**Conclusions:** This study assessed validated laparoscopic simulation models, with particular interest in content and construct validity. Articles reported an increased use of simulation models in laparoscopic training with positive feedback from trainees, but few studies reported validation of training model. Validation strategies are not standardized, limiting comparability between them.

## Background

Development of technical laparoscopic surgical skills (LTS) in General Surgery residents is of paramount importance (1). To develop laparoscopic ability in trainees, multiple tools have been implemented, such as video games (2), virtual reality (VR) training models, box simulators (BS), surgical simulation (SS) on synthetic, virtual, biological models, and human cadaveric models which have shown to be an effective educational method for surgeons (3).

While tools such as video games have had positive correlation with the development of technical skills in the operating room (OR), training programs currently focus on models that allow simulation and training of laparoscopic movements (2). These training models can be further divided into laparoscopy simulation models and SS models, depending on whether aspect simulated is the technique (simple laparoscopic movements) or procedure (emulating a surgical procedure on a synthetic or biological model) (2, 4).

From available training models to develop LTS, SS has been heavily favoured due to evidence of superiority given by shortening of the learning curve for LTS acquisition, improvement of technical dexterity, and faster development of skills that are necessary to perform these types of procedures (4).

Most laparoscopic simulations are done using VR, integrating visual and haptic equipment; synthetic models, live or cadaveric animal models, and human cadaveric specimens preserved by freezing are other popular ST currently in use (2, 4–7). The earliest laparoscopic simulators to be widely utilized were known as training boxes, or BS; nowadays most recent models of BS are used for training basic coordination skills or as casings to simulate intraabdominal environment in SS. All mentioned simulation tools are used to train skills through simple and standardized exercises and repetition (7), with the objective of strengthening basic skills related to precision and movement of objects in space, which are fundamental for any laparoscopic procedure. New simulation models have evolved from BS, adding synthetic or biological components to enhance the experience by adding SS to basic dexterity exercises (8, 9).

SS in VR offers an additional opportunity to measure skills of students and evaluate their progress objectively by evaluating time, number of movements, errors committed, amongst other variables (2, 4–7, 10). Mechanical VR simulators or haptic simulators allow training of the resistance and strength sensation required (11). Nevertheless, due to the diversity and heterogeneity of available models, there is no clear comparison between ST models validation results reported in the literature (4, 12).

Although multiple systematic reviews evaluate the degree in which skills acquired in simulated models can be transferred to a real-life scenario (3–5) or compare the reported effectiveness of specific ST (13, 14), to the best of our knowledge there is no current literature that assesses available evidence on the effectiveness of simulation between different ST through face, content, and construct validation for acquisition of LTS in surgery residents.

These validation strategies are of particular interest amongst a variety of available options (15) when evaluating and validating a model due to their aims: face validity considers the realism and difficulty level similarity in comparison to real tasks, content validity assesses the model's effectiveness during a specific skill training to improve participants' technique and construct validity evaluates the ability of a simulator to differentiate between different levels of expertise (16). This validation strategies, when combined, could determine a model's effectiveness on training (17). As such, the aim of this SR is to summarize current published evidence on ST used for general surgery education and training, validated through two validation strategies between either face, content or context.

## Materials And Methods

### Search Strategy and Assessment of eligibility:

An initial protocol was published in PROSPERO (Registration code: CRD42020205397) and based on it a SR was undertaken in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (18) guidelines up until October 2021 using predefined search terms. Complete published articles in English or Spanish that validate either content or construct, plus another form of validation of ST to acquire LTS in general surgery were

included. Articles that used only one validation strategy or those who presented an ST but did not validate the model, evaluated an ST for other specialties or did not include general surgery residents were excluded.

## Data Extraction:

Databases reviewed included PubMed (Medline), Cochrane, Google Scholar, EMBASE, SciELO and the references of systematic reviews found during the search process were searched for additional manuscripts. The review was updated up to October 2021. The search strategy was replicated across all databases, MeSH and non-MeSH terms were used as part of search algorithm:

("laparoscopy"[MeSH Terms] OR "laparoscopy"[Title] OR "laparoscopic"[Title] OR "minimally invasive"[Title]) AND ("simulation training"[MeSH Terms] OR "box trainer"[Title] OR "augmented reality simulat\*"[Title] OR "virtual reality simulat\*"[Title] OR "virtual patient\*"[Title] OR "training model"[Title] OR "cadaver model\*"[Title]).

Articles obtained were imported to Rayyan Systems Inc (19) online platform. Two reviewers independently reviewed articles obtained by search strategy in a blind extraction by title and abstract; once done, reviewers discussed any conflicts, and a third peer had final decision on disagreements, after which a full text review was carried out. Full papers in conflict between reviewers were discussed, no third-party mediation was required at this point. Data extraction was carried out independently by reviewers to an Excel template previously designed with data of interest (Annex 1).

## Validation:

In the context of simulation for medical education, validation is considered a complex concept that aims to establish that a proposed simulation model accurately achieves its training purpose and is comprised of a series of principles including face, content, construct, concurrent, discriminate, and predictive validity amongst others (15). To assess ST that provided objective evidence of their utility to develop LTS and enhance reliability, only ST which were evaluated for either content or context validity and included at least another form of validation were included. Validity principles appraised in this study were defined as follows:

**Construct validity:** The degree in which the model can identify the ability it is designed to simulate (15), in simulation this is represented by the capability of the ST to differentiate between experts and novices in a specific task (20).

**Content validity:** A detailed evaluation of the content of simulation to determine if its content is appropriate (21). In context of ST for medical education, it assesses usefulness of the ST as a training tool.

**Face Validity:** Face validity describes the realism of a simulator and difficulty level similarity in comparison to real training tasks (22).

## Risk of Bias:

PRISMA guidelines suggest Risk of bias should be assessed using Cochrane's RoB2 (23), however this tool is exclusive for clinical trials, and it is not appropriate for medical education studies, for that reason Medical Education Research Study Quality Instrument (MERSQI) (24), which better suits the evaluation of quality in education research, was used.

MERSQI is a measurement score designed in 2007, validated and widely used to evaluate methodological quality of medical education research (24). It consists of ten items, reflecting six domains of study quality: Study design, sampling, type of data, validity, data analysis and outcomes; each of which carries a maximum score of 3; maximum possible MERSQI score is 18 and has potential range from 5 to 18. Higher MERSQI scores have been associated with a higher acceptance rate to competitive journals and increased likelihood of external funding (24).

## Results

The search strategy produced 1,052 publications across all searched databases. Title review of the manuscripts using selection criteria identified 204 studies which were eligible for full text screening, of which 10 studies were included in the final review, results of the search and screening are summarized in Fig. 1. General characteristics of the study, simulation type, and the ST's characteristics of validity are summarized in Table 1.

Initial article selection excluded 812 of 1052 studies found. Reasons for exclusion were: 1) No validation of the model was included in the study (340 articles excluded); 2) Subjects in the study did not include General Surgery Residents (384 articles excluded); 3) Articles not in English or Spanish, or those that did not pertain to laparoscopic surgical simulation training (88 articles were excluded).

A total of 240 full texts were screened, 156 papers were excluded due to either a lack of, incomplete or not properly described validation strategy, (FIG. 1 Reason 4) (146 articles were excluded); or wrong outcome (FIG. 1 Reason 5) (10 articles were excluded). Papers that did not have at least two different validation strategies, at least one of them being content or construct validity were excluded as well (FIG. 1 Reason 7) (74 articles were excluded). Finally, articles were clustered for analysis by validation strategies employed: face and content, face and construct and, face, content and construct.

## Face and content validity

Two articles validated ST by face and content (25, 26) both presented a BS training model. **Bahsoun et al.** proposed a cardboard-based model which replaced traditional laparoscopic camera with a tablet (25) included 10 participants, and face validation was done by experienced laparoscopic surgeons that answered an evaluation form for tablet Feedback. For content validation the authors looked at quality of feed and compared it with the laparoscopic camera and stack. This study researched components of a laparoscopic training setup and recreated a simplified, less expensive model based on the use of a tablet, to overcome the obstacle of laparoscopic stack and camera costs (25).

**Achurra et al.** presented a high-definition BS, (26) in which 76 expert bariatric surgeons answered a four-item questionnaire for face validity, result showed that 98% considered that BS had a high fidelity and was optimum for advanced laparoscopic simulation. Content validity was evaluated by assessing improvements on global and specific rating scores of 15 first-year general surgery residents. Model evidenced a statistically significant improvement in scores after training with the BS (26).

## Face and construct validity:

A total of four articles that assessed face and construct validity were included. **De Moura Júnior et al.** (27) described a BS for laparoscopic surgery in a fiberglass-based console with 37 participants including both surgeons and general surgery residents with no further stratification, assessed both face and construct by means of self-administered surveys which tested both satisfaction with model as well as perception on performance and found consistently higher scores on model evaluated when compared with a standard laparoscopic trainer. Authors conclude that the model is a useful tool in the development and evaluation of residents in training (27).

Another three authors presented ST based on VR with different variations, **Sánchez-Peralta et al.** described a model based in SINERGIA (28) laparoscopic virtual reality simulator, participants were divided into novice and experts. Both groups positively rated VR simulator as similar to reality, and a statistically significant difference was found in at least one of the metrics for each task between groups, providing construct validity (28).

**Lahanas et al.** proposed an augmented reality simulator for skills assessment in minimal invasive surgery (29) tasks evaluated included instrument navigation, peg transfer and clipping. Participants were divided in two groups, novices and experienced surgeons. Construct validity was evaluated by a statistical comparison of two experience groups while performing the three tasks described previously (29). All participants answered a five-point Likert questionnaire regarding the realism and difficulty of each task obtaining face validity of the instrument.

Finally, **Sankaranarayanan et al.** presented a ST consisting of a VR simulator called Gen2-VR (30), participants had a different experience between them and were not stratified, the study compared the use of the model under different scenarios with an older VR model. Face validity was assessed by means of a five-point Likert-scale, while a statistically significant difference was found between scenarios and VR models providing construct validity (30). This ST is of particular interest as it includes the assessment of factors that may interfere with surgery outcomes such as noise and interruptions which can be added to immersive VR environment and should be considered while training is performed (30).

## Face, content, and construct validity

Four articles amongst all the publications had all three: face, content and construct validity, two of them (31, 32) proposed BS training models while the other two presented VR models (33, 34). **Van Empel et al.** proposed a BS model named TrEndo (31), participants were divided into two groups based on prior laparoscopic experience, novices and experts. Face and content validities were assessed by a questionnaire and construct validity was appraised by a series of tasks such as needle position and executing a square knot. Results for Content validity showed experts and trainees rated procedural functioning, hand-eye coordination, and depth perception above 3 on the five-point Likert-scale (31). Construct validity showed differences between the two groups, with significant better results in the experts' group.

**Perez Escamiroso et al.** presented a ST called EndoViS (32), a training system for psychomotor skills development in laparoscopic surgeons, sample groups were divided by expertise levels (medical students, residents and surgeons), face and content validity were assessed by a questionnaire of 13 statements that combined design, realism of cavity, functionality of simulator and training capacities, answered with a five-point Likert-scale (32), construct validity was measured by evaluation of four skill tasks peg transfer, rubber band, pattern cutting, and intracorporeal knot suture.(32) Authors of this article stated that EndoViS training system has been presented and successfully validated. They proposed that the system provides a non-obstructive alternative to the traditional tracking systems and a reliable method to capture and analyze motion of surgical instruments offering potential for surgical training programs (32).

**Kawaguchi et al.** presented a ST named LAP-X to train basic laparoscopic skills, for this paper groups were divided by novice and experts, all participants completed a questionnaire designed to evaluate the face and content validity, content validity was evaluated by experts only, questions were answered by selecting a mark on an ordinal scale ranging from 1 to 5 and construct validity was tested by parameters of a task completed using the VR simulator (33).

**Arts et al.** Presented an augmented reality VR model called EoSim (34). Participants were divided in experienced, intermediate and inexperienced. To assess face and content validity, each participant completed a questionnaire comprised of a five-point Likert-scale on realism, didactic value, and usability, construct validity was assessed using the outcome parameters generated by the tracking software. Construct, face and content validity are presented for the ST (34).

## Methodological Quality and Risk of bias

PRISMA guidelines Risk of bias using Cochrane's RoB2 (23) was not conducted due the educational research nature of the included manuscripts. Methodological quality of the studies was reviewed, classified, and scored using MERSQI scoring system (24). MERSQI scores were compiled for each domain and totalled to allow comparisons between papers; Information was summarized in an Excel template, with a Confidence Level of 95,0% resulting in a mean of 11,69 (CI 95% 11,40 - 11,98). Individual scores can be found at Table 1.

## Discussion

Evidence provided by educational papers and validity definitions suggests that ST for LTS development should be validated (15, 18, 35). Our Systematic Review highlights evidence of feasibility, utility, and importance of multiple validation strategies in development and presentation of laparoscopic simulation models. Heterogeneity between papers is rampant, and widely explained due to unavailability of standardized validation strategies for simulation models in

surgical education. One of the most vital difficulties lies in the fact that terms utilized across studies are not standardized, such as conflicting and incomplete information on participant experience levels. Terms such as “novice” are not punctually defined in literature; authors take novices as medical students, first year residents or any year residents, even though these represent extremely different populations with high variability in informal training and skill acquisition. Furthermore, in some cases they provide no data regarding previous training or laparoscopic experience when differentiating experience levels, thus increasing the difficulty in cross study comparisons of ST.

Reviewed literature showed that most validated simulations are carried out on VR, using both visual and haptic equipment. Other ST currently in use include synthetic models, live or cadaveric animal models, and human cadaveric specimens preserved by freezing (2, 4–7), all these simulation tools are used to train skills through standardized exercises and repetition (7), offering the opportunity to objectively measure the skills and evaluate progress (10), and allowing the training of resistance and strength sensation required for laparoscopic surgery (11).

While previous studies suggest residents prefer training on frozen human or animal cadaveric models (12), no validation of these ST was found in the reviewed literature, even when it was the preferred model between students, high costs and difficulties in preservation of cadaveric specimens are strong drawbacks that may contribute to scarce literature objectively evaluating their usefulness. Multiple systematic reviews focus on the effectiveness of specific ST's such as virtual or cadaveric simulation (13, 15), but don't compare results across different ST's.

Most authors attempted to assess face validity; however, no standardized validated score exists for this purpose, questionnaires and scores tend to vary, even questions on surveys are not comparable and thus represent a source of heterogeneity. Due to the high variability of questionnaires used, and in the absence of a standardized strategy to validate ST, comparison between papers is not objectively possible at this time.

All studies evidenced some measure of improvement in LTS when using simulation, which is supported by current literature and suggests simulation training in laparoscopy is useful to develop LTS in residents (1–3). Simulation exercises allow students to practice in a controlled environment, giving a safe space to make mistakes while receiving feedback from teachers and experienced clinicians, thus reducing risk of errors in future procedures (36). As a requirement of general surgery residency programs, laparoscopic simulation of any kind should certainly be included (1), ideally before residents perform any surgery in patients (4, 16).

As stated previously, it is not possible to determine which ST has a better performance for general surgery residents due to the heterogeneity of models and lack of standardized definitions, measuring tools and validation strategies, which limit comparability between them (4, 35). Even when the quality of individual studies was moderate, as evidenced by the MERSQI tool, these differences in methodological procedure prohibit adequate comparison between the different ST's. Despite this limitation, improvement has been described in LTS acquisition after simulation, regardless of the model, methods for validation and validation type which reinforces the need for simulation during training in surgery (1, 4, 5, 6, 9, 21, 27–32, 34–36).

## Face and content validity

Both studies evaluated BS models, Bahsoun et al. (25) and Achurra et al. (26). Although they had a similar ST, there were substantial differences between them, particularly in the subjects included for each study. Due to a lack of clarification on expertise levels, study populations were not comparable. While Bahsoun et al. described the study population as “residents and experts”, Achurra et al. defined them as “novices and experts”, furthermore each author used different selection criteria to classify the participants expertise. Face validation for these ST was performed in both studies with non-standardized questionnaires developed for the singularities of each BS model, which in accordance with literature, a lack of standardized and validated questionnaires can be a confusion factor in multiple scenarios (36).

## Face and construct validity

Four studies assessed both Face and construct validity. In this sub-group most ST were VR, consistent with current major prevalence of this type of simulations for laparoscopy training (2, 4–7). De Moura Júnior et al.(27) presented a BS model, face validity was assessed by means of a self-administered survey which tested both satisfaction with the model as well as perception on performance. Sánchez-Peralta et al.(28) and Sankaranarayanan et al. (30) presented a different VR model. The main distinction between those two models, was that Sankaranarayanan et al. included assessment of environmental factors that may interfere with surgery outcomes such as operating room noise and interruptions which can be included in the immersive VR environment and should be considered while training is performed, giving this model an interesting additional perspective.

Lahanas et al.(29) proposed an augmented reality simulator, in which virtual objects were superimposed on reality. Participants were described as either novices with no previous laparoscopic experience, or experts but no further information was provided. It is the only study that proposed an augmented reality simulator (29).

## Face, content, and construct validity

Only four of included studies assessed all three: face, content, and construct validity. Two assessed VR ST's and two BS's. Van Empel et al.(21), Pérez Escamiroso et al. (32), Kawaguchi et al.(33) and Arts et al. (34). Van Empel et al.(21) assessed face and content validation by means of a questionnaire while construct validity was appraised by a series of tasks. Perez Escamiroso et al.(32) presented a ST and assessed face and content validity by means of a questionnaire comprised of 13 statements answered with a five-point Likert-scale, construct validity was measured by skill performance in four tasks.

Kawaguchi et al. (33) assessed face and content validities with a questionnaire designed by authors, questions were answered by selecting a mark on an ordinal scale ranging from 1 to 5 and content validity was evaluated by experts only. Construct validity was tested by parameters of a task completed using the VR simulator. Arts et al. (34) described an augmented reality VR model and assessed face and content validity through a questionnaire comprised of a five-point Likert-scale, while construct validity comprised the outcome parameters generated by a tracking software.

As already stated, an important variation on Likert-scales and questionnaire design is apparent in the included studies, which hinders data analysis and comparability of results (36).

## COVID-19 and Simulation Models:

Development of novel learning modalities such as synthetic models and VR, have been accelerated during the COVID-19 pandemic (37). However, many simulation models are nascent in development and represent high costs for their application (16). Currently BS and VR models are the most prevalent simulation tools available, BS models are favoured due to their visual and haptic realism while maintaining relatively low costs. Rapid advent of technology may expand on the possibility of VR models using haptic feedback while lowering development and usage costs.

## Strengths and Limitations

As already stated in the literature, limitations in this type of reviews are related to restrictions in medical education research, ethical barriers and the fact that no gold standard has been described to assess surgical training outcomes (16, 38).

A major limitation of this review is the lack of inclusion of predictive validity, considered an “ultimate” assessment to establish validity of an educational tool (16, 36). Nevertheless, it remains difficult to implement in practical terms. Also, the inability to compare data acquired by diverse researchers due to heterogeneity of gathering tools or variations of Likert-scales hinders the possibility to acquire strong evidence about the use of simulation in different contexts.

The fact that only manuscripts which considered two types of validation with at least one being content, or context excluded many partially validated ST models which are being used worldwide. And the specific population being general surgery residents also diminished multiple simulation scenarios that could be of importance.

Furthermore, research participants definition and selection for our review limited studies considering exclusively medical students, which were excluded, nevertheless several manuscripts validated content by assessing improvement in performance in laparoscopically “naive” subjects, defined as undergraduate students (39).

However, according to our knowledge, this review is the first to address face, content and construct validity for laparoscopic surgery training ST validation. A recent SR published by Shah et al. (40) aimed to describe the status of simulation-based training tools in general surgery in the current literature, assess their validity and determine their effectiveness (40); that systematic review may be complementary to ours. While Shah et al. included all published simulation models and described a variety of validations proposed by authors, no distinction was made to highlight and review models with at least two kinds of validation, one of them being either content or construct validity which have been previously established validations that support validity of ST as training tools to effectively develop LTS in surgical education (15).

Future work should aim to standardize scales for validation which incorporates both subjective and objective validation assessments such as the GOALS and OSATS scores for laparoscopic training. This will allow researchers to collect adequate data to validate multiple models and, in the future, compare results between published ST's.

## Conclusion

This SR assessed laparoscopic simulation models that were validated, with particular interest in content and construct validity. Articles reported an increased use of simulation models in laparoscopic training. Although studies demonstrated that laparoscopic surgery trainees preferred the use of frozen human or animal cadaveric models, synthetic and BS models are currently the most convenient and practical for developing skills in residents, VR models provided visual realism and improved haptic feedback, but still need future development.

Although surgical simulation models receive positive feedback from participants, proper validation of the models is not standardized. For that reason, comparing results from different publications is limited by the methodological heterogeneity amongst articles, even when individually articles are well executed. Moreover, studies examining simulation methods seldom use objective validation methods such as construct and content validity. Future work should examine how to facilitate and standardize objective validation strategies that are applicable to a variety of ST's.

## Abbreviations

LTS  
Laparoscopic technical surgical skills  
ST  
Simulation tools  
SR  
Systematic review  
VR  
Virtual reality  
BS  
Box simulators  
SS  
surgical simulation

OR  
operating room  
PRISMA  
Preferred Reporting Items for Systematic Reviews and Meta-Analyses

## Declarations

### Ethical approval

IRB approval was acquired under endorsement **No. 201910224** by Ethics Committee of investigation, Universidad de los Andes School of Medicine.

### Consent for publication

Not applicable

### Competing interests

The authors declare that they have no competing interests.

### Funding

No funding was received for this research.

### Authors' contributions

RR-E: Conceived original idea for the article, planned its execution, carried out data gathering and was a major contributor in redaction and correction of final manuscript.

JL-M: Conceived original idea for the article, planned its execution, carried out data gathering and was a major contributor in redaction and correction of final manuscript.

SR-B: Planned its execution, carried out data gathering and was a major contributor in redaction and correction of final manuscript.

MA: Carried out the data gathering and helped in the redaction and correction of the final manuscript.

JH: Data analysis, redaction, and correction of final manuscript.

ET: Data analysis, redaction, and correction of final manuscript.

### Acknowledgements

This article has been developed in the context of a doctoral thesis to obtain the PhD title in Surgery and Morphological Sciences. Escola de Docotrado, Universitat Autònoma de Barcelona, Departamento de Cirugía. Unidad docente de la Vall d'Hebron.

### Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## References

1. ACGME. Program. Requirements for Graduate Medical Education in General Surgery.
2. A F, SY G. The cutting-edge training modalities and educational platforms for accredited surgical training: A systematic review. *Journal of research in medical sciences: the official journal of Isfahan University of Medical Sciences* [Internet]. 2017 [cited 2021 Oct 24];22(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/28567070/>.
3. Shichinohe T, Kobayashi E. Cadaver surgical training in Japan: its past, present, and ideal future perspectives. *Surgery Today* [Internet]. 2021 Jul 5 [cited 2021 Nov 28];1:1–5. Available from: <https://link.springer.com/article/10.1007/s00595-021-02330-5>.
4. B Z, TP SS, KE G. R, AJ D. Perceptions, training experiences, and preferences of surgical residents toward laparoscopic simulation training: a resident survey. *Journal of surgical education* [Internet]. 2014 [cited 2021 Oct 24];71(5):727–33. Available from: <https://pubmed.ncbi.nlm.nih.gov/24794063/>.
5. D L, S B, P G, L H. Outcomes of an innovative training course in laparoscopic hernia repair. *Annals of the Royal College of Surgeons of England* [Internet]. 2017 Nov 1 [cited 2021 Oct 24];99(8):614–6. Available from: <https://pubmed.ncbi.nlm.nih.gov/28682133/>.
6. F N, JD H, KF TB, HG K, BP M-S K, et al. Successful learning of surgical liver anatomy in a computer-based teaching module. *International journal of computer assisted radiology and surgery* [Internet]. 2016 Dec 1 [cited 2021 Oct 24];11(12):2295–301. Available from: <https://pubmed.ncbi.nlm.nih.gov/26872808/>.
7. F N FJ, JD AMMG. H, T G, et al. One or two trainees per workplace in a structured multimodality training curriculum for laparoscopic surgery? Study protocol for a randomized controlled trial - DRKS00004675. *Trials* [Internet]. 2014 Apr 23 [cited 2021 Oct 24];15(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/24754961/>.

8. SR D, JA W, PC JAB, PJ C. H, GJ M. A systematic review of surgical skills transfer after simulation-based training: laparoscopic cholecystectomy and endoscopy. *Annals of surgery* [Internet]. 2014 Feb [cited 2021 Oct 24];259(2):236–48. Available from: <https://pubmed.ncbi.nlm.nih.gov/24100339/>.
9. Andrade M, Rueda R, Lopez-McCormick JS. Existing models for minimally invasive surgery simulation training in post-grad education. [Internet]. 2019. Available from: [https://www.crd.york.ac.uk/PROSPEROFILES/205397\\_STRATEGY\\_20200820.pdf](https://www.crd.york.ac.uk/PROSPEROFILES/205397_STRATEGY_20200820.pdf).
10. PJ G, AH M, TM K. Simulation and virtual reality in surgical education: real or unreal? *Archives of surgery (Chicago, Ill: 1960)* [Internet]. 1999 [cited 2021 Oct 24];134(11):1203–8. Available from: <https://pubmed.ncbi.nlm.nih.gov/10555634/>.
11. MS P, M M. G M, SM C. Objective Assessment of Laparoscopic Force and Psychomotor Skills in a Novel Virtual Reality-Based Haptic Simulator. *Journal of surgical education* [Internet]. 2016 Sep 1 [cited 2021 Oct 24];73(5):858–69. Available from: <https://pubmed.ncbi.nlm.nih.gov/27267563/>.
12. F N, KF K, F R, JD H, BF M, HG K, et al. Face validity of the pulsatile organ perfusion trainer for laparoscopic cholecystectomy. *Surgical endoscopy* [Internet]. 2017 Feb 1 [cited 2021 Oct 24];31(2):714–22. Available from: <https://pubmed.ncbi.nlm.nih.gov/27317031/>.
13. GR MA. W, T A. Virtual reality training in laparoscopic surgery: A systematic review & meta-analysis. *International journal of surgery (London, England)* [Internet]. 2016 May 1 [cited 2021 Oct 24];29:85–94. Available from: <https://pubmed.ncbi.nlm.nih.gov/26992652/>.
14. DA R, DT P. L H, PA K. Cadaveric Simulation Training in Cardiothoracic Surgery: A Systematic Review. *Anatomical sciences education* [Internet]. 2020 May 1 [cited 2021 Oct 24];13(3):413–25. Available from: <https://pubmed.ncbi.nlm.nih.gov/31232510/>.
15. Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training. *Surgical endoscopy* [Internet]. 2003 Oct [cited 2021 Nov 18];17(10):1525–9. Available from: <https://pubmed.ncbi.nlm.nih.gov/14502403/>.
16. Chawla S, Devi S, Calvachi P, Gormley WB, Rueda-Esteban R. Evaluation of simulation models in neurosurgical training according to face, content, and construct validity: a systematic review. *Acta Neurochirurgica* [Internet]. 2022 Feb 4 [cited 2022 Mar 6];1:1–20. Available from: <https://link.springer.com/article/10.1007/s00701-021-05003-x>.
17. Renterghem K van, Nieuwenhove Y van. Training Models in Laparoscopy: a Systematic Review Comparing their Effectiveness in Learning Surgical Skills. 2014 [cited 2022 Mar 6]; Available from: <https://www.researchgate.net/publication/237068276>.
18. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* [Internet]. 2021 Mar 29 [cited 2022 Mar 6];372. Available from: <https://www.bmj.com/content/372/bmj.n71>.
19. Rayyan. – Intelligent Systematic Review [Internet]. [cited 2021 Nov 11]. Available from: <https://www.rayyan.ai/>.
20. Fink A. Construct Validity - an overview | ScienceDirect Topics [Internet]. *International Encyclopedia of Education*. 2010 [cited 2021 Nov 6]. Available from: <https://www.sciencedirect.com/topics/social-sciences/construct-validity>.
21. Van Empel PJ, van Rijssen LB, Commandeur JP, Verdam MGE, Huirne JA, Scheele F, et al. Validation of a new box trainer-related tracking device: The TrEndo. *Surgical Endoscopy* [Internet]. 2012 Feb 21 [cited 2021 Nov 18];26(8):2346–52. Available from: <https://link.springer.com/article/10.1007/s00464-012-2187-6>.
22. Face validity -. Oxford Reference [Internet]. [cited 2021 Nov 6]. Available from: <https://www.oxfordreference.com/view/10.1093/oi/authority.20110803095807498>.
23. Risk of Bias 2 (RoB 2) tool | Cochrane Methods [Internet]. [cited 2021 Nov 28]. Available from: <https://methods.cochrane.org/risk-bias-2>.
24. Smith RP, Learman LA. A Plea for MERSQI: The Medical Education Research Study Quality Instrument. *Obstetrics and gynecology* [Internet]. 2017 Oct 1 [cited 2021 Nov 21];130(4):686–90. Available from: <https://pubmed.ncbi.nlm.nih.gov/28885409/>.
25. Bahsoun AN, Malik MM, Ahmed K, El-Hage O, Jaye P, Dasgupta P. Tablet based simulation provides a new solution to accessing laparoscopic skills training. *Journal of surgical education* [Internet]. 2013 Jan [cited 2021 Nov 6];70(1):161–3. Available from: <https://pubmed.ncbi.nlm.nih.gov/23337687/>.
26. Achurra P, Lagos A, Avila R, Tejos R, Bucke E, Alvarado J, et al. Allowing New Opportunities in Advanced Laparoscopy Training Using a Full High-Definition Training Box. *Surgical innovation* [Internet]. 2017 Feb 1 [cited 2021 Nov 6];24(1):66–71. Available from: <https://pubmed.ncbi.nlm.nih.gov/27729567/>.
27. de Moura Júnior LG, de Vasconcelos PRL, Fechine FV, de Moura MS, de Moura RLS, Rocha HAL, et al. Endosuture trainer box simulator as a tool for training and teaching in bariatric laparoscopic surgery. *BMC surgery* [Internet]. 2018 Oct 9 [cited 2021 Nov 6];18(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/30301475/>.
28. Sánchez-Peralta LF, Sánchez-Margallo FM, Moyano-Cuevas JL, Pagador JB, Enciso-Sanz S, Sánchez-González P, et al. Construct and face validity of SINERGIA laparoscopic virtual reality simulator. *International journal of computer assisted radiology and surgery* [Internet]. 2010 [cited 2021 Nov 6];5(4):307–15. Available from: <https://pubmed.ncbi.nlm.nih.gov/20422300/>.
29. Lahanas V, Loukas C, Smailis N, Georgiou E. A novel augmented reality simulator for skills assessment in minimal invasive surgery. *Surgical endoscopy* [Internet]. 2015 Aug 25 [cited 2021 Nov 6];29(8):2224–34. Available from: <https://pubmed.ncbi.nlm.nih.gov/25303925/>.
30. Sankaranarayanan G, Li B, Manser K, Jones SB, Jones DB, Schwaitzberg S, et al. Face and construct validation of a next generation virtual reality (Gen2-VR) surgical simulator. *Surgical endoscopy* [Internet]. 2016 Mar 1 [cited 2021 Nov 6];30(3):979–85. Available from: <https://pubmed.ncbi.nlm.nih.gov/26092010/>.
31. Van Empel PJ, van Rijssen LB, Commandeur JP, Verdam MGE, Huirne JA, Scheele F, et al. Validation of a new box trainer-related tracking device: the TrEndo. *Surgical endoscopy* [Internet]. 2012 [cited 2021 Nov 6];26(8):2346–52. Available from: <https://pubmed.ncbi.nlm.nih.gov/22350239/>.
32. Escamirosa FP, Flores RMO, García IO, Vidal CRZ, Martínez AM. Face, content, and construct validity of the EndoViS training system for objective assessment of psychomotor skills of laparoscopic surgeons. *Surgical endoscopy* [Internet]. 2015 Nov 1 [cited 2021 Nov 6];29(11):3392–403. Available from: <https://pubmed.ncbi.nlm.nih.gov/25515985/>.
33. Kawaguchi K, Egi H, Hattori M, Sawada H, Suzuki T, Ohdan H. Validation of a novel basic virtual reality simulator, the LAP-X, for training basic laparoscopic skills. *Minimally invasive therapy & allied technologies: MITAT : official journal of the Society for Minimally Invasive Therapy* [Internet]. 2014

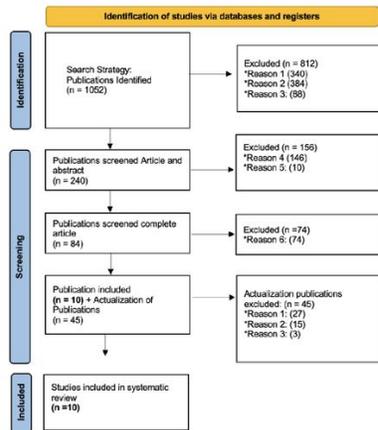
- Oct 1 [cited 2021 Nov 6];23(5):287–93. Available from: <https://pubmed.ncbi.nlm.nih.gov/24773373/>.
34. Arts EEA, Leijte E, Witteman BPL, Jakimowicz JJ, Verhoeven B, Botden SMBl. Face, Content, and Construct Validity of the Take-Home EoSim Augmented Reality Laparoscopy Simulator for Basic Laparoscopic Tasks. *Journal of laparoendoscopic & advanced surgical techniques Part A* [Internet]. 2019 Nov 1 [cited 2021 Nov 6];29(11):1419–26. Available from: <https://pubmed.ncbi.nlm.nih.gov/31613679/>.
  35. Fink A. Construct Validity - an overview | ScienceDirect Topics [Internet]. *International Encyclopedia of Education*. 2010 [cited 2021 Oct 30]. Available from: <https://www.sciencedirect.com/topics/social-sciences/construct-validity>.
  36. Arora S, Sevdalis N, Nestel D, Woloshynowych M, Darzi A, Kneebone R. The impact of stress on surgical performance: A systematic review of the literature. *Surgery*. 2010 Mar;147(3)(1):318–30.e6.
  37. Hau HM, Weitz J, Bork U. Impact of the COVID-19 Pandemic on Student and Resident Teaching and Training in Surgical Oncology. *Journal of clinical medicine* [Internet]. 2020 Nov 1 [cited 2021 Nov 17];9(11):1–15. Available from: <https://pubmed.ncbi.nlm.nih.gov/33114552/>.
  38. Bohl MA, McBryan S, Spear C, Pais D, Preul MC, Wilhelmi B, et al. Evaluation of a Novel Surgical Skills Training Course: Are Cadavers Still the Gold Standard for Surgical Skills Training? *World neurosurgery* [Internet]. 2019 Jul 1 [cited 2022 Mar 6];127:63–71. Available from: <https://pubmed.ncbi.nlm.nih.gov/30930320/>.
  39. Bongers PJ, van Hove PD, Stassen LPS, Dankelman J, Schreuder HWR. A new virtual-reality training module for laparoscopic surgical skills and equipment handling: can multitasking be trained? A randomized controlled trial. *Journal of surgical education* [Internet]. 2015 Mar 1 [cited 2022 Mar 6];72(2):184–91. Available from: <https://pubmed.ncbi.nlm.nih.gov/25439179/>.
  40. Shah S, Aydin A, Fisher R, Ahmed K, Froghi S. P. Dasgupta. Current status of simulation-based training tools in general surgery: A systematic review [Internet]. *International Journal of Surgery Open*. 2021 [cited 2021 Nov 25]. Available from: <https://reader.elsevier.com/reader/sd/pii/S2405857221001182?token=8984AEE918EC73DBF413A2A922DAD8BFFBDF9B65CDC50C927637501B578C69C6D4C8FBF3D69994133F6C813D8AA44CD9&originRegion=us-east-1&originCreation=20211126160705>.

## Tables

**TABLE 1: Study characteristics**

Title	Author/Year	Participants	(n =X) / Model	Validation
Construct and face validity of SINERGIA laparoscopic virtual reality simulator.	<b>Sánchez-Peralta et al.</b> [2010]	Novices Experts	n=20 / VR	Face - Construct
Validation of a new box trainer-related tracking device: the TrEndo.	<b>Van Empel et al.</b> [2012]	Novices Experts	n=64 / BS	Face – Construct – Content
Tablet Based Simulation Provides a New Solution to Accessing Laparoscopic Skills Training	<b>Bahsoun et al.</b> [2013]	Residents Surgeons	n=10 / BS	Face – Content
A novel augmented reality simulator for skills assessment in minimal invasive surgery	<b>Lahanas et al.</b> [2014]	Novices Experts	n= 20 / VR	Face – Construct
Face, content, and construct validity of the EndoViS training system for objective assessment of psychomotor skills of laparoscopic surgeons.	<b>Perez Escamirosa et al.</b> [2014]	Novices Experts	n= 35 / VR	Face – Construct – Content
Validation of a novel basic virtual reality simulator, the LAP-X, for training basic laparoscopic skills.	<b>Kawaguchi et al.</b> [2014]	Novices Experts	n= 20 / VR	Face – Construct – Content
Face and construct validation of a next generation virtual reality (Gen2-VR) surgical simulator.	<b>Sankaranarayanan et al.</b> [2015]	Fellows Attending PGY5's	n= 16 / VR	Face - Construct
Allowing New Opportunities in Advanced Laparoscopy Training Using a Full High-Definition Training Box	<b>Achurra et al.</b> [2016]	Residents Experts	n= 15 / BS	Face – Content
Endosuture trainer box simulator as a tool for training and teaching in bariatric laparoscopic surgery.	<b>De Moura Júnior et al.</b> [2018]	Residents Experts	n= 37 / BS	Face - Construct
Face, Content, and Construct Validity of the Take-Home EoSim Augmented Reality Laparoscopy Simulator for Basic Laparoscopic Tasks.	<b>Arts et al.</b> [2019]	Novices Intermediates Experts	n= 102 / VR	Face – Construct – Content

## Figures



\*Reason 1: No validation, \*Reason 2: Other specialty, \*Reason 3: Wrong study or Language, \*Reason 4: No complete/described validation, \*Reason 5: Wrong Outcome, \*Reason 6: Don't have Two kinds of validation, at least one must be Content Validity or Construct Validity.

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: <http://www.prisma-statement.org/>

Figure 1

**PRISMA Flow Diagram:** From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi: 10.1136/bmj.n71. For more information, visit: <http://www.prisma-statement.org/>

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [3.ANNEX1SV.docx](#)