Robot-Assisted Medical Training for Safety-Critical Environments

Huajie Jay Cao
Media & Info, Michigan State Univ
Michigan, USA
caohuaji@msu.edu

Michael Joseph Sack
Lili Mkrtchyan
mjs596,lm688@cornell.edu
Info Sci, Cornell Univ
Ithaca, New York, USA

Kevin Ching
Emergency Med, Weill Cornell Med
New York, USA
kec9012@med.cornell.edu

Tariq Iqbal
School of Eng and Appl Sci, Univ of Virginia
Virginia, USA
tiqbal@virginia.edu

Hee Rin Lee
Media & Info, Michigan State Univ
Michigan, USA
heerin@msu.edu

Angelique Taylor
Info Sci, Cornell Univ
New York, USA
amt298@cornell.edu

ABSTRACT

While resuscitation training is critical, healthcare workers (HCWs) with high workload have limited chance to get trained and retrained due to time and resource constraints. To address this gap, we engaged in a co-design process of robots that facilitate and prepare HCWs for resuscitation procedures (i.e., codes). First, we investigated what resuscitation training consists of, including challenges faced by trainees and trainers. Second, we collaboratively explored how a crash cart robot, that guides users to medical supplies and equipment, could assist trainers and trainees synchronously–during team-based clinical simulations and asynchronously–during one-on-one training. We found that robots could 1) serve as a learning assistant by providing real-time feedback and supporting personalized training needs; and 2) an evaluating assistant by monitoring multiple trainees and tracking critical timing of interventions in the training. Through this new training paradigm, we hope to demonstrate opportunities for crash cart robots to aid HCWs for their sustainable training and reskilling. We discuss the role of robots in training beyond cognitive knowledge, situating them within two underexplored contexts: practical skill training and team-based training.

CCS CONCEPTS

• **Human-centered computing → Participatory design**; • **Computer systems organization → Robotic autonomy**.

KEYWORDS

robots, emergency medicine, clinical training, co-design

ACM Reference Format:
Huajie Jay Cao, Michael Joseph Sack, Lili Mkrtchyan, Kevin Ching, Tariq Iqbal, Hee Rin Lee, and Angelique Taylor. 2026. Robot-Assisted Medical Training for Safety-Critical Environments. In *Proceedings of the 21st ACM/IEEE International Conference on Human-Robot Interaction (HRI '26), March 16–19, 2026, Edinburgh, Scotland, UK*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3757279.3785607

HRI '26, March 16–19, 2026, Edinburgh, Scotland, UK

ACM ISBN 979-8-4007-2128-1/2026/03
https://doi.org/10.1145/3757279.3785607

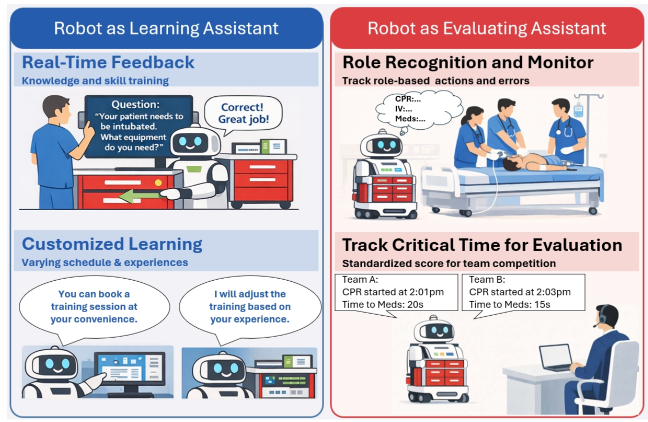

Figure 1: Crash cart training is essential for healthcare worker readiness for high-stakes patient care. We engaged in the design space of robots embodied in crash carts and their potential to deliver effective training.

1 INTRODUCTION

Medical crash carts are mobile units that store medical materials, equipment, and medications [8]. Healthcare workers (HCWs) use crash carts to treat patients in safety-critical environments, including Emergency Departments (EDs) and Intensive Care Units (ICUs). However, HCWs often face challenges locating relevant supplies in the cart due to limited exposure to crash carts, conducting resuscitation procedures, and treating patients in the hospital, particularly for new HCWs. Our research explores how medical crash cart robots can facilitate crash cart training to HCWs, offering an automated education tool for synchronous and asynchronous training to improve learning outcomes.

Despite its importance, many HCWs do not receive formal crash cart training in school, and the occurrence and frequency of crash cart training varies across healthcare institutions. Thus, many HCWs lack the opportunity for scenario-based learning or simulation training to gain practical skills and engage in critical thinking to use the crash cart effectively. To mitigate this issue, HCWs are trained to utilize the crash cart during orientation in the hospital. Crash cart training involves teaching medical personnel how to effectively use a crash cart, quickly identify and locate relevant supplies, and use those supplies to respond to critical situations such as cardiac arrest or other life-threatening patient conditions. However, HCWs can spend long periods of time without using the crash cart; thus, when a critical patient arrives in the hospital and requires use of a crash cart for patient care, HCWs must quickly recall their training, which is problematic due to the high-stress, chaotic nature of rare, high-stakes patient cases.

Robotic systems offer unique opportunities to address challenges in training [27]. We envisioned two different ways robots could be used in crash cart training. From a teaching perspective, robots can help walk trainees through standardized evidence-based practices, leaving the trainers to focus on processes unique to a given healthcare institution (i.e., synchronized training). Furthermore, from a trainee perspective, robots can provide one-on-one personalized training curriculum to tailor educational materials to the HCWs' level of experience and expertise (i.e., asynchronous training).

Prior studies have shown that robots are increasingly being used in nurse training programs and healthcare simulation labs, primarily to provide realistic and immersive learning experiences [4, 25]. The most popular example is the semi-autonomous robots used for non-invasive surgery, particularly using the da Vinci robot [28]. Robotic patient manikins are a long-standing technology that simulates patients in terms of bleeding, breathing, and blinking, among others. Prior work has built robotic heads to replace the static heads on standard patient simulators to improve HCWs' recognition of patients' emotions (e.g., pain) [18–20]. Other efforts primarily focused on designing robots for nurse training, including identifying problems that need addressing to enable nurse training for onboarding and reassessment [22], telepresence robots for clinical simulation training [17], and training to reduce healthcare-associated infections [21].

Despite recent advances in robot developments for nurse training, several research gaps remain. First, there is limited knowledge about how robots can support trainers during hospital onboarding, which could enable HRI researchers to build robots that facilitate institution-specific curriculum. Second, limited work provides insights into effective design practices for robots that train nurses or pharmacists one-on-one to support their unique needs (i.e., asynchronous). Third, there is a lack of work that provides actionable design guidelines for robots to support team-based training of HCWs (i.e., synchronous) in safety-critical environments such as codes.

In this paper, we address these gaps in a co-design study with HCWs to explore opportunities for synchronous and asynchronous crash cart training in the hospital. We used our recently built medical crash cart robot as a design probe to help participants reflect on the challenges faced during crash cart training from the trainer and trainee perspectives. Furthermore, we explore opportunities for robots to support synchronous and asynchronous training for nurses and pharmacists to uncover their unique skills training needs. This study explores these goals in the following **research questions**:

RQ1: How can crash cart robots enable HCWs to identify, locate, and use supplies in the crash cart during scenario-based training?

RQ2: How can crash cart robots help HCWs to practice critical thinking skills in synchronous and asynchronous training schemes?

RQ3: How can robots be designed to support the unique needs of trainers and trainees, in dyadic and team interactions in safety-critical environments?

The contributions of this paper are threefold. First, we present an understanding of simulation training from key stakeholders, including Nurses and Pharmacists. Second, we provide insights into how crash cart robots can engage in synchronous and asynchronous training to prepare nurses and pharmacists to effectively locate and use supplies in the cart. Third, we shed light on new opportunities for robots to train and assess learners' performance in dyadic and team interactions.

2 BACKGROUND

2.1 Healthcare Robotics

Healthcare robotics has emerged as a transformative field addressing critical challenges in medical practice, including workforce shortages, training standardization, and quality-of-care delivery [24, 31, 33]. Robots in healthcare settings span a broad spectrum of applications, including surgical assistance, rehabilitation, logistics, and clinical training support [34]. In surgical contexts, robots have demonstrated their capacity to enhance precision and reduce invasiveness in complex procedures. The da Vinci surgical system, for instance, has become widely adopted for minimally invasive surgeries, enabling surgeons to perform delicate operations with enhanced dexterity and visualization [23]. Similarly, rehabilitation robots provide consistent, measurable therapy for patients recovering from strokes or traumatic injuries, offering repetitive task training that would be physically demanding for human therapists to sustain [2]. Beyond direct patient care, robots increasingly support healthcare operations and logistics. Service robots manage logistics tasks such as medication delivery and supply transport, reducing the non-clinical burden on healthcare workers [30]. Social robots have shown promise in eldercare settings, providing companionship and cognitive stimulation while monitoring patient wellbeing [26]. These diverse applications underscore robotics' potential to augment rather than replace human healthcare workers, addressing efficiency gaps while maintaining the human-centered nature of medical care [33].

2.2 Robots for Clinical Training

Robots offer unique capabilities for clinical training, including adaptive responses, real-time performance feedback, and objective skill metrics [1]. This is especially critical in emergency medicine, where healthcare workers must maintain proficiency in high-stakes procedures they may rarely perform [8]. The integration of robots into medical education represents a natural evolution from static simulation toward more interactive, personalized training experiences [6].

Robots are increasingly used in nurse training programs, both to familiarize nurses with robots in clinical services [16, 24] and to

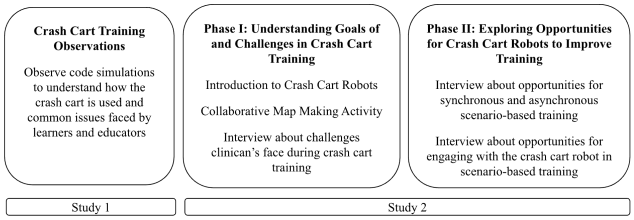

Figure 2: Study Overview

deliver high-fidelity simulations for practicing clinical skills [5, 14]. Recent work has focused on robots that directly train nursing skills through personalized feedback. For example, Qian et al. [21] introduced ASTRID, a robotic system that provided real-time guidance on physical interventions to help nurses master sterile dressing changes. Building on this understanding, we extend the focus to life-critical contexts and examine how robots may enhance nursing skills in code situations through direct observations of code simulations and co-design sessions with nurses.

Other work, such as Huang et al. [7] developed a robotic patient that could simulate limb movements for patient-transfer training and reported significant skill gains among trainees. However, recent studies have focused on using robots to directly train nursing skills. In the aftermath of the COVID-19 pandemic and the resulting nursing shortages, the cultivation of nursing skills has become increasingly critical. Robot tutors who could provide personalized feedback hold promise for reducing the time required to acquire essential competencies and for enabling trainers to train larger cohorts of nurses [22].

2.3 Research Methods

To investigate how crash carts are currently used for training and how such training is conducted and can be improved by robots, we carried out two studies (see Figure 2): **Study 1 (Observations):** Many hospitals conduct code simulations involving crash carts, we observed these simulations to understand how the carts are used and how staff are trained. **Study 2 (Co-design):** We engaged healthcare workers in co-design sessions to explore how a robotic crash cart could support training and to gather design requirements for future systems.

2.4 Study 1: Observations of Clinical Training Sessions

2.4.1 Overview. As part of our design process, two of the coauthors observed two monthly mock code simulations at a hospital in the eastern global north. The goal of these observations was to understand how nurses are trained to manage code situations, with particular attention to how they interact with and train on the crash cart, as well as to identify potential opportunity areas for improving crash cart training. These simulations were designed to mirror real resuscitations. Scenarios began without giving trainees prior notice. The simulations involved ICU nurses who were on shift at the time of observation. They covered multiple code roles, including the primary nurse of the "patient" (a mannequin), the recorder, the crash cart manager, and several nurses for chest compressions. Besides, a physician worked as code leader, and an anesthesiologist joined halfway.

2.4.2 Procedure. The simulation began when the primary nurse found the "patient" unresponsive, started compressions, and called a code. Then the rest of the team entered the room, and the code leader (the physician) assigned roles. One nurse (recorder) recorded the time and tracked intervals for medications and pulse checks. Two nurses stood next to the crash cart, one preparing and delivering medications and materials, and the other managing the devices such as the defibrillator. One IV nurse coordinated between the bed and crash cart to administer medications. Two to three nurses rotated through chest compression. The primary nurse reported to the code leader about the "patient" situation, and the code leader was responsible for medication orders and the entire code. An anesthesiologist entered partway to replicate real-world crowding and noise by raising their voice and moving within the team's space. After several rounds of coordinated action, the team restored the "patient" to spontaneous circulation, and the simulation ended.

Each simulation was conducted twice in succession. After the first round, all participants engaged in a debrief in which they reflected on what went well and identified areas for improvement. The facilitators of this debrief were not dedicated trainers or educators, but rather physicians who had taken on the responsibility of organizing and leading the simulation. Although everyone contributed feedback, the facilitators were primarily responsible for guiding the process. The simulation was then run again to address the issue that emerged in the first simulation, followed by a second round of debriefing. Following these observations, we spoke with healthcare workers on site about their experiences with the simulation and the use of crash cart in the training.

2.5 Results of Hospital Site Observation

The observation of two code simulations at the site informed the use of the crash cart in current training and several challenges.

The role of the crash cart in simulation: In contrast to our initial vision of a robotic crash cart with interactive functions—such as providing trainees with patient background information—the observed simulations revealed that such features were not considered important. According to two simulation leaders, during a real code and simulation, the team's primary goal is to restore stability rather than diagnose or address the underlying disease. The use of the crash cart itself signals that the patient is unstable and requires immediate resuscitation, not investigative reasoning. As a result, during the simulation, the cart remained stationary against the wall, occasionally being moved only to make room for staff to pass.

Discrepancies of Using Crash Cart in Simulation and in Real Code: The simulation sought to replicate the actual code situation, but not the use of the crash cart. The anesthesiologist deliberately entered partway through the simulation, raised their voice, and disrupted the environment to mimic the chaos in real codes and increase the pressure on trainees. They noted that during real code situations, especially when less experienced, healthcare workers may panic and act abnormally under the pressure of life-safety. For example, they saw novices remove multiple medications without properly preparing them and delaying the delivery. Such behaviors were observed less frequently in the simulation. One important reason we learned was that the simulations employed a 'dummy' crash cart stocked with expired and reduced supplies. While this setup allowed staff to rehearse logistical coordination, it failed to reflect the actual organization and contents of a real crash cart. As a result, it undermined the ability to build cart-specific literacy and constrained opportunities to practice accurate and timely retrieval of medications.

This study provides useful insights about how HCWs utilize the crash cart during crash cart training. In the next study, we engage with HCWs to explore how robots embodied in crash carts can support learners and trainees in medical training environments.

2.6 Study 2: Co-Design Study

2.6.1 Participants. We recruited 13 participants as part of an IRB-approved study, during an event at a medical school in the global north, through recruitment email and word-of-mouth. 3 out of 12 of the participants did not complete the demographic form to its completion, so we report the demographic data to the best of our ability. Nine participants identified as female and three as male, with three participants' ages ranging from 25-34 and six participants' ages ranging from 35-44. Participants were healthcare workers with clinical experience in roles such as 3 Registered Nurses RN (1 Nursing Professional Development Specialist), 1 Clinical Nurse Specialist, 1 Critical Care Clinical Nurse Specialist, 1 Pharmacist, 1 Certified Registered Nurse Anesthetist, 1 Emergency Department Nurse Educator, 1 Nursing Professional Development Specialist, 1 Emergency Medicine Nurse Practitioner. [1]

2.6.2 Phase I: Understanding Goals of and Challenges in Crash Cart Training. We first introduced our robot prototype to participants by walking them through the key capabilities of the robot—voice query, LED guidance, tablet-based step-by-step instructions, and inventory management system. We emphasized that the prototype served as a design probe and invited participants to codesign the robot for the later session (see Figure 3).

We then conducted a Collaborative Map-Making Activity to understand the matter of concern in crash cart training such as goals and challenges [12]. We asked participants to write 15 words associated with the keywords: Nurse, training/education, crash cart. Then we asked participants to identify 2-3 most important words and describe the rationale. In the end, we asked participants to group and relate those words into 2-3 themes, and interpret their map.

2.6.3 Phase II: Exploring Opportunities for Crash Cart Robots to Improve Training. In the second phase, we conducted a semi-structured interview to explore how a robotic crash cart could support training. We began with a general question about how robots could be used to enhance crash cart training. We then asked participants to consider the robot's role across different types of training, including simulation-based exercises, synchronous training with trainers and other trainees, and asynchronous practice in the hospital. In addition, drawing on the most important words and themes identified in Phase I, we asked participants to reflect on how a robot could address those specific issues in training. Participants were encouraged to propose ideas both grounded in the demonstrated prototype and extending beyond it.

2.6.4 Data Collection & Analysis. We recorded video of interviews via Zoom and transcribed the data using Whisper, an automatic speech recognition model, which we deployed locally to ensure data security and privacy. The transcripts were then analyzed using a thematic analysis grounded in constructivist approach[3], which enabled us to identify themes beyond predefined hypotheses. For instance, although we initially distinguished three potential types of robot-supported training—code simulation, asynchronous, and synchronous—we found participants highly valued training efficiency and often treated synchronous training as team-based code simulation, which they regarded as the most effective format. Accordingly, we adopt this participant-centered definition in the paper: asynchronous training refers to one-on-one learner–robot sessions for skills and knowledge, where the robot serves as a learning assistant (Section 3.3); synchronous training refers to team-based code simulation, where the robot serves as an evaluating assistant (Section 3.4). Along with the interview analysis, we also constantly examined and reflected on participants' maps of crash cart training by identifying words and themes that frequently emerged. We analyzed how these themes and words were related to each other [12].

3 RESULTS

3.1 Understanding Crash Cart Training

We reflected on our analysis of the collaborative maps to learn about participants' thought processes with respect to crash cart training. A crash cart, or code cart, is a mobile unit that contains the materials, medications, and devices necessary to perform a code. Code refers to medical emergencies where a patient needs immediate and professional intervention. We first analysed the words participants came up with for crash cart training. The most commonly used words were "code roles" and "supplies". Eight participants stated the importance of understanding different roles in a code (e.g., recorder, medication nurses)and diverse supplies in the crash cart(i.e., materials, medications, and devices). Seven participants also chose "experience" (or "exposure to code") as the keyword. Six participants connected crash cart training with clinical protocols such as Advanced Cardiac Life Support (ACLS), Basic Life Support (BLS), and Cardiopulmonary Resuscitation (CPR). These responses suggested that participants considered crash cart training inseparable from code practice. Mastery of the crash cart and its supplies requires not only familiarity with the equipment but also a solid foundation in code knowledge and resuscitation protocols.

Then, we analyzed the themes that were grouped based on participants' own keywords. These themes reflected a broader interpretation of crash cart training. Notably, two themes appeared consistently across five participants. The first emphasized knowledge, processes, and clinical considerations, while the second highlighted application, skills, and operational considerations in training. As participants elaborated on these themes, we recognized that they

[1] Role counts were chosen to reflect the real-world stakeholder distribution in crash cart use. Nurses are the primary users during resuscitation, so most participants were nurses or nurse educators. In some hospitals, pharmacists serve as secondary users responsible for managing and verifying high-risk medications in the cart; accordingly, we included one pharmacist. Although physicians typically lead diagnosis and clinical decision-making during codes, they are rarely involved in crash cart management and are therefore outside the scope of this study.

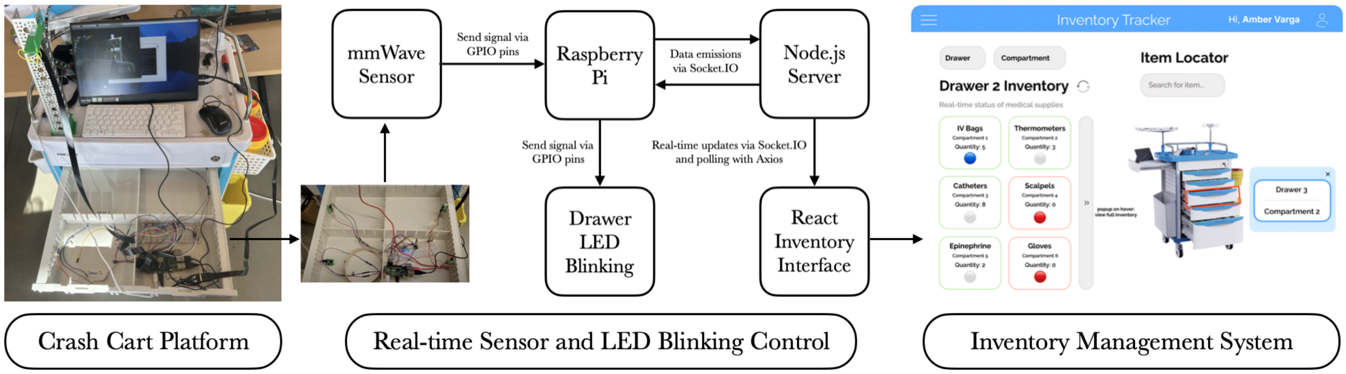

Figure 3: Crash Cart System Platform Overview

represent two fundamental components of training. In addition, a third section corresponded to the simulation practices that we observed in Study 1.

We therefore concluded these three main components of training: knowledge, skill, and simulation. **Knowledge** training refers to understanding of protocols, algorithms, and pharmacology— such as causes and symptoms of cardiac and respiratory arrest, the functions of different medications and devices in the crash cart, and content of different algorithms (e.g., ACLS; CPR). **Skill** training involves practice-based ability and hands-on activities that translate clinical knowledge into action, such as medication administration, chest compressions, and rhythm-guided defibrillation. Importantly, all participants recognized using a crash cart as a vital skill. **Simulation** training embeds knowledge and skills into scenario-based practice designed to replicate real-life code situations.

3.2 Challenges in Crash Cart Training

Based on the above understanding, we further analyzed our interviews about challenges in current crash cart training and the possibilities of robot design.

3.2.1 Challenges of knowledge and skill development. **Codes as life-critical events demanding knowledge and skill.** Codes represent life-threatening emergencies where immediate and effective intervention can determine survival. Ideally, every nurse is prepared to act on codes, as P7 explained, *"If it is their patient [who codes], they're the ones in charge of controlling the room until the doctor arrives."* However, participants underscored that it was exceptionally demanding to learn. P1 noted, *"There are four different algorithms [in resuscitation], and medicines are dangerous if not used in the correct algorithm, so ensuring we're using the right drugs for the right scenarios and then dosing and administration is very important."* Patient-specific factors add another layer of uncertainty in code intervention: the same condition may have multiple possible causes, and while knowing the textbook list of causes is not difficult, identifying what is driving the condition in a particular patient is the real challenge. As P3 added, even a common medical intervention could become tricky if the patient had an allergy.

Besides, in practice, many HCWs are not familiar with code because it occurs less frequently than other clinical situations. This scarcity of exposure makes even the training of basic knowledge a challenge, such as the steps of resuscitation. P4 noted that *"If you don't go to codes often, you'll just forget."* For this reason, training sessions often dedicate several hours to reviewing foundational knowledge.

Skills, however, can be even more difficult to acquire. As P5 explained, *"Lecture, orientation, roles in an emergency—that's all content that's covered, but until they can, like, apply it in real life, it's really hard for them to practice it."* For example, even if nurses know the functions of all supplies in the crash cart, less experienced nurses may struggle to locate the needed items quickly within its drawers. Indeed, 8 out of thirteen participants described this lack of crash cart literacy as a persistent obstacle in both training and real-life codes. Furthermore, as P13 emphasized, enacting skills in life-critical situations is even more challenging: *"You're stressed out if you're using a crash cart, it might take you a little bit longer than usual to just like read and then open the drawer for what you're looking for."*

Time Constraints in Training. The unfamiliarity of code translates into greater time demands during training, which stands in direct tension with HCWs' ongoing patient care responsibilities. Coordinating time for training, therefore, becomes a main challenge. P3 talked about the difficulty of coordinating training time between the trainer and trainees: *"Like a lot of times, if I come [as the trainer], the number of nurses that I capture is dependent on [their] patient acuity, availability of nurses, right. So sometimes I could be there for like 3 hours and not get a lot of nurses to come."*

P4 further explained, their hospital had to cut what was considered the most effective training session because of time constraints: *"We had this crazy session where we went through every single dysrhythmia known to man and then how the nurses would go through the code cart, but unfortunately, due to volume and time restrictions, we had to get rid of it."*

3.2.2 Challenges of engaging trainees in code simulation. Simulation training aims to replicate real-world codes. However, the gaps between training and real-world practice also make it challenging to stimulate trainees' engagement. Four participants noted that current simulation training is perceived as useful primarily for familiarizing HCWs with the overall flow of code resuscitation rather than reproducing its realities, as explained by P8, *"It's not like a real depiction of what's going on. And we don't take it as seriously because sometimes the supplies are being used for training constantly. You're not actually opening the supplies; you're not actually unlocking the cart. Or sometimes the supplies could be missing and then you just say like, 'oh, just pretend that this is here' or something, you know."*

Besides, simulation training frequently lacks the kind of feedback inherent to real-world practice. Mistakes rarely carry meaningful consequences—for example, taking the wrong item or reacting slowly may be overlooked, since trainees simply assume that the correct step was taken. Although several trainers are present to provide feedback and evaluation, it's hard to capture every detail, given the abundance of supplies and simultaneous role-specific actions. As P9 explained, *"there's been so many times that we've been either in a real code or a training situation where I'm focusing on one thing, and I look over and the nurses are already in the wrong drawer, drawing up the wrong thing or getting the wrong supply."*

Two participants offered contrasting evaluations of simulated codes depending on whether they involved an interactive mannequin or a stationary one. P7 described how their hospital used mannequins that could breathe, talk, and be controlled in detail, which they considered highly useful. In contrast, P12 noted that in their setting, mock codes mostly involved only a stationary mannequin, and at times, this exercise was skipped.

3.3 Robot as Learning Assistant

Considering the demanding knowledge and skills for crash cart training, participants envisioned robots as an assistant to facilitate trainees' learning. As discussed above, current training sessions offer limited opportunities for hands-on practice or individualized feedback. A robot embodied in the crash cart could help address this gap by enabling interactive, scenario-based activities and delivering timely, personalized feedback to trainees.

3.3.1 Robot provides real-time feedback for knowledge and skill training. The demanding knowledge and skills required for codes make real-time feedback especially critical. Some existing training formats already provide interaction with trainers, such as Q&A sessions, but sustaining this is difficult when one trainer may be responsible for ten or more trainees. Participants envisioned a role for robots in filling this gap. As P4 explained, *"We do this without a robotic cart, and we have an educator there, and you (trainer) ask these questions kind of like jeopardy style. So, it'd be great if the code cart [robot] can just do it itself, you know."* In addition, robots could also support interactive exploration of the crash cart to strengthen trainees' operational skills. As P10 described, *"You might go through each drawer and be asked what you think is inside, then [robot] uses prompts to get them (trainees) more comfortable with the cart."*

Five participants further suggested that robots could integrate knowledge and skill training by coming up with a question and then prompting trainees to perform the relevant action on the cart. P9 explained: *"It [Robot] could say, 'Your patient needs to be intubated. What equipment do you need to get and where is it?'"* Meanwhile, robots could provide multimodal feedback to guide trainees to make the right actions, for example, blinking lights on the right drawer and voice prompts. Furthermore, P6 stated that the feedback should be more than right or wrong, they suggested that robots could build on the way trainees often 'think aloud' during interaction with trainers, and give feedback on the reasoning trainees had: *"Kind of not just say you got it right or wrong, but give some details on your reasoning behind why you did what you did, that would probably be way more valuable."*

Building on this, three participants highlighted the value of *error-focused feedback*, noting that training is meant to be a space for trial-and-error. Rather than directly providing the correct answer, offering trainees sufficient cues to guide them toward discovering the right action themselves was seen as the most effective approach. For instance, P10 suggested signaling mistakes by locking the drawer to prompt reflection of trainee's practice: *"Maybe I touched the wrong drawer and it's going to then alert me that no, it's in this one. Like it locks. I think that's the most helpful thing because sometimes that's half the battle. The things are in there. They're like buried, flipped over. So, if you didn't immediately find it, you then might open another drawer and it locks like [it's telling you], 'no, wait, just pause. It is in this drawer. Let's look for it.'"*

3.3.2 Robot can satisfy different learning needs of trainees. Participants also suggested programming the robot to meet the unique needs of trainees. For new nurses and those working in units where codes occur infrequently, training is often the primary context for exposure to code knowledge and skills. These HCWs may be unfamiliar with resuscitation algorithms, and a robot could support them by walking them through the basic algorithmic steps. Inspired by our design probe, three participants suggested flowchart-style visualizations could be useful making these algorithms explicit and easier to follow. Besides, P10 suggested targeted training for high-use items to save their time: *" You could tailor it to finding a few items that are used a lot in whatever hospital system or unit."*

In contrast, more seasoned nurses could be reinforced in specific areas aligned with their experience of code roles. For example, if a trainee is less familiar with medications in the crash cart, the system could provide targeted training, as P8 noted, *"If they want to focus on a specific aspect of the crash cart, programming it to reinforce medication selection, dosing, or timing specifically. It seems like the possibilities are pretty customizable, and would depend on what needs of the person are. "* Moreover, because nurses from different units encounter varying causes of codes, participants suggested that robots should provide scenario-based customization. P11 envisioned a library of clinical scenarios accessible on demand: *"The most optimal thing in that environment would be to be able to program multiple different clinical scenarios within the robot that nurses could access, like on demand of a nurse's specialty."*

Moreover, participants highlighted how robots could make training more flexible. As discussed, there are challenges in coordinating in-person training with nurses' patient care responsibilities. With the assistance of robots for training, P3 suggested that, *"If you put it on the nurse to say, hey, we have the robotic cart in room XYZ, you have until this date to review the code cart [...] they can go whenever they want, at a time that works best for them."* Such a new training format could not only respect HCWs' schedules but also provide records of training completion for recertification.

3.4 Robot as Evaluating Assistant in Simulation

For simulation training, two participants worried that assistance from a robot might undermine the very purpose of mock codes—preparing HCWs for real codes where no robot is present. As P7 explained, *"If you were to do it (using a robot in simulation), you would just use it the same way you would now for a regular code. That's the whole purpose of a mock code."* For simulation training, therefore, participants emphasized that the robot should avoid direct intervention or assistance to trainees. The most discussed role of the robot was an evaluation assistant to reduce the workload of trainers. Five participants envisioned that robots could quietly monitor and record the performance of trainees.

3.4.1 Robots monitor and record different roles in code simulations. In simulations, HCWs are assigned different code roles, each responsible for a distinct part of resuscitation. To evaluate their performance, trainers must closely monitor a wide range of role-specific behaviors. For example, for the compressor, this includes rate and depth of compressions; for the recorder, adherence to algorithm reminders such as administering epinephrine and checking the pulse, as well as documenting medication use and time stamps; for medication nurses, preparation of appropriate dosage and on-time delivery. As discussed, the abundance of supplies and simultaneous actions often makes it impossible to capture every detail.

Participants anticipated that robots could recognize different code roles and monitor role-specific actions and errors. As P7 described, *"The robot recognizes their faces and knows what they're supposed to be doing. They even have the cameras now that follow you, like once it identifies your face. And then, you know, you start the code, they push a drug, the robot knows. It assesses for errors and things that can be improved for the future."* Based on the robot's records, even a single trainer is able to oversee the simulation and provide evaluations for different trainees. Besides, P6 thought that the robot could reconstruct how an error was made and provide opportunities for trainees to redo the task, turning mistakes into learning opportunities.

Beyond individual monitoring, participants also envisioned robots supporting team-based evaluation through competitive comparisons. P9 suggested that error rates and task performance could be tracked across groups, fostering camaraderie and realism: *"When we assign the roles like the airway nurse, the IV, the medication nurse. If you almost are competing in teams of like who can get the things most accurate—which team did the best compressions, at the right depth, at the right rate, who opened the correct drawer on the first try—something like that, like almost you're trying to work together to really try, instead of just having someone stand there and be like, now open this drawer, like more of a camaraderie."*

3.4.2 Robots track time for evaluation. In resuscitation, time is directly tied to patient survival. Therefore, time also provides an important basis for evaluating performance during codes, yet participants noted that it was difficult to record with precision. P9 reflected how trainers sensed time in simulation: *"Sometimes there's a group no matter what they do, like they can't get the CPR rates right or they're taking forever to drop the meds. And sometimes it just has to be almost obvious—there's no real way to track, it's just kind of what we saw whenever we could see it."* A crash cart robot, however, could simplify and enhance existing practices of time tracking, since many critical time points are tied directly to the cart itself. As P8 explained, *"The robot could track the time it takes from the time the prompted script scenario is given to the time that it takes for the nurse to actually take action—open that drawer and detect that they pulled the right supply. [Or even earlier,] how long it takes to get the cart to the emergency situation and start opening it and using it."* P10 compared crash cart robots to robotic patient mannequins in their hospital that are equipped with sensors to evaluate the rate, depth, and recoil of CPR and then generate a performance grade. They suggested that by tracking the timing of crash cart use, robots could provide a similar form of scoring, offering a more standardized way to assess competencies.

3.5 Concerns of using a robot in training

Participants showed two main concerns about using robots for resuscitation training.

Improving the validity of robots. Four participants raised concerns about whether healthcare workers would trust robots in training. P3 emphasized that a robotic trainer should be rigorously validated; otherwise, a senior trainer would still be needed to supervise: *"I think it would be more efficient if you could do it without trainers, but if you developed a system that did that, you would have to have it, you know, rigorously validated."* Participants stressed that just as HCWs require certification for different resuscitation algorithms, robots would also need transparent validation of how they are programmed. Besides, participants noted that more seasoned nurses, especially those in the ICU, would have higher expectations for training quality. As P4 explained, *"When you get to ICU, they'll probably want more specific things from the cart that the nurses need to know about versus other specialties."* Meeting such expectations would be critical to the successful deployment of robots in practice.

Risk of Overreliance on Robotic Support. Two participants cautioned that excessive reliance on robotic prompts could undermine preparedness for real codes. For example, P10 warned: *"We can't just hand it to trainees, because then they're always going to want that blinking light. We need to consider how much prompting we provide, because that won't happen in real-life scenarios."*

4 DISCUSSION

4.1 The Role of Robot for Training

This study contributes to the HRI community by extending the understanding of the robot's role in teaching and training. Consistent with prior work, we recognize the value of robot training in addressing workforce shortages across many professions, particularly in healthcare [29, 32]. By providing individualized feedback to trainees, robots could promise effective learning outcomes [10, 11, 13].

Nevertheless, our study advances existing research in two important ways. First, we situated our analysis within the context of nursing—a domain that emphasizes not only cognitive learning but also practical, skill-based training. Second, we recognized the team-based HRI, where role-specific actions should be coordinated simultaneously. Therefore, while previous studies have used the term "robot tutor", we propose a more nuanced characterization by introducing the notions of the robot as *learning assistants* and *evaluation assistants*.

Robot as learning assistants for skill-based training. Our study examined both knowledge-based and skill-based training. A direct comparison between the two provides clearer insight into their differences. Robots for cognitive learning emphasize verbal interaction to enhance engagement, using this interaction as a basis for evaluation and personalized feedback [11, 15]. For example, participants highlighted the importance of encouraging trainees to "think aloud" so that the robot could analyze the rationale behind their answers.

In contrast, robots as learning assistants for skill-based training emphasize perception-driven evaluation [21] and provide non-verbal feedback to prompt trainees to continue or abandon current actions. In our study, this was exemplified by locking or unlocking drawers to guide the correct use of the crash cart. In this context, the robot does not act as a tutor directly providing correct answers, but instead scaffolds the process so that trainees ultimately discover the correct actions themselves. This approach encourages active exploration of tools and strengthens operational skills. We argue that this new role expands the potential of robots in training and could be applied to a broader range of domains in the future.

Robot as evaluating assistants for team-based training. Prior HRI studies on robot training have largely emphasized dyadic interaction, with relatively little attention paid to multi-user or team-based training that more closely mirrors real-world contexts [17, 18, 21]. Robots are often considered less advantageous in these scenarios because they cannot fully interpret the complex social dynamics of group interaction. However, training contexts differ in that, despite the simultaneous role-specific actions taking place, these behaviors are typically structured around standardized protocols [22]. Our findings indicate that robots may, in fact, be better suited than humans to capture such details, providing more accurate and quantifiable records of team performance. By time-stamping actions, these records also enable retrospective reference among different actors.

Unlike studies that emphasize the robot's intervention in coordinating complex team dynamics [9], training contexts benefit more from robots that remain quiet and precise, providing a space for trainees to reflect rather than shape their behavior. Embodying the robot in a familiar artifact—the crash cart—further enhances this advantage, as it does not disrupt the existing social dynamics of the team. We argue that such unobtrusive evaluation is particularly valuable in high-fidelity simulations.

4.2 Design Guidelines for Crash Cart Robots for Code Training

Multimodality. As discussed, medical training involves both knowledge and skill. Although our initial aim was to support crash cart training, during the co-design sessions, most participants closely associated crash cart use with code practice. They envisioned the robot as a platform that integrates both knowledge of code protocols and skill training on the crash cart. This raises higher expectations for multimodality. While our current design of the crash cart robot already incorporates several modalities, participants suggested additional physical interventions, such as locking or unlocking drawers, to scaffold exploration and reinforce operational accuracy.

Time measurement for life-critical training. Robots can provide a critical form of measurement in code training: precise timing. Time is a defining feature of codes: in real situations, the timeliness of actions is directly tied to patient outcomes, and healthcare workers operate under this pressure. In simulation, however, the absence of real patient feedback makes it challenging to fully engage trainees. By introducing precise time measurement, the robot both reflects the realities of code practice and links performance to timing, enabling trainees to recognize the seriousness of their actions.

Validity of Robot Feedback. A major concern raised by participants was the validity of individualized feedback provided by the robot. We argue that ensuring validity may be even more important in training than in real codes, since trainees are expected to face challenges in recognizing errors or inconsistencies. Robots, therefore, need to be transparent about the clinical protocols on which they are programmed and offer evidence-based explanations for their feedback, so that trainees can build trust in the training process.

4.3 Limitations & Future Work

Our study has several limitations that should be acknowledged. First, the context of our research is rooted in nursing education in the United States. While we propose general guidelines for crash cart robots, training practices may differ across regions due to variations in culture and healthcare systems. Future research could build on our findings by grounding investigations in diverse contexts to enhance ecological validity.

Second, although we initially learned that crash cart training often originates in hospital orientation programs, we did not have the opportunity to directly observe such sessions. Instead, our work focused on co-design sessions that combined crash cart training with code training. The feasibility and effectiveness of this integration will need to be examined in actual training contexts in the future.

Finally, although our study contributes a more nuanced understanding of the role of robots in training, our findings are situated within the specific context of code training. Codes are highly distinctive even within medicine—life-critical but infrequent events—which may limit the generalizability of these models of HRI to other domains.

5 CONCLUSION

Our study provides new perspectives on understanding the role of robots in training and education. By comparing practical skill versus cognitive knowledge, and dyadic versus team-based interaction, we extend the conceptualization of robots beyond the traditional "tutor" to more contingent roles as learning assistants and evaluating assistants. We translate these insights into design guidelines to support these roles and highlight the importance of accuracy in safety-critical environments such as codes. Finally, we encourage future studies to adopt and refine these interaction techniques, advancing the integration of robots into broader contexts.

REFERENCES

[1] Michael Nnaemeka Ajemba, Chinweike Ikwe, and Judith Chioma Iroanya. 2024. Effectiveness of simulation-based training in medical education: assessing the impact of simulation-based training on clinical skills acquisition and retention: a systematic review. *World Journal of Advanced Research and Reviews* 21, 1 (2024), 1833–1843.

[2] Adriana Daniela Banyai and Cornel Brișan. 2024. Robotics in physical rehabilitation: Systematic Review. In *Healthcare*, Vol. 12. MDPI, 1720.

[3] Kathy Charmaz. 2014. *Constructing grounded theory*. sage.

[4] Hsin-Yu Chen, Pei-Ying Chen, Gwo-Jen Hwang, and Shan-Hung Wu. 2026. Beyond surgical applications: A systematic review of educational robotics in health professional development. *Nurse Education in Practice* (2026), 104734.

[5] Angelo Dante, Carmen La Cerra, Vittorio Masotta, Valeria Caponnetto, Luca Bertocchi, Alessia Marcotullio, Fabio Ferraiuolo, Celeste M Alfes, and Cristina Petrucci. 2022. The use of robotics to enhance learning in nursing education: a scoping review. In *International Conference in Methodologies and intelligent Systems for Techhnology Enhanced Learning*. Springer, 217–226.

[6] Chukwuka Elendu, Dependable C Amaechi, Alexander U Okatta, Emmanuel C Amaechi, Tochi C Elendu, Chiamaka P Ezeh, and Ijeoma D Elendu. 2024. The impact of simulation-based training in medical education: A review. *Medicine* 103, 27 (2024), e38813.

[7] Zhifeng Huang, Chingszu Lin, Masako Kanai-Pak, Jukai Maeda, Yasuko Kitajima, Mitsuhiro Nakamura, Noriaki Kuwahara, Taiki Ogata, and Jun Ota. 2016. Impact of using a robot patient for nursing skill training in patient transfer. *IEEE Transactions on Learning Technologies* 10, 3 (2016), 355–366.

[8] Gabrielle A Jacquet, Bachar Hamade, Karim A Diab, Rasha Sawaya, Gilbert Abou Dagher, Eveline Hitti, and Jamil D Bayram. 2018. The Emergency Department Crash Cart: A systematic review and suggested contents. *World journal of emergency medicine* 9, 2 (2018), 93.

[9] Malte F Jung, Nikolas Martelaro, and Pamela J Hinds. 2015. Using robots to moderate team conflict: the case of repairing violations. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*. 229–236.

[10] Roshni Kaushik, Rayna Hata, Aaron Steinfeld, and Reid Simmons. 2025. Choosing Robot Feedback Style to Optimize Human Exercise Performance. In *2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 658–666.

[11] James Kennedy, Paul Baxter, Emmanuel Senft, and Tony Belpaeme. 2016. Social robot tutoring for child second language learning. In *2016 11th ACM/IEEE international conference on human-robot interaction (HRI)*. IEEE, 231–238.

[12] Hee Rin Lee, Selma Šabanović, and Sonya S Kwak. 2017. Collaborative map making: A reflexive method for understanding matters of concern in design research. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 5678–5689.

[13] Daniel Leyzberg, Samuel Spaulding, and Brian Scassellati. 2014. Personalizing robot tutors to individuals' learning differences. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. 423–430.

[14] Noel Maalouf, Abbas Sidaoui, Imad H Elhajj, and Daniel Asmar. 2018. Robotics in nursing: a scoping review. *Journal of Nursing Scholarship* 50, 6 (2018), 590–600.

[15] Nikolas Martelaro, Victoria C Nneji, Wendy Ju, and Pamela Hinds. 2016. Tell me more designing HRI to encourage more trust, disclosure, and companionship. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 181–188.

[16] Margaret McAllister, Krystle Kellenbourn, and Denise Wood. 2021. The robots are here, but are nurse educators prepared? *Collegian* 28, 2 (2021), 230–235.

[17] Margie Molloy, Ryan J Shaw, Jackie Vaughn, and Remi Hueckel. 2016. An innovative use of telepresence robots for educating healthcare professional. In *Nursing informatics 2016*. IOS Press, 989–990.

[18] Maryam Moosaei, Sumit K Das, Dan O Popa, and Laurel D Riek. 2017. Using facially expressive robots to calibrate clinical pain perception. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. 32–41.

[19] Maryam Moosaei, Cory J Hayes, and Laurel D Riek. 2015. Facial expression synthesis on robots: An ros module. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts*. 169–170.

[20] Maryam Moosaei, Maryam Pourebadi, and Laurel D Riek. 2019. Modeling and synthesizing idiopathic facial paralysis. In *2019 14th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2019)*. IEEE, 1–8.

[21] Peizhu Qian, Filip Bajraktari, Carlos Quintero-Peña, Qingxi Meng, Shannan Hamlin, Lydia Kavraki, and Vaibhav Unhelkar. 2024. ASTRID: A Robotic Tutor for Nurse Training to Reduce Healthcare-Associated Infections. *Decision Making* 20, 47 (2024), 116.

[22] Carlos Quintero-Pena, Peizhu Qian, Nicole M Fontenot, Hsin-Mei Chen, Shannan K Hamlin, Lydia E Kavraki, and Vaibhav Unhelkar. 2023. Robotic tutors for nurse training: Opportunities for hri researchers. In *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 220–225.

[23] A Masie Rahimi, Ezgi Uluç, Sem F Hardon, H Jaap Bonjer, Donald L van der Peet, and Freek Daams. 2024. Training in robotic-assisted surgery: a systematic review of training modalities and objective and subjective assessment methods. *Surgical Endoscopy* 38, 7 (2024), 3547–3555.

[24] Laurel D Riek. 2017. Healthcare robotics. *Commun. ACM* 60, 11 (2017), 68–78.

[25] Yeisson Rivero-Moreno, Sophia Echevarria, Carlos Vidal-Valderrama, Luigi Pianetti, Jesus Cordova-Guilarte, Jhon Navarro-Gonzalez, Jessica Acevedo-Rodríguez, Gabriela Dorado-Avila, Luisa Osorio-Romero, Carmen Chavez-Campos, et al. 2023. Robotic surgery: a comprehensive review of the literature and current trends. *Cureus* 15, 7 (2023).

[26] Bartosz Sawik, Sławomir Tobis, Ewa Baum, Aleksandra Suwalska, Sylwia Kropińska, Katarzyna Stachnik, Elena Pérez-Bernabeu, Marta Cildoz, Alba Agustin, and Katarzyna Wieczorowska-Tobis. 2023. Robots for elderly care: review, multi-criteria optimization model and qualitative case study. In *Healthcare*, Vol. 11. MDPI, 1286.

[27] Henk WR Schreuder, Richard Wolswijk, Ronald P Zweemer, Marlies P Schijven, and René HM Verheijen. 2012. Training and learning robotic surgery, time for a more structured approach: a systematic review. *BJOG: An International Journal of Obstetrics & Gynaecology* 119, 2 (2012), 137–149.

[28] Ashwin N Sridhar, Tim P Briggs, John D Kelly, and Senthil Nathan. 2017. Training in robotic surgery—an overview. *Current urology reports* 18, 8 (2017), 58.

[29] Adel Tutuo Tamata and Masoud Mohammadnezhad. 2023. A systematic review study on the factors affecting shortage of nursing workforce in the hospitals. *Nursing open* 10, 3 (2023), 1247–1257.

[30] Angelique Taylor, Hee Rin Lee, Alyssa Kubota, and Laurel D Riek. 2019. Coordinating clinical teams: Using robots to empower nurses to stop the line. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–30.

[31] Angelique Taylor, Tauhid Tanjim, Huajie Cao, and Hee Rin Lee. 2024. Towards collaborative crash cart robots that support clinical teamwork. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 715–724.

[32] Candice Vander Weerdt, Jessica A Peck, and Tracy Porter. 2023. Travel nurses and patient outcomes: a systematic review. *Health care management review* 48, 4 (2023), 352–362.

[33] Induni N Weerarathna, David Raymond, and Anurag Luharia. 2023. Human-robot collaboration for healthcare: a narrative review. *Cureus* 15, 11 (2023).

[34] Guang-Zhong Yang, James Cambias, Kevin Cleary, Eric Daimler, James Drake, Pierre E Dupont, Nobuhiko Hata, Peter Kazanzides, Sylvain Martel, Rajni V Patel, et al. 2017. Medical robotics—Regulatory, ethical, and legal considerations for increasing levels of autonomy. eaam8638 pages.

Received 2025-09-30; accepted 2025-12-01