

International Journal of Human–Computer Interaction

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hihc20

Usability of Automated External Defibrillators: A Randomized, Comparative Simulator Study

Beatrice Coldewey, Andreas Klausen, Fabian Otto-Sobotka, Rainer Röhrig & Myriam Lipprandt

To cite this article: Beatrice Coldewey, Andreas Klausen, Fabian Otto-Sobotka, Rainer Röhrig & Myriam Lipprandt (08 Oct 2023): Usability of Automated External Defibrillators: A Randomized, Comparative Simulator Study, International Journal of Human–Computer Interaction, DOI: 10.1080/10447318.2023.2260973

To link to this article: https://doi.org/10.1080/10447318.2023.2260973

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.

View supplementary material

Published online: 08 Oct 2023.

Submit your article to this journal

Article views: 855

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at https://www.tandfonline.com/action/journalInformation?journalCode=hihc20

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
https://doi.org/10.1080/10447318.2023.2260973

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Usability of Automated External Defibrillators: A Randomized, Comparative Simulator Study

Beatrice Coldewey[a,b], Andreas Klausen[b,c], Fabian Otto-Sobotka[d], Rainer Röhrig[a,b,c], and Myriam Lipprandt[a,b,c]

[a]Institute of Medical Informatics, Medical Faculty of RWTH Aachen University, Aachen, Germany; [b]Department of Health Services Research, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany; [c]Oldenburger Research Network of Emergency Services and Intensive Care, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany; [d]Division of Epidemiology and Biometry, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

ABSTRACT

The aim of this study was to investigate whether one particular automated external defibrillator (AED) is superior to another in use by a layperson such that there is a clinically relevant difference from first contact with the device to the time of shock delivery (TimeToShock (TTS)). Causes for interaction problems and delays were analyzed. The study was conducted as a randomized, two-arm manikin study using the AED training devices "Lifepak CR-T AED Trainer" (Medtronic Physio-Control) and "AED Trainer 3" (Philips Medical Systems). In each study arm, the participants used one of the AEDs in a simulated resuscitation. The interaction was recorded by a mobile eye tracker. The TTS was compared with the Wilcoxon rank-sum test. The study was conducted with 28 participants inexperienced in AED use. A significant difference ($p < 0.05$) in TTS was measured at 32.5 s median between the two devices (61 s (IQR: 58 s–76 s, range: 51 s–115 s) for the AED Trainer 3 and 93.5 s (IQR: 91 s–103 s, Range: 80 s–129 s) for the Lifepak CR-T AED Trainer). Causes of delays and interaction problems included differences in the use of voice instruction (length, pauses, and terminology), iconography on the device and electrodes, and device opening mechanisms. The measured clinically relevant difference can be attributed to the usability of the interaction elements, which differ significantly between the devices. Besides specific design elements, the extent of user guidance provided by the AEDs is crucial for efficient and error-free interaction.

KEYWORDS

Resuscitation; automated external defibrillators; ergonomics; usability testing; patient safety; use error

1. Introduction

Out-of-hospital cardiac arrest (OHCA) is a life-threatening condition (Gräsner et al., 2021; Gräsner et al., 2020; Kiguchi et al., 2020). In Europe, OHCA occurs with an incidence of 67–170 per 100,000 inhabitants per year (Gräsner et al., 2021). The later the resuscitation is started, the less likely people are to survive or survive without neurological damage (Gold et al., 2010; Larsen et al., 1993). The survival rate after an OHCA at the time of hospital discharge in Europe is 8% on average, varying from 0–18% (Gräsner et al., 2021).

In the case of ventricular fibrillation (VF) or pulseless ventricular tachycardia (pVT), early defibrillation can improve the survival rate in an OHCA (Bækgaard et al., 2017; Brooks et al., 2022; Hawkes et al., 2017; Olasveengen et al., 2021; Weaver et al., 1986). For delayed defibrillation, there was a decrease in survival of 1.1% per minute for witnessed cardiac arrest due to VF, calculated using linear regression (Larsen et al., 1993). A linear model shows that shortening the time to defibrillation by an average of 3 min 30 s appeared to improve overall yearly survival by 10–13% (Weaver et al., 1986). If continuous CPR is started immediately, there is an absolute risk reduction of about 2% per 30 s (Valenzuela et al., 1997). This is equivalent to a number needed to treat of 50 for 30 s time reduction.

Up to 20 min can elapse from the time of the emergency call to the arrival of the emergency services (Schehadat et al., 2017). This interval must be bridged by laypersons (Dainty et al., 2022; Olasveengen et al., 2021). The European Guidelines for Resuscitation recommended the use of automated external defibrillators (AEDs) in lay resuscitation since 1998 (Handley et al., 1998; Olasveengen et al., 2021). For non-trained persons, resuscitation is challenging, especially when it involves the use of an AED. This results in high demands on the human–machine interaction to quickly deliver a shock: it can be assumed that the users, often laypersons without emergency medical training, are operating this type of device for the first time. Safe and efficient use requires an ad hoc understanding of the interface and

CONTACT Myriam Lipprandt mlipprandt@ukaachen.de University Hospital Aachen, Institute of Medical Informatics, Pauwelsstraße 30, 52074 Aachen, Germany

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10447318.2023.2260973.

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

unambiguous instructions (Fairbanks & Caplan, 2004; Food and Drug Administration, 2016; IEC 62366-1, 2015). Various AEDs with different use concepts are available on the market (Ho et al., 2016; Torney et al., 2018). Studies have demonstrated that effectiveness and efficiency differ between AEDs (Andre et al., 2004; Callejas et al., 2004; Dong et al.., 2020; Eames et al., 2003; Fleischhackl et al., 2004; Lei et al., 2022; Polikaitis, 2005) and ergonomic features affect the speed of delivering a shock (Mosesso et al., 2009). An analysis of user interface features in terms of effectiveness, efficiency and possible use errors, especially with eye trackers, has not yet been carried out or published. We set 30 s as a clinically relevant value for the different duration of use of AEDs.

The objective of this study was to investigate whether one of the two AEDs is superior in terms of efficiency of use, such that there is a clinically relevant time difference from first contact with the device to the time of first shock delivery (TimeToShock (TTS)). The secondary objective was to identify the causes of problems in human-machine interaction and delays during the use of the devices.

2. Methods

2.1. Study design

A randomized, comparative manikin study was conducted. The participants used an AED in a simulated resuscitation. The two study arms differed in the type of AED used. The participants were randomly assigned and equally divided between the study arms. The study was positively evaluated by the ethics committee (ethics vote 2018-082, Medical Ethics Committee of the Carl von Ossietzky University of Oldenburg, Chairman Prof. Dr Frank Griesinger).

2.2. Setting and equipment used

AED trainers used for the study were AED Trainer 3 (Philips Medical Systems, model 861467) with a soft case and LIFEPAK CR-T AED Trainer (Medtronic Physio-Control, model 3201804-012). In this paper, AED Trainer 3 is referred to as AED1 and LIFEPAK CR-T AED Trainer as AED2. The device features are described in Table 1. The CPR training model "Resusci Anne" (Laerdal Medical AS) served as a manikin. The participants wore eye-tracker glasses (Tobii Pro Glasses 2).

The simulated resuscitation was performed in a quiet meeting room on the manikin lying on a solid base. The participant was asked to use the AED while a supervisor from the study team performed the CPR (Figure 1). The AED was positioned 3 m away in a location known to the participant.

2.3. Participants

Medical laypersons were recruited for the study. The inclusion criteria were age 18–65 years and written informed consent. The exclusion criteria were medical or nursing training or professional experience with emergencies (e.g. members of fire brigades) as well as an individual training in defibrillators (but not participation in a demonstration of an AED, e.g., as part of a first aid course) so that the tests would not be influenced by people's previous experience and skills to deal with emergency situations. Additionally, people with health limitations that restrict or prevent resuscitation of a person (manikin) were excluded (see Figure 3).

2.4. Procedures

After giving informed consent, the participant completed a questionnaire on demographic data. After the start signal to begin the simulation, the participant had the task of fetching the AED and attaching it to the manikin to deliver a shock while a supervisor of the study team performed the CPR continuously. The simulation follows the ERC BLS guide (Olasveengen et al., 2021): 1. Check response and breathing (no response or breathing; both actors); 2. Call for help (both actors); 3. Start CPR (supervisor) and get AED (participant); 4. Once AED arrives – switch on and follow instructions (participant, AED instructions see Table 1) while continuing CPR (supervisor). The use of the AED includes: 1. Switching on the AED; 2. Calling for help (AED2 only); 3. Applying the electrodes (clothing was removed from the patient's chest by the supervisor before CPR is started); 4. Rhythm analysis (initiated by AED); 5. Delivering a shock. The simulation was terminated after the shock was delivered. Subsequently, the participant evaluated the usability of the AED and their comments on the usage were discussed.

The AED model was randomly assigned (drawn without return by the participant). The aim of the study was kept from the participant to avoid an exam-like situation. The participant was told that personal interaction during resuscitation was the main objective. After the study, the participant was debriefed and explained the study's objectives.

2.5. Outcomes

Figure 2 shows an overview of the collected times. The primary outcome is the difference between study arms in time from first contact with the AED (T1) to delivery of the first shock (T5)–TimeToShock (TTS).

The secondary outcomes regarding the individual subtasks are the task completion and difference between time from first contact to the start of the AED (TimeToStart), time from the start of the AED to the start of the analysis (TimeStartToAnalysis) and time from the end of the analysis to the delivery of the first shock (TimeEndAnalysisToShock). Further secondary outcomes for comparing the usability of the devices are the subjective evaluation of usability by the participants and use errors during interaction with the AEDs and underlying causes.

2.6. Data acquisition and analysis

The simulation was recorded by video camera and eye-tracker glasses worn by the participant in order to examine

Table 1. Features of AED1 and AED2 (voice instructions translated into English, original German version see supplemental material).

Features	AED1	AED2
		
Power-on mechanism	On/off button on inside of a box with velcro	On/off button for opening the lid on inside of a box with velcro
Start of Analysis of the heart rhythm	The test supervisor triggered the start of the analysis via the training AED remote control.	The test supervisor triggered the start of the analysis via the training AED remote control.
Visual instructions	Power LED; shock button LED; illustration of the manikin with the position of both electrodes on AED	Shock button LED; handling instructions and illustration of the manikin with the position of both electrodes on electrode package; illustration of the manikin with the position of both electrodes and electrode pad indicator LEDs on AED
Pad design	Two electrodes on a foil inside the box next to the AED, preconnected plug; frontal view on a manikin with the position of both electrodes, indication of the respective electrode position by arrows	Two electrodes on a foil in a tear-open package under the opening lid, pre-connected plug; frontal/lateral view on a manikin with the position of the respective electrode
Voice instructions		
Power up	Beep sound + 3.5 s break	Beep sound + "Attention, follow the instructions of the device"
Call for help	/	"Call for emergency help" + 2 s break
Pad placement	"Apply pads to the exposed chest of the patient" (repetition after 12 s (varying instructions) until electrodes are connected)	"Remove all clothing from patient's chest" + 10s break + "pull red handle to open bag and pull out pads" + 4s break + "look at the pictures on the pads" + 1s break + "peel one pad off blue plastic" + 1s break + "apply pad to bare skin exactly as shown in the picture" + 1s break + "press pad firmly to chest" + 1.5s break + "peel other pad off blue plastic" + 1s break + "apply pad to bare skin exactly as shown in the picture" + 1s break + "press pad firmly to chest" (repetition after 3s (varying instructions) until electrodes are connected)
Rhythm analysis	"Do not touch the patient. Analysis running" + 6.5s break	"Attention, do not touch the patient. Evaluating heart rhythm" + 3.5s break
Analysis result	"Shock recommended. Do not touch patient" + 1.5s break	"Stand by, preparing to shock" + 3.5s break + "Attention, now no one is allowed to touch the patient"
Shock delivery	"Deliver shock now" + beep sound (repetition after 3s: "Press flashing orange button now")	"Press flashing button" + beep sound
Resuming CPR	"Shock delivered" + 1.5s break "provide chest compressions and rescue breaths"	"Shock delivered" + 1s break + "provide 30 chest compressions and 2 rescue breaths alternately"
Overall operation time (power up to shock delivery)	27 s	83 s

the simulation in detail following the study with regard to primary and secondary endpoints. The eye tracker was calibrated with every participant before the start of the simulation. The demographic data and the subjective evaluation of usability were collected in questionnaires.

The digital transfer of questionnaires was performed by two researchers. In the case of deviations, the results were checked against the original data. The analysis of the quantitative data was carried out by two independent researchers and double programming with the statistical software



Figure 1. Picture of the setup of the simulated resuscitation scenario (two study managers as actors).

packages SPSS 26 and R 3.6.3. Additionally, the use errors were consolidated by discussion in order to reduce or control for individual confounding and avoid errors in data analysis.

2.6.1. Analysis of task performance and time intervals

The tasks were considered completed when the respective goal (start AED/position electrodes/deliver shock) had been achieved. The calculation of the time intervals was carried out by determining the time points T1–T5 to an accuracy of one second from the recordings. The determined periods were averaged and the variability between the results was calculated. Delays in AED operation caused by the test supervisor, e.g., due to the remote control of the AEDs to start the analysis, were taken into account and subtracted from the overall time. For AED1, some participants placed the electrodes before the AED was switched on contrary to the regular procedure. After that, the voice prompt for attaching the electrodes was given and lasted for 3.5 s. In the main analysis, difference of TTS was assessed using the Wilcoxon rank-sum test.

2.6.2. Analysis of the usability questionnaire

Usability was evaluated using 11 closed questions on a four-point Likert scale: eight questions from the standardized questionnaire System Usability Scale (SUS) (J. Brooke, 1996) and three individual questions. The procedure described by Brooke (J. B. Brooke, 2013; J. Brooke, 1996) for calculating a global usability score was adapted accordingly. The analysis was descriptive with the calculation of mean values, standard deviations and 95% confidence intervals (bootstrapping, sample size 1000) for both groups.

2.6.3. Analysis of use errors and underlying causes

The participants' statements as well as the recordings of the simulation were examined for interaction problems with the AED and usability issues as possible cause for unsuccessful or delayed shock delivery. Analysis of the use errors is based on risk management concepts according to ISO 14971, 2019. Use errors during the use of a medical device lead to a different result than intended by the manufacturer or expected by the user (IEC 62366-1, 2015). The consequences of use errors are interaction errors (e.g. wrong button pressed), a loss of time, which is a clinically relevant parameter, or user uncertainty. The interpretation of the underlying causes is based on assumptions about the context of use and users' mental models, as well as interaction principles (ground rules) for designing usable products (Geis & Johner, 2020; IEC 62366-1, 2015; ISO 9241-110; Wiklund et al., 2016). We have determined the occurrence (frequency) for each use error per study arm (AED1 vs. AED2) and specified the 95% confidence interval.

2.7. Sample size calculations

To test the primary hypothesis, a group sequential design with two stages was planned. There was no information about the expected distribution of possible differences and standard deviations of the data. For the first stage, the assumption was made that the mean difference between the two groups was one standard deviation (significance level: $\alpha = 0.05$, power $= 0.8$). This resulted in a required sample

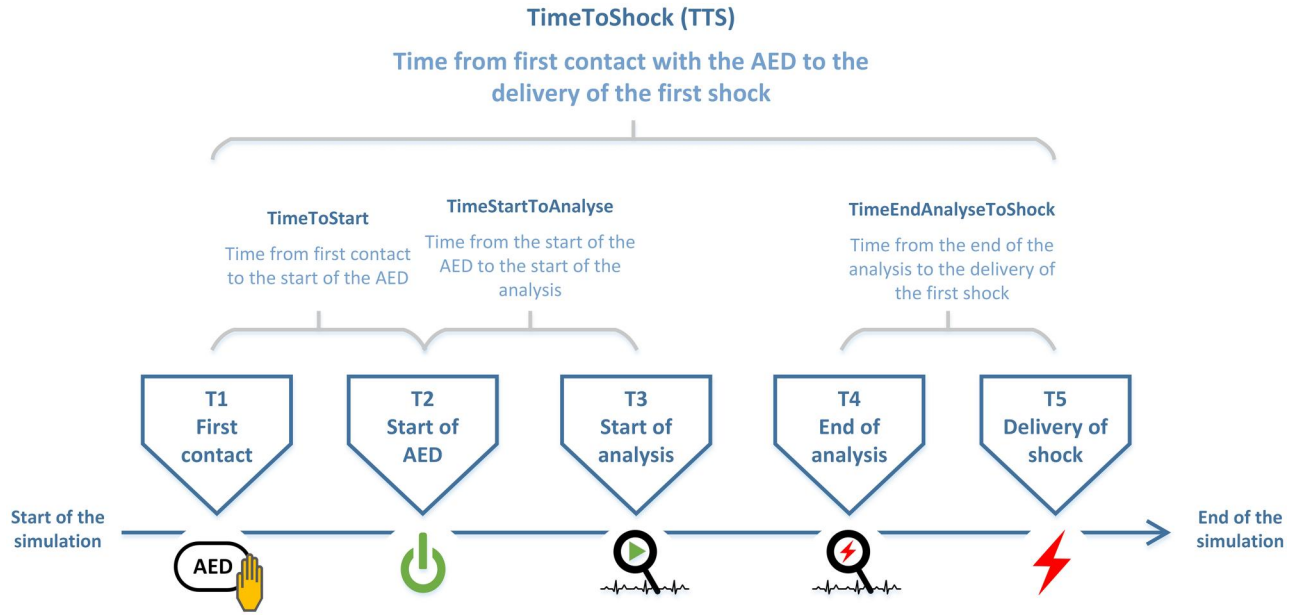


Figure 2. Relevant time points during the interaction with the AED and corresponding time intervals. The time points are defined as follows: T1: First contact: first haptic contact when fetching the AED; T2: Start of AED: pressing the power button; T3: Start of analysis: start of the acoustic audio output at the start of the analysis; T4: End of analysis: start of the acoustic audio output for the analysis result; T5: Delivery of shock: pressing the shock button after being instructed by the AED.

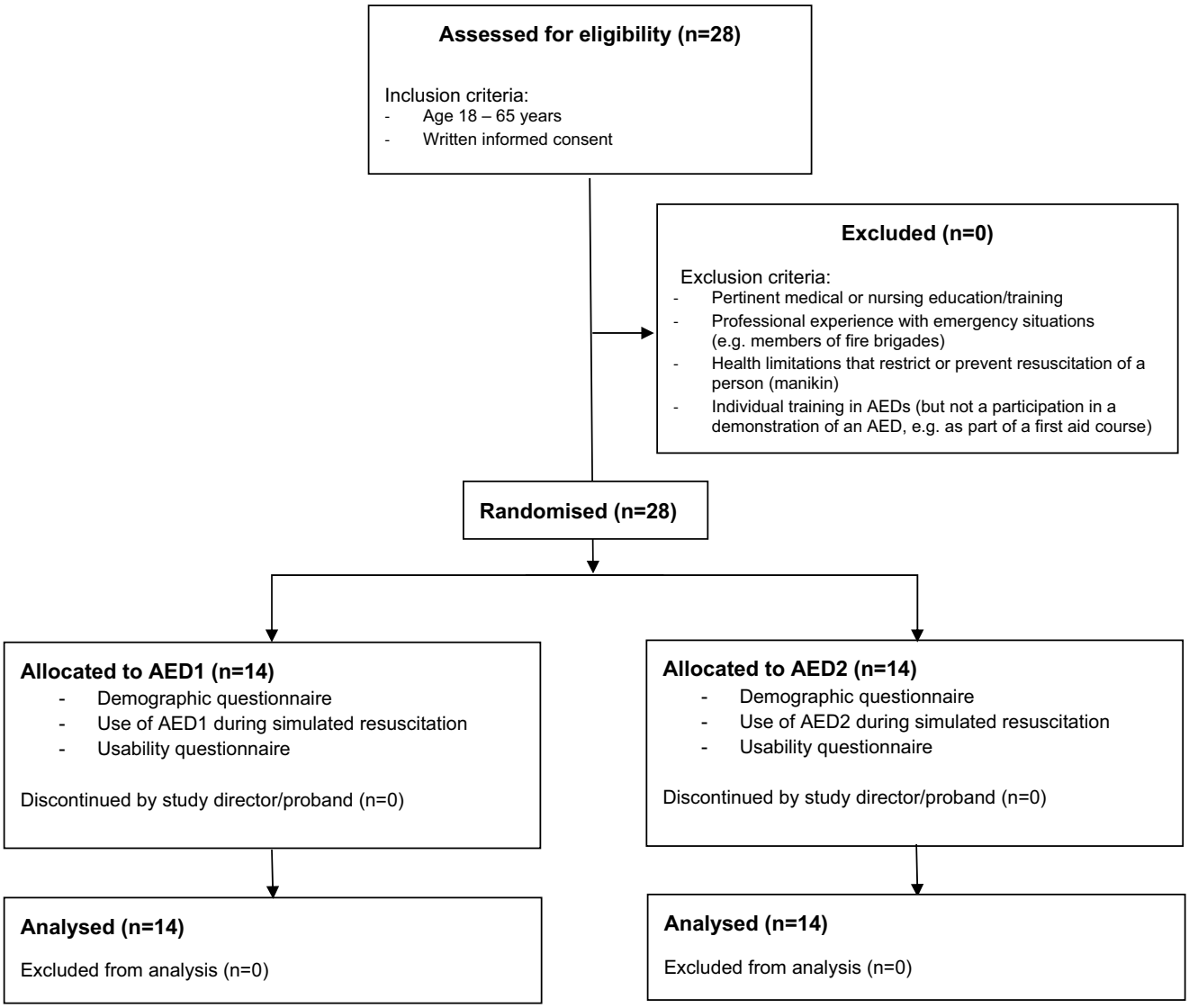


Figure 3. Flow diagram of the study procedure.

Table 2. Demographic data of the participants.

	Total		AED1		AED2	
	N	%	*n*	%	*n*	%
Gender						
Female	20	71.4	11	78.6	9	64.3
Male	8	28.6	3	21.4	5	35.7
Diverse	0	0.0	0	0.0	0	0.0
Total	28	100	14	100	14	100
Age						
18–25 years	4	14.3	1	7.1	3	21.4
26–35 years	17	60.7	9	64.3	8	57.1
36–45 years	2	7.1	0	0.0	2	14.3
46–55 years	1	3.6	1	7.1	0	0.0
56–65 years	4	14.3	3	21.4	1	7.1
Total	28	100	14	100	14	100
Highest level of school education						
Hauptschulabschluss	0	0.0	0	0.0	0	0.0
Realschulabschluss	4	14.3	4	28.6	0	0.0
Abitur	8	28.6	4	28.6	4	28.6
Hochschulabschluss	16	57.1	6	42.9	10	71.4
Total	28	100	14	100	14	100

of 28 participants (equal distribution) for a Wilcoxon rank-sum test. For the early successful termination of the study after the interim analysis (futility level: $\alpha_0 = 0.5$), the significance level of α_1=0.0044 was calculated according to the method of inverse combination of the normal distribution. In the case that $\alpha_1 < p < \alpha_0$ applied to the p-value of the interim analysis, the case number planning for the entire study would be carried out using the available data. In the final analysis, the original significance level $\alpha = 0.05$ could then be used. Due to the compact nature of the study, no dropouts were expected.

3. Results

3.1. Characteristics of study subjects

The study was conducted in three sessions between April 2019 and July 2019. A total of 28 participants were included (Figure 3). The demographic characteristics of the participants and their distribution between the study arms are shown in Table 2. All participants were able to communicate fluently in German. None of the participants had practical experience in the use of AEDs.

3.2. Main results

3.2.1. Task performance and execution times

3.2.1.1. Primary outcome. All participants successfully delivered a shock using an AED. For the primary outcome (TTS), a median of 61 s (IQR: 58.5 s – 75.8 s, range: 51.5 s – 114.5 s) was measured for AED1 and 93.5 s (IQR: 91.3 s – 102.1 s, range: 80 s – 128.5 s) for AED2 (Figure 4). The median difference in operating time was 32.5 s. The p-value of the interim analysis was p_1=0.0005 and thus below the significance level defined for the interim analysis. The study was terminated accordingly after the interim analysis.

3.2.1.2. Secondary outcomes. The subtask of starting the AED was successfully performed by all participants. Only one participant pressed the shock button on AED1 directly without first switching on the device (AED1: $n = 1$, 7%, 95%-CI: 0%–22%). AED1 is designed in such a way that this action switches on the device.

The positioning of the electrodes was incorrect in three cases: (1) One participant (AED1: $n = 1$, 7%, 95% CI: 0%−22%) swapped the electrodes. (2) One participant (AED2: n = 1, 7%, 95% CI: 0%−22%) positioned the electrodes on the other side of the body (left instead of right and vice versa). (3) One participant (AED2: $n = 1$, 7%, 95% CI: 0%−22%) positioned both electrodes on the same side of the body. He realized the mistake after delivering the shock.

For the subtasks, the median differences in execution times were: TimeToStart: −10.5 s (AED1: 18.8 s, AED2: 8.3 s), TimeStartToAnalyse: 39.3 s (AED1: 26 s, AED2: 65.3 s) and TimeEndAnalyseToShock: 4.2 s (AED1: 7.8 s, AED2: 12 s). Figure 4 shows a graphical representation of the data. The discrepancy between observer 1 and observer 2 for the calculation of TTS is a median of −1.0 s (IQR: −1.0 s – −1.0 s, range: −2.0 s − 1.0 s) for AED1 and a median of 0.0 sec (IQR: 0.0 s − 1.0 s, range: −1.0 s − 1.0 s) for AED2. Detailed data for the execution times and discrepancies are given in the supplemental material.

3.2.2. Usability questionnaire

For the SUS, the mean score for AED1 was 86.6 (SD: 5.2; 95% CI: 83.9 − 89.3) and for AED2 79.3 (SD: 10.5; 95% CI: 73.9 − 84.8). Only one question was not answered by a participant (Missing data: AED1 $n = 0$, AED2 $n = 1$). Figure 5 shows the detailed results of the usability questionnaire. The original German questionnaire is attached in the supplemental material.

3.2.3. Analysis of use errors and potential underlying causes

Table 3 shows a list of use errors for AED1 and AED2 as a possible cause for unsuccessful or delayed shock delivery. In total, there were 20 use errors for AED1 and 19 errors for AED2. On average, each participant made 1.39 errors (AED1: $n = 1.43$, AED2: $n = 1.36$). Selected screenshots of the eye-tracker recording of the use errors that occurred are given in the supplemental material.

4. Discussion

This randomized confirmatory study showed a significant, clinically relevant difference in time to shock delivery between two AEDs (AED1: 61 s, AED2: 93.5 s, $p < 0.05$). In the study of Eames et al. (2003), three AEDs were tested with a mean TTS of 74.8 s (range 60–90 s), 83.0 s (range 54–111 s) and 153.4 s (range 92–240 s). In the study of Fleischhackl et al. (2004), six AEDs were tested with a mean TTS from 78 s (95% CI: 68–89 s) to 128 s (95% CI: 110–146 s) and eight AEDs were tested with a mean TTS from 48 s (SD: ±2 s) to 101 s (SD: ±1 s) by Müller et al. (2015).

In addition to the time to first shock delivery, the correct placement of the electrodes is important. There was one

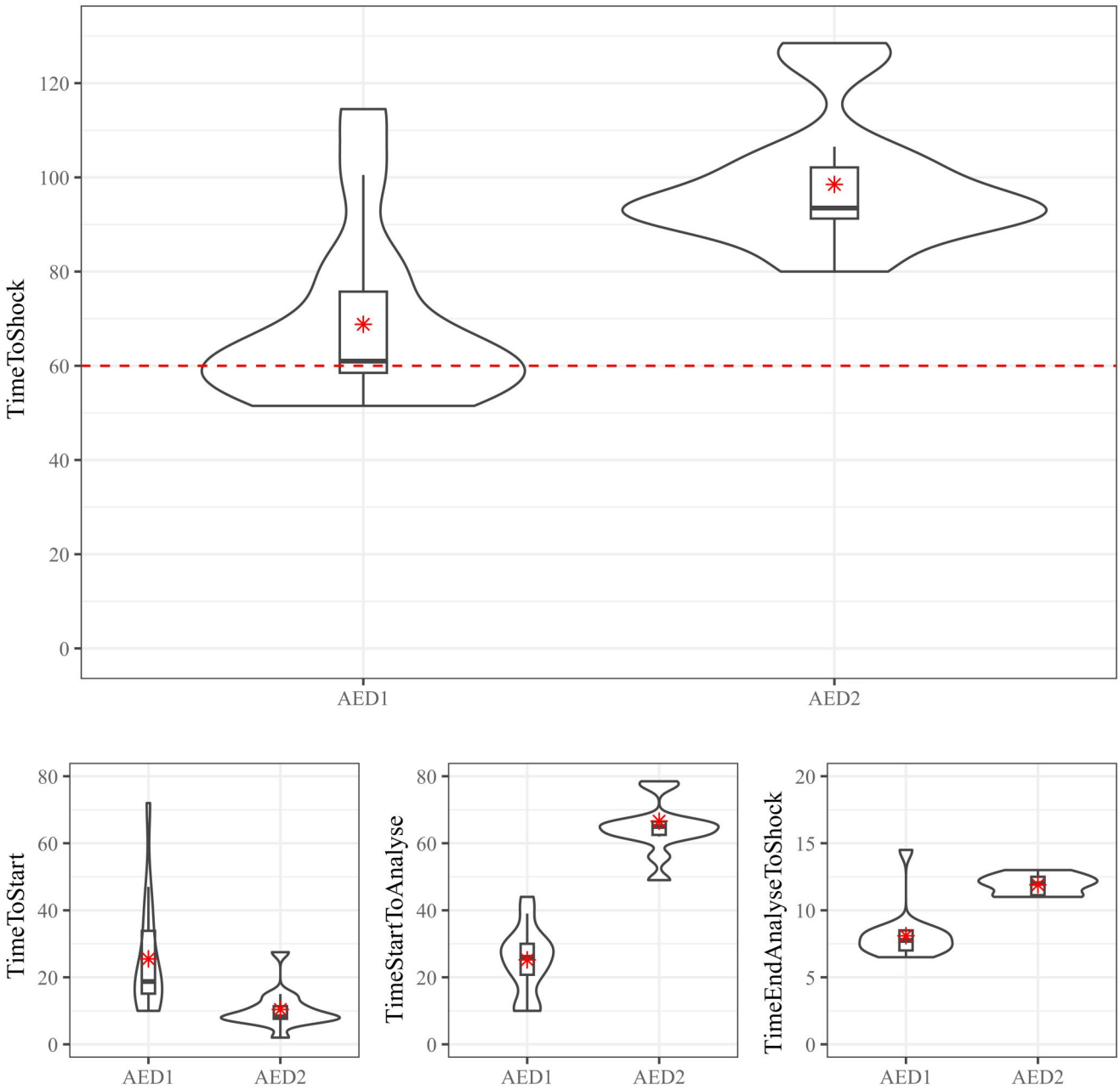


Figure 4. Box-whisker plots with the execution times for AED1 and AED2. The median, quartiles and the largest and smallest values within the 1.5 times the interquartile range above and below the quartiles are plotted. The dots represent measured values outside this range. The violin plot function shows the smoothed distribution within the range (minimum–maximum values). The red stars mark the mean value.

misplacement with AED1 and two with AED2. Of these, two were critical for AED2 and none for AED1. We identified several categories of use errors and usability issues as possible causes of unsuccessful or delayed shock delivery. This demonstrates that usability has a relevant impact on the use and therefore safety- and time-critical performance of an AED.

The AEDs have a minimum operating time set by the manufacturer because of the voice instructions and pauses (see Table 1). AED1 had less instructions and therefore an operation time of 27 s from pressing the power button to delivering a shock. In our study, the shortest time needed was 51.5 s. Seven participants needed less than 60 s. AED2 had an intended duration of 83 s with detailed voice instructions. The only way to reduce the TTS was to apply the electrodes independently of the voice instructions. Two participants "skipped" the voice instructions and required minimum 80 s. Most participants followed the voice instructions completely.

An additional delay occurred for a large number of participants due to the occurrence of use errors. The identified use errors could be divided into five topics (unpacking, attaching electrodes, delivering shocks, a different sequence of actions and asynchronous sequence of actions; see Table 3).

When unpacking the AED1, participants had problems identifying the opening mechanism of the packaging (a large velcro fastener). They also unnecessarily tried to remove the AEDs from their packaging. Longer operating times due to packaging have been observed by Mosesso et al. (2009) and offer an opportunity for improvement.

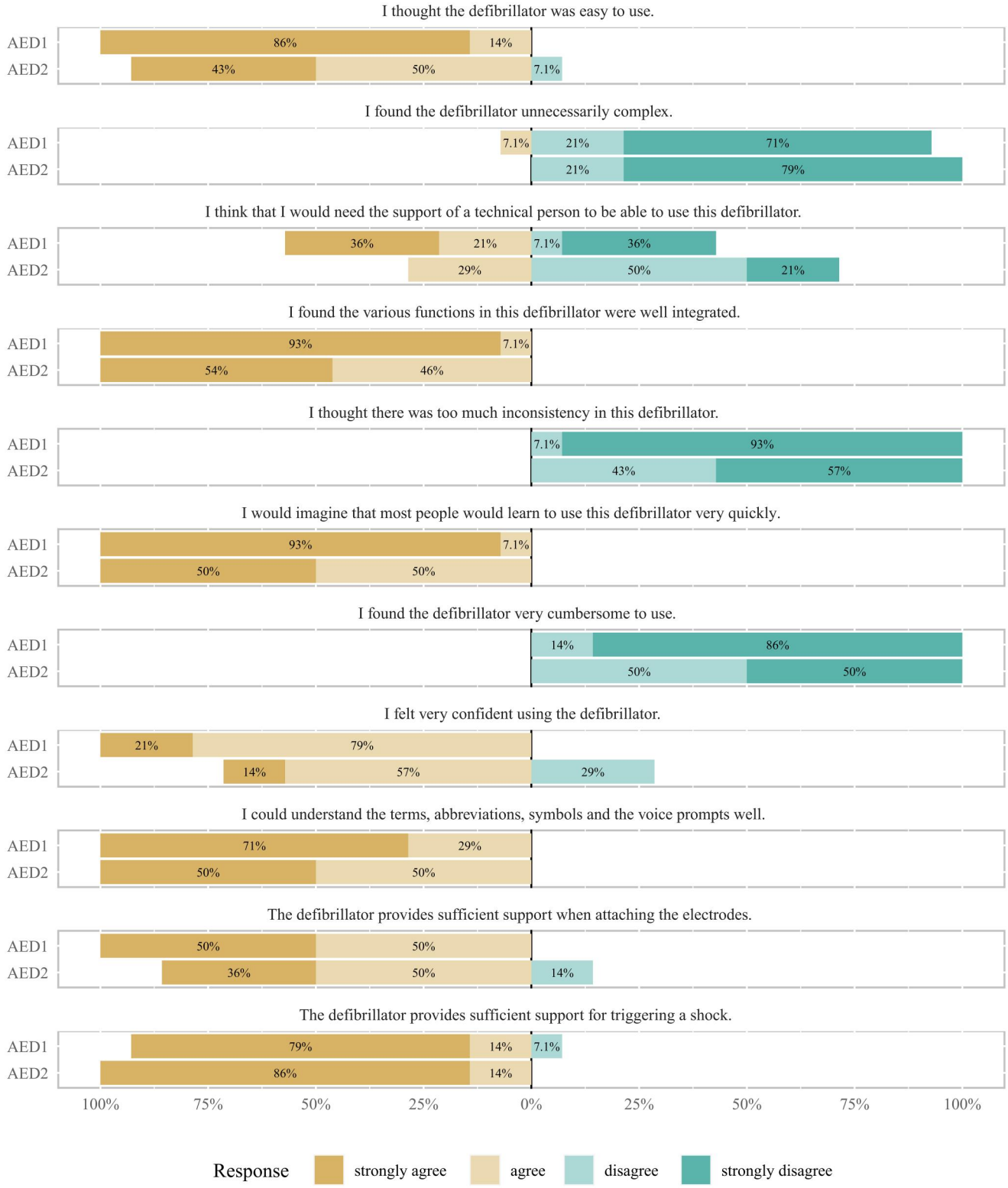


Figure 5. Results of the usability questionnaire for AED1 and AED2 (%-values rounded).

The errors in attaching the electrodes suggest problems regarding the design of the pictograms on the electrodes, which are crucial for correct positioning (Andre et al., 2004; Bødtker & Rosendahl, 2018; Foster & Deakin, 2019; Nurmi & Castrén, 2005; Polikaitis, 2005). While Nurmi and Castrén (2005) have observed that a lateral view results in a better placement of the apical electrode, in this study, electrode misplacement was observed for AED2 in two cases due to the perspective view. According to Foster and Deakin (2019), an optimal representation of images on adhesive electrodes is a rotated anterolateral view.

During shock delivery, the wording of the verbal instructions of AED1 was such that they did not provide an unambiguous instruction to act ("Deliver shock now" for AED1 vs "Press flashing button" for AED2). The instruction should be sufficiently informative or formulated as a precise request for

Table 3. Use errors that occurred in the simulated scenario, their frequency, consequences, and possible causes.

Category	Use error	Frequency		Consequences of use errors		Causes for use errors + violated interaction principles (ISO 9241-110)
		AED1	AED2	AED1	AED2	
Unpacking	Search for a mechanism to open the AED package (AED1 only) and/or (attempt to) remove the device from its package	9 (64%) 95%-CI: 35%–86%	6 (43%) 95%-CI: 14%–72%	Time loss of up to 20 s	Time loss of up to 5 s	Unclear recognition of the mechanism for opening the device (black border of the cover was mistaken for a zip) (self-descriptiveness) The envelope is perceived as packaging to be removed (conformity with user expectations)
Attaching electrodes	Electrodes incorrectly positioned: - swapping the electrodes (AED1 only) - Correction of positioning (AED1 only) - Deviation from specified positioning (side reversed) (AED2 only) - Deviation from specified positioning (both on one side of the body) (AED2 only)	2 (14%) 95%-CI: 0%–36%	2 (14%) 95%-CI: 0%–36%	Time loss due to correction of up to 8 s Ineffective shock delivery due to incorrect pad position	Ineffective shock delivery due to incorrect pad position	The arrow pictograms representing the electrode position are not perceived correctly (AED1) (self-descriptiveness) Pictograms on electrodes shows a perspective view of the patient, view of the AED user on the right side of the patient is assumed (AED2) (Suitability for the user's tasks)
Delivering shock	Shock delivery delayed after request	1 (7%) 95% CI: 0%–22%	0 (0%)	Time loss of up to 6 seconds, shock button pressed only after the second, differently worded prompt		No request for action was recognised in the voice instruction "Deliver shock now" (Suitability for the user's tasks)
A different sequence of actions	Performing actions deviating from the intended sequence of actions: - Attaching the electrodes before switching on (AED1 only) - Pressing the shock button instead of the start button (AED1 only) - (thinking about) pressing the shock button before being prompted by the device - Searching for information about the functionality	3 (21%) 95% CI: 0%–43%	2 (14%) 95%-CI: 0%–36%	Reading safety instructions on the back of the packaging of the electrodes, loss of time of up to 21 s before switching on Fault tolerance of system: switch-on also possible via shock button, no shock delivery if not intended according to analysis Uncertainty of the user	Uncertainty of the user	Lack of user control and information on system status, need for action and sequence of action not clear: - Initial activation via button necessary to start user support - Shock is only delivered after analysis and explicit request by the unit (Self-descriptiveness) Display of nontask-relevant information (AED1) (Suitability for the user's tasks)
Asynchronous sequence of actions	Actions asynchronous to voice instructions: - Performing activities before being prompted by the device e.g. unpacking and attaching electrodes - Performing activities slower than the sequence provided by voice prompts in the device	5 (36%) 95%-CI: 14%–65%	11 (79%) 95%-CI: 57%–100%	Uncertainty of the user due to unnecessary or inappropriate instructions for action Time-saving through abbreviation	Time loss of up to 30 seconds due to a faulty attempt to open the electrode packaging Uncertainty of the user due to unnecessary/ inappropriate instructions for action Time-saving through abbreviation Loss of time due to missing/waiting for appropriate action instruction, in some cases compensation through independent action	Opening mechanism and voice prompt "Pull the red handle" after an action has already been performed is misunderstood (AED2) (Use error robustness) Timed actions, insufficient support of the user (Controllability)

Literaturverzeichnis.
ISO 9241-110: ISO 9241-110:2020-05, Ergonomics of human-system interaction - Part 110: Interaction principles.

action (Andre et al., 2004; Eames et al., 2003; Mosesso et al., 2009).

Some participants performed actions deviating from the intended process. With AED1, three participants attached the electrodes first and started the AED afterwards. On the AED2, the first required step was to open the lid; the AED does not allow for alternative actions. The alternative sequence is not considered incorrect, but it is not the intended use of the AED. This can and led to errors in perception, cognition and action. It is likely that the subjects are unaware that the voice instruction after starting the AED, explicitly refers to individual steps to assist the user in performing them. The same applies to the awareness of the need to analyze the heart rhythm and the timing of the shock delivery. Similar to our observations, Mosesso et al. (2009) found that the time to power on was shorter for devices with open lids and pull handle mechanisms than for those with a push button.

The AED2 in particular experienced interaction errors and asynchronous action sequences due to the length of and pauses between the voice prompts. The time slot for delegating the emergency call was too short and, as the patient was already undressed, there was a waiting time that made the participants feel insecure or led them to initiate the next steps before the appropriate announcement. In total, 11 people (79%) deviated from the procedure intended by the voice instructions due to the small-step instructions for positioning the electrodes. Nevertheless, it has been shown that voice instructions can improve electrode positioning (Andre et al., 2004; Mosesso et al., 2009). In our study, the voice instructions were a potential source of error: participants did not follow the prescribed sequence. If there are too many voice interactions, some of may be missed, especially in a noisy environment or under stress. In addition, a study of AEDs in hospitals has shown that TTS took in median only about half of the time of 30 s (IQR: 28–32 s) (Wurmb et al., 2015). This may be due to the trained users (nurses). However, there were differences, suggesting that interrupting the voice instructions results in a huge gain in shock delivery efficiency. It is debatable whether voice instructions help inexperienced users adhere to correct use and slow down experienced users in the process. However, voice instructions may not be audible in noisy environments. Even for non-native speakers, voice instructions can be a barrier and therefore a hazard.

In the evaluation of the SUS, the usability was rated as acceptable (J. B. Brooke, 2013), with AED1 being rated better than AED2. In the study by Polikaitis (2005), the five AEDs evaluated also scored well in the satisfaction assessment using individual questions (mean: 4.33 of 5, SD: 0.33). A significantly lower level of satisfaction was found for the AED whose median time to shock was also significantly longer (Polikaitis, 2005).

4.1. Strengths and weaknesses of the study

The strength of the comparative study is the randomized confirmatory study design in a simulator study. The study design followed the guidelines for clinical trials. As a result, a clinically relevant effect was significantly demonstrated. This shows the importance of usability concepts for patient safety. By testing the AEDs in a simulated scenario their usability could be compared and evaluated in a controlled environment. In particular, the use of eye-tracker recordings to analyse use errors provided detailed insights into the causes of the differences found. This not only allowed device-specific interaction problems to be identified but also revealed which design aspects were better or worse in comparison.

The selected scenario (shockable rhythm and end of simulation after first shock delivery) enables evaluation of the essential functionalities of the AEDs. Other scenarios, e.g., a non-shockable rhythm, were not tested.

A detailed evaluation of the correct electrode position (Andre et al., 2004; Eames et al., 2003; Polikaitis, 2005) was not performed. Firstly, the manikin was already undressed by the second actor at the beginning. Secondly, the positioning of the pads was distorted by the ECG connections on the manikin's torso.

For this study, the two AEDs with the largest market share (in Germany) were examined exemplarily. The training devices are not identical to devices intended for real-life use, e.g., with regard to labelling and electrode packaging. The problems identified are therefore only partially transferable to other devices.

ILCOR recommends a short period of CPR until the defibrillator is ready for analysis or defibrillation in unmonitored cardiac arrest (weak recommendation, low-certainty evidence) (Nolan et al., 2020; Wyckoff et al., 2022). This was accounted for in this study by a two-helper resuscitation operation. Despite standardized procedures, the influence of human–human communication on the results cannot be completely excluded.

In a real emergency, stress, characteristics of the first responder such as language skills or hearing impairment, as well as contextual factors such as ambient noise, communication with other actors, the patient's clothing or bad weather also influence efficiency and use errors. We have omitted these external factors from the simulation to analyze solely the use of the AED. Since the participants primarily come from the age group 26–35 with a higher level of education, the population may not be representative of all potential AED users.

4.2. Conclusion

Testing the two AEDs in a simulated scenario shows a clinically relevant difference in TTS. The duration of use was different due to the interaction design of the AEDs. This includes differences in the use of voice instruction (length, pauses, and terminology), iconography on the device and electrodes, and device opening mechanisms. This resulted in lost time, use errors, and participant uncertainty during the interaction.

No general conclusions can yet be drawn from this study regarding an interface design that works best for AEDs, but it provides a basis for further research. An investigation of further devices currently available on the market in this study design to identify differences in the performance of the devices and further user interface issues is a logical

consequence of these study results. The next step is to (re-)start research to improve the interfaces and develop a catalog of criteria and best practices for AED interfaces. Manufacturers can draw on and factor the revealed problems into their risk management. The use concept for an AED should ensure certain freedom of control and abbreviation of processes for faster operation. If voice instructions are used, short instructions are preferable to long ones. Training should not be the sole solution for effective, efficient, and safe use. Instead, it should serve the purpose of communicating the relevance of AEDs and removing any fears about using them. For future development, there should be benchmark tests for the minimum time of use, which should serve as a limit value for development. The time of use must no longer be a game of luck depending on the AED that is used.

Acknowledgments

The authors would like to thank Stefan Thate and Eike Mitschler (Oldenburg Fire Department) for their assistance in recruiting volunteers.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Author contributions

ML, AK, and RR conceived the original idea. BC, AK, ML, and RR have developed the study design and performed the data acquisition, analysis and interpretation. FO supported the study design, sample size estimation and data analysis as a biometrician. BC wrote the draft of the manuscript. AK, FO, RR, and ML contributed substantially to its revision.

ORCID

Beatrice Coldewey http://orcid.org/0000-0001-9564-9467
Andreas Klausen http://orcid.org/0009-0003-6203-7386
Fabian Otto-Sobotka http://orcid.org/0000-0002-9874-1311
Rainer Röhrig http://orcid.org/0000-0002-0032-5118
Myriam Lipprandt http://orcid.org/0000-0001-9371-0551

Data availability statement

The data that support the findings of this study and the R-script are available via zenodo. The video recordings will not be provided due to privacy protection.

References

Andre, A. D., Jorgenson, D. B., Froman, J. A., Snyder, D. E., & Poole, J. E. (2004). Automated external defibrillator use by untrained bystanders: Can the public-use model work? *Prehospital Emergency Care*, *8*(3), 284–291. https://doi.org/10.1016/j.prehos.2004.02.004

Bækgaard, J. S., Viereck, S., Møller, T. P., Ersbøll, A. K., Lippert, F., & Folke, F. (2017). The effects of public access defibrillation on survival after out-of-hospital cardiac arrest. *Circulation*, *136*(10), 954–965. https://doi.org/10.1161/CIRCULATIONAHA.117.029067

Bødtker, H., & Rosendahl, D. (2018). Automated external defibrillators and defibrillation electrodes from major manufactures depict placement of the left apical defibrillation electrode poorly!. *Resuscitation*, *125*, e11–e12. https://doi.org/10.1016/j.resuscitation.2018.01.046

Brooke, J. (1996). SUS - A quick and dirty usability scale. In *Usability Evaluation in Industry*. https://doi.org/10.1201/9781498710411-35

Brooke, J. B. (2013). SUS: A retrospective. *Journal of Usability Studies*, *8*(2), 29–40. https://dl.acm.org/doi/10.5555/2817912.2817913

Brooks, S. C., Clegg, G. R., Bray, J., Deakin, C. D., Perkins, G. D., Ringh, M., Smith, C. M., Link, M. S., Merchant, R. M., Pezo-Morales, J., Parr, M., Morrison, L. J., Wang, T. -L., Koster, R. W., & Ong, M. E. (2022). Optimizing outcomes after out-of-hospital cardiac arrest with innovative approaches to public-access defibrillation: A scientific statement from the International Liaison Committee on Resuscitation. *Resuscitation*, *172*, 204–228. https://doi.org/10.1016/j.resuscitation.2021.11.032

Callejas, S., Barry, A., Demertsidis, E., Jorgenson, D., & Becker, L. B. (2004). Human factors impact successful lay person automated external defibrillator use during simulated cardiac arrest. *Critical Care Medicine*, *32*(9 Suppl), S406–S413. https://doi.org/10.1097/01.ccm.0000139948.46663.3a

Dainty, K. N., Colquitt, B., Bhanji, F., Hunt, E. A., Jefkins, T., Leary, M., Ornato, J. P., Swor, R. A., & Panchal, A. (2022). Understanding the importance of the lay responder experience in out-of-hospital cardiac arrest: A scientific statement From the American Heart Association. *Circulation*, *145*(17), e852–e867. https://doi.org/10.1161/CIR.0000000000001054

Dong, X-J., Zhang, L., Yu, Y-L., Shi, S-X., Yang, X-C., Zhang, X-Q., Tian, S., Myklebust, H., Li, G-H., & Zheng, Z-J (2020). The general public's ability to operate automated external defibrillator: A controlled simulation study (No. 4). *World Journal of Emergency Medicine*, *11*(4), 238–245. https://doi.org/10.5847/wjem.j.1920-8642.2020.04.006

Eames, P., Larsen, P. D., & Galletly, D. C. (2003). Comparison of ease of use of three automated external defibrillators by untrained lay people. *Resuscitation*, *58*(1), 25–30. https://doi.org/10.1016/S0300-9572(03)00103-5

Fairbanks, R. J., & Caplan, S. (2004). Poor interface design and lack of usability testing facilitate medical error. *Joint Commission Journal on Quality and Safety*, *30*(10), 579–584. https://doi.org/10.1016/S1549-3741(04)30068-7

Fleischhackl, R., Losert, H., Haugk, M., Eisenburger, P., Sterz, F., Laggner, A. N., & Herkner, H. (2004). Differing operational outcomes with six commercially available automated external defibrillators. *Resuscitation*, *62*(2), 167–174. https://doi.org/10.1016/j.resuscitation.2004.03.018

Food and Drug Administration (2016). *Applying human factors and usability engineering to medical devices: Guidance for industry and food and drug administration staff*. FDA.

Foster, A. G., & Deakin, C. D. (2019). Accuracy of instructional diagrams for automated external defibrillator pad positioning. *Resuscitation*, *139*, 282–288. https://doi.org/10.1016/j.resuscitation.2019.04.034

Geis, T., & Johner, C. (2020). *Usability Engineering als Erfolgsfaktor: Effizient DIN EN 62366-1- und FDA-konform dokumentieren*. Beuth Verlag GmbH (Beuth Praxis).

Gold, L. S., Fahrenbruch, C. E., Rea, T. D., & Eisenberg, M. S. (2010). The relationship between time to arrival of emergency medical services (EMS) and survival from out-of-hospital ventricular fibrillation cardiac arrest. *Resuscitation*, *81*(5), 622–625. https://doi.org/10.1016/j.resuscitation.2010.02.004

Gräsner, J., Herlitz, T., Tjelmeland, J., Wnent, I. B., Masterson, J., Siobhan, S., Bein, G., Böttiger, B., Rosell-Ortiz, B. W., Nolan, F., Bossaert, J. P., Perkins, L., & G. D., Lilja. (2021). European Resuscitation Council Guidelines 2021: Epidemiology of cardiac arrest in Europe. *Resuscitation*, *161*, 61–79. https://doi.org/10.1016/j.resuscitation.2021.02.007

Gräsner, J.-T., Wnent, J., Herlitz, J., Perkins, G. D., Lefering, R., Tjelmeland, I., Koster, R. W., Masterson, S., Rossell-Ortiz, F., Maurer, H., Böttiger, B. W., Moertl, M., Mols, P., Alihodžić, H., Hadžibegović, I., Ioannides, M., Truhlář, A., Wissenberg, M., Salo, A., … Bossaert, L. (2020). Survival after out-of-hospital cardiac arrest in Europe - Results of the EuReCa TWO study. *Resuscitation*, *148*, 218–226. https://doi.org/10.1016/j.resuscitation.2019.12.042

Handley, A., Bahr, J., Baskett, P., Bossaert, L., Chamberlain, D., Dick, W., Ekström, L., Juchems, R., Kettler, D., Marsden, A., Moeschler,

O., Monsieurs, K., Parr, M., Petit, P., & van Drenth, A. (1998). The 1998 European Resuscitation Council guidelines for adult single rescuer basic life support: A statement from the Working Group on Basic Life Support, and approved by the executive committee of the European Resuscitation Council. *Resuscitation*, *37*(2), 67–80. https://doi.org/10.1016/S0300-9572(98)00036-7

Hawkes, C., Booth, S., Ji, C., Brace-McDonnell, S. J., Whittington, A., Mapstone, J., Cooke, M. W., Deakin, C. D., Gale, C. P., Fothergill, R., Nolan, J. P., Rees, N., Soar, J., Siriwardena, A. N., Brown, T. P., & Perkins, G. D, OHCAO collaborators (2017). Epidemiology and outcomes from out-of-hospital cardiac arrests in England. *Resuscitation*, *110*, 133–140. https://doi.org/10.1016/j.resuscitation.2016.10.030

Ho, C. L., Cheng, K. W., Ma, T. H., Wong, Y. H., Cheng, K. L., & Kam, C. W. (2016). Characterization of available automated external defibrillators in the market based on the product manuals in 2014. *World Journal of Emergency Medicine*, *7*(2), 138. https://doi.org/10.5847/wjem.j.1920-8642.2016.02.010

IEC 62366-1. (2015). IEC 62366-1:2015/Amd 1:2020, Medical devices - Part 1: Application of usability engineering to medical devices.

ISO 14971. (2019). ISO 14971:2019-12, Medical devices—Application of risk management to medical devices.

Kiguchi, T., Okubo, M., Nishiyama, C., Maconochie, I., Ong, M. E. H., Kern, K. B., Wyckoff, M. H., McNally, B., Christensen, E. F., Tjelmeland, I., Herlitz, J., Perkins, G. D., Booth, S., Finn, J., Shahidah, N., Shin, S. D., Bobrow, B. J., Morrison, L. J., Salo, A., … Iwami, T. (2020). Out-of-hospital cardiac arrest across the World: First report from the International Liaison Committee on Resuscitation (ILCOR). *Resuscitation*, *152*, 39–49. https://doi.org/10.1016/j.resuscitation.2020.02.044

Larsen, M. P., Eisenberg, M. S., Cummins, R. O., & Hallstrom, A. P. (1993). Predicting survival from out-of-hospital cardiac arrest: A graphic model. *Annals of Emergency Medicine*, *22*(11), 1652–1658. https://doi.org/10.1016/S0196-0644(05)81302-2

Lei, C., Qin, H., Dong, X., You, J., & Zhang, L. (2022). Layperson's performance on an unconversant type of AED device: A prospective crossover simulation experimental study (No. 2). *World Journal of Emergency Medicine*, *13*(2), 98–105. https://doi.org/10.5847/wjem.j.1920-8642.2022.024

Mosesso, V. N., Shapiro, A. H., Stein, K., Burkett, K., & Wang, H. (2009). Effects of AED device features on performance by untrained laypersons. *Resuscitation*, *80*(11), 1285–1289. https://doi.org/10.1016/j.resuscitation.2009.07.016

Müller, M. P., Poenicke, C., Kurth, M., Richter, T., Koch, T., Eisold, C., Pfältzer, A., & Heller, A. R. (2015). Quality of basic life support when using different commercially available public access defibrillators. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, *23*(1), 48. https://doi.org/10.1186/s13049-015-0123-1

Nolan, J. P., Maconochie, I., Soar, J., Olasveengen, T. M., Greif, R., Wyckoff, M. H., Singletary, E. M., Aickin, R., Berg, K. M., Mancini, M. E., Bhanji, F., Wyllie, J., Zideman, D., Neumar, R. W., Perkins, G. D., Castrén, M., Morley, P. T., Montgomery, W. H., Nadkarni, V. M., … Hazinski, M. F. (2020). Executive Summary 2020 International Consensus on Cardiopulmonary Resuscitation and Emergency Cardiovascular Care Science with treatment recommendations. *Resuscitation*, *156*, A1–A22. https://doi.org/10.1016/j.resuscitation.2020.09.009

Nurmi, J., & Castrén, M. (2005). Layperson positioning of defibrillation electrodes guided by pictorial instructions. *Resuscitation*, *64*(2), 177–180. https://doi.org/10.1016/j.resuscitation.2004.08.014

Olasveengen, T. M., Semeraro, F., Ristagno, G., Castren, M., Handley, A., Kuzovlev, A., Monsieurs, K. G., Raffay, V., Smyth, M., Soar, J., Svavarsdottir, H., & Perkins, G. D. (2021). European Resuscitation Council Guidelines 2021: Basic life support. *Resuscitation*, *161*, 98–114. https://doi.org/10.1016/j.resuscitation.2021.02.009

Polikaitis, A. (2005). *The usability of five automated external defibrillators by minimally trained bystanders*. University of Illinois Medical Center, Chicago, Illinois. https://www.semanticscholar.org/paper/The-Usability-of-Five-Automated-External-by-Trained-Polikaitis/3b418ed1c63b8627a9b4dd47816a6b64f1bce4f3

Schehadat, M. S., Groneberg, D. A., Bauer, J., & Bendels, M. H. K. (2017). Hilfsfristen des Rettungsdienstes in den deutschen Bundesländern. *Zentralblatt für Arbeitsmedizin, Arbeitsschutz Und Ergonomie*, *67*(5), 255–260. https://doi.org/10.1007/s40664-017-0203-3

Torney, H., Harvey, A., Finlay, D., Magee, J., Funston, R., & Bond, R. (2018). *Eye-tracking analysis to compute the visual hierarchy of user interfaces on automated external defibrillators* [Paper presentation]. https://doi.org/10.14236/ewic/HCI2018.42

Valenzuela, T. D., Roe, D. J., Cretin, S., Spaite, D. W., & Larsen, M. P. (1997). Estimating effectiveness of cardiac arrest interventions: A logistic regression survival model. *Circulation*, *96*(10), 3308–3313. https://doi.org/10.1161/01.CIR.96.10.3308

Weaver, W. D., Cobb, L. A., Hallstrom, A. P., Fahrenbruch, C., Copass, M. K., & Ray, R. (1986). Factors influencing survival after out-of-hospital cardiac arrest. *Journal of the American College of Cardiology*, *7*(4), 752–757. https://doi.org/10.1016/S0735-1097(86)80332-1

Wiklund, M., Dwyer, A., & Davis, E. (2016). *Medical device use error: Root cause analysis* (1st ed.). CRC Press. https://doi.org/10.1201/b19265

Wurmb, T., Vollmer, T., Sefrin, P., Kraus, M., Happel, O., Wunder, C., Steinisch, A., Roewer, N., & Maier, S. (2015). Monitoring of in-hospital cardiac arrest events with the focus on automated external defibrillators–A retrospective observational study. *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine*, *23*(1), 87. https://doi.org/10.1186/s13049-015-0170-7

Wyckoff, M. H., Greif, R., Morley, P. T., Ng, K.-C., Olasveengen, T. M., Singletary, E. M., Soar, J., Cheng, A., Drennan, I. R., Liley, H. G., Scholefield, B. R., Smyth, M. A., Welsford, M., Zideman, D. A., Acworth, J., Aickin, R., Andersen, L. W., Atkins, D., Berry, D. C., … Berg, K. M, Collaborators (2022). 2022 International Consensus on Cardiopulmonary Resuscitation and Emergency Cardiovascular Care Science with treatment recommendations: Summary from the basic life support; advanced life support; pediatric life support; neonatal life support; education, implementation, and teams; and first aid task forces. *Circulation*, *146*(25), e483–e557. https://doi.org/10.1161/CIR.0000000000001095

About the authors

Beatrice Coldewey is a PhD-Candidate in Medical Informatics at the RWTH Aachen University. She studied Computer Science at the Carl von Ossietzky University Oldenburg and Assistive Technologies at the Jade University of Applied Sciences. Her research is focused on usability and risk-management in medical socio-technical systems.

Andreas Klausen is MD and Post-Doctoral-Researcher in the department of Big Data in Medicine at the Carl von Ossietzky University Oldenburg and Emergency Physician in the Emergency Department of the Oldenburg Hospital. He has experience in both prehospital emergency care and clinical acute and emergency medicine.

Fabian Otto-Sobotka is PhD in medical Biostatistics and Post Doctoral researcher in the devision of epidemiology and biostatistics of the Carl von Ossietzky University Oldenburg.

Rainer Röhrig is Fullprofessor for Medical Informatics and Director of the Institute of Medical Informatics at the RWTH Aachen University. The main research areas are research data infrastructure and medical socio-technical systems. He worked for many years as a physician and researcher in Anaesthesiology, Intensive Care and Emergency Medicine at the University Hospital Giessen and Marburg - Campus Giessen.

Myriam Lipprandt is Post-Doctoral / Senior-Researcher in Computer Science and deputy director of the Institute of Medical Informatics at the RWTH Aachen University. Her focus is on medical software engineering and human-machine interaction in medicine. She leads a simulation and usability laboratory for medical software and digital medical devices.