university of groningen

faculty of science and engineering

Development of an Interactive Video Practice Application for Rare Procedures in Emergency Medical Services

Floor Kuipers

university of groningen / faculty of science and engineering

University of Groningen

Development of an Interactive Video Practice Application for Rare Procedures in Emergency Medical Services

Master's Thesis

To fulfill the requirements for the degree of
Master of Science in Computational Cognitive Science
at the University of Groningen under the supervision of
Dr. F. Cnossen (Artificial Intelligence, University of Groningen)
and
B. Dercksen (UMCG, Groningen)

Floor Kuipers (s4213335)

December 4, 2022

Acknowledgments

I would like to express my gratitude to my supervisor, Dr. Fokie Cnossen, who initiated this project. Her expertise and wide research interests always led to interesting dialogue and discussions.

I also would like to thank my external supervisor, Bert Dercksen, for always finding a way to make time for me, even though he faces a lot of different pressing matters each day. Both Fokie and Bert provided an approachable atmosphere, in which issues could easily be discussed and worked through.

In addition, I would like to thank the training coordinators of the ambulance posts in Leeuwarden and Emmen, Peter de Ruiter and Wilma Hulsman, for facilitating my experiments.

Finally, I would like to give thanks to my partner, Casper Maas, for always being open for a discussion and brainstorm and for his support.

Abstract

Nurses in the Emergency Medical Services face a broad variety of medical situations. Most of these medical situations require frequently used skills, but other skills are rarely used. In both situations, EMS teams have to be able to take the appropriate steps. It is known that if a skill requires declarative knowledge to practice it, it is prone to decay over time if it is not used. It can be very impactful and distressing if a rare procedure has to be performed and the EMS team does not know how to take the appropriate steps anymore. This research focuses on a solution to this issue of skill decay of rarely executed procedures. This is done by developing an interactive video application, named '*IVEA*', with which nurses can practice the necessary knowledge of such a rare skill. This research focused specifically on the cricothyroidotomy procedure. Nurses used *IVEA* for two weeks, after which they performed a cricothyroidotomy procedure on a phantom. Their performance of this procedure was compared to the performance of nurses who did not use the application. No firm conclusion could be drawn as to whether or not performance of the nurses improved, because of a low number of participants. It was, however, apparent that nurses would appreciate and use an app such as *IVEA* as a supporting learning tool. There were clear indications that there is a need for a more coherent application for refresher education. Overall, *IVEA* shows promise in supporting nurses in practicing skills that rarely have to be used, in an interactive and accessible way.

Contents

1 Introduction **7**

2 Theoretical Background **9**
2.1 Skill acquisition . . . 9
2.1.1 Declarative and procedural knowledge . . . 9
2.1.2 ACT-R . . . 9
2.1.3 Stages of skill acquisition . . . 13
2.2 Skill retention . . . 15
2.2.1 Emergency Medical Services . . . 15
2.2.2 Spacing . . . 16
2.2.3 Refresher interventions . . . 18
2.3 Mobile medical education . . . 19
2.3.1 Media in education . . . 19

3 Methods **21**
3.1 App Development . . . 21
3.1.1 Cognitive Task Analysis . . . 21
3.1.2 Videos . . . 23
3.1.3 App development . . . 23
3.2 Experiment . . . 37
3.2.1 Participants . . . 37
3.2.2 Materials . . . 38
3.2.3 Procedure . . . 38

4 Results **43**
4.1 Performance measure . . . 43
4.1.1 Evaluation by medical experts . . . 43
4.2 Usability . . . 43
4.2.1 App usage . . . 43
4.2.2 Observation of videofootage . . . 43
4.2.3 Quantitative usability . . . 44
4.2.4 Qualitative usability . . . 44

5 Discussion **46**
5.1 Did *IVEA* improve performance? . . . 46
5.1.1 Age and experience as confounding factors . . . 46
5.2 Is *IVEA* useful? . . . 47
5.3 Limitations . . . 48
5.3.1 Development . . . 48
5.3.2 Experiment . . . 48
5.4 Recommendations . . . 49
5.4.1 Feedback in the app . . . 49
5.4.2 Content of the app . . . 49
5.4.3 Future research . . . 49

6 Conclusion **51**

References **52**

A Questions of interactive videos **56**

B Walkthrough of Figma designs **57**

C MoSCoW table **59**

D Source code **60**

E Information for participants **60**
E.1 General information e-mail (in Dutch) 60
E.2 Extensive experiment information (in Dutch) 60

F Usability questionnaire **62**

G Synthesis of expert evaluations **63**

1 Introduction

Nurses in the Emergency Medical Services (EMS) face a broad variety of medical situations. From frequently performed procedures, such as peripheral venous access or cardiac monitoring (Carlson et al., 2015), to rare procedures such as the cricothyroidotomy (Petrosoniak, Ryzynski, Lebovic, & Woolfrey, 2017), EMS teams have to be able to take the appropriate steps. Generally, there is little to no time to consult sources when driving or flying to a situation that requires emergency care. It is therefore of the utmost importance that the knowledge and skills are readily available. Benneck and Bremer (2019) researched situations that almost went wrong, which they call ‘near misses’, in ambulance care. They describe that near misses that are caused by, for instance, lacking knowledge, leads to stress, guilt and shame in nurses working on that ambulance. It can be very impactful if the EMS team lacks the necessary knowledge for a rare procedure, both for the patient and for the nurses. Since cricothyroidotomy procedures are considered a last resort, it is argued that the maintenance of this skill is an important issue (Chang, Hamilton, & Carter, 1998). Chang et al. (1998) mention that whereas surgical residents will probably acquire proficiency in this procedure in the operating room, EMS nurses may never see or perform the procedure in a clinical setting during their residency training. It is thus likely that the necessary knowledge of the cricothyroidotomy procedure decays over time. Skill maintenance of this specific procedure is therefore necessary.
This skill maintenance should preferably not be happening while driving (or flying) to a situation that requires emergency medical care, since this context does not allow for much preparation. The refreshing of this knowledge could be done when nurses are at home, when they are not paged to an emergency or when they have a break. Because the knowledge of these rare skills, such as the cricothyroidotomy, is vital to remember on the spot when with a patient, the skill is re-evaluated periodically. When these evaluation moments are scheduled, it is usually the case that nurses practice the skill through textbooks or several different internet sources in the weeks leading up to the evaluation. This method of preparation is, however, not very engaging or consistent. A more immersive method of learning could be created using interactive technologies. Interactive technologies are quickly becoming key instruments in displaying a real environment that can be used for education in healthcare (Herault, Lincke, Milrad, Forsgärde, & Elmqvist, 2018). The interactivity of these technologies has been shown to positively influence user attitude towards learning and enhance learners’ performance (Kettanurak, Ramamurthy, & Haseman, 2001). Usually, these technologies are aimed at learning entirely new skills. It is less extensively researched to what extent these interactive technologies are applicable for autonomous knowledge retention and skill maintenance after a skill has been learned already.

In collaboration with the Emergency Medical Services of the University Medical Center Groningen (UMCG), the following research question thus arose from the issues described above:

> *"To what extent does an interactive video practice application support EMS nurses in refreshing declarative knowledge of rare EMS procedures, in particular the cricothyroidotomy?"*

This research question can be split up into two sub-questions:

1. *"To what extent does the use of an interactive video application improve the performance of EMS nurses in the cricothyroidotomy skill?"*
2. *"To what extent is an interactive video application useful in supporting the learning process of nurses"*

This research aims to answer these questions by developing an interactive video application for EMS nurses to practice the cricothyroidotomy skill. This is done by refreshing the necessary declarative knowledge of the skill through the application. The hypothesis for the first sub-question is that if the declarative knowledge is strengthened, the performance of nurses should increase. An experiment is therefore conducted in which the performance of nurses who have used the developed application for two weeks is compared to the performance of nurses who have prepared only with their regular study material.

The hypothesis for the second sub-question is that nurses would find such an interactive video application useful as a supporting educational tool.

2 Theoretical Background

2.1 Skill acquisition

2.1.1 Declarative and procedural knowledge

In the topics of learning and memory, a distinction is often made between procedural knowledge (knowing how to do a task) and declarative knowledge (factual knowledge, knowing the ‘what’). Both types of knowledge support skill acquisition (J. R. Anderson, 1982; J. R. Anderson et al., 2004, as cited in Fernandes, 2018). Procedural knowledge is more implicit, whereas declarative knowledge is more explicit. The person who is practicing the skill will likely find it hard to explain the ‘how’ of the task (i.e. how they know where to click to type the letter ’j’ on a keyboard), but find it easier to explain the ‘what’ (i.e. that the ’#’ is located above to the ’3’ on the keyboard). Another distinction between the two knowledge types is that declarative knowledge can decay over time, when the knowledge is not used enough. Procedural knowledge, however, does not decay (Cnossen, 2015). The necessary knowledge for a task can be declarative, procedural or a combination of both (J. W. Kim, Ritter, & Koubek, 2013).

2.1.2 ACT-R

J. R. Anderson (2007) created a theory called ACT-R (Adaptive Control of Thought–Rational), for simulating and understanding human cognition. This theory is more than just a theory, however. It is referred to as an ’architecture of cognition’. This cognitive architecture provides a further, in depth, understanding of the underlying mechanisms of the two knowledge types, declarative and procedural. One of the three essential theoretical commitments that are made in the ACT-R knowledge representation is that ACT-R has two long-term repositories of knowledge, declarative knowledge and procedural knowledge (J. R. Anderson, 2014). The second and third commitment are that the basic unit of knowledge in declarative knowledge is referred to as a ’*chunk*’ and the basic unit of knowledge in procedural knowledge is a ’*production rule*’ (J. R. Anderson, 2014). Both knowledge types have two layers of representation, namely a symbolic level and a sub-symbolic level (Taatgen, 1999). The symbolic level is a high-level, human-readable, representation, where the sub-symbolic level is a representation with a lower explainability (no longer human-readable).

Declarative knowledge

As explained in the previous section, declarative knowledge represents the factual knowledge. Declarative memory contains facts and goals, which are represented by a semantic network at a symbolic level (Taatgen, 1999). Declarative memory is a knowledge base that describes the semantic relationships between different declarative knowledge chunks. A chunk consists of a type and slots that each contain a value. The type of a chunk can be viewed as the category (e.g. dogs) and the slots as the attributes of that category (e.g. their breed).

At a sub-symbolic level, the representation looks more like neural networks (Taatgen, 1999). Each chunk has their own so-called ’activation’, which represents how available a chunk is at a given time (Van Rijn, van Maanen, & van Woudenberg, 2009). The activation of a chunk thus determines whether or not it can be retrieved from memory at a certain moment (Taatgen, 1999). When a chunk is used, its activation increases. Over time, when it is not used, this activation decreases again. This process is called decay. The activation of chunks in ACT-R is described by the following Equation (J. R. Anderson and Schooler, 1991, as cited in Van Rijn et al., 2009):

$$A_i(t) = \sum_{j=1}^{n} (t - t_j)^{-d_j} \tag{1}$$

Equation 1: Activation Equation

In this equation, the activation, A, of a certain chunk, i, at a certain time, t, is described. The decay parameter, d_j, represents the speed of decline. In ACT-R, this decay parameter is almost always set to 0.5 (ACT-R, 2022b). If also the current context of a certain chunk should be taken into account, Equation 2 is suitable to work with.

$$A_i = B_i + C_i + M_i + \varepsilon \tag{2}$$

Equation 2: Activation Equation with context taken into account

The B_i represents the base-level activation, previously described in Equation 1 as A_i. The C_i represents the influence of the current context on the activation, A_i, which is called the 'spreading activation'. This spreading activation entails that there is an increased probability of needing a chunk if it is associated with the current context (Borst, 2019b). This implies that if a learner, for instance, activates a chunk with information about a specific dogbreed (e.g. 'Labrador'), another chunk with a different dogbreed also slightly increases in activation. The ε represents the noise of a chunk. The base-level activation, B_i, of a chunk can also be viewed as the odds that you need a chunk based on frequency and recency (Borst, 2019a). Often the log is used of the base-level activation, resulting in Equation 3.

$$B_i(t) = ln\left[(\sum_{j=1}^{n} (t - t_j)^{-d_j} \right] \tag{3}$$

Equation 3: Base-level activation

Since Equation 3 implies that different presentations of a chunk are additive (implied by the summation sign), the activation over time follows the pattern described in Figure 1.

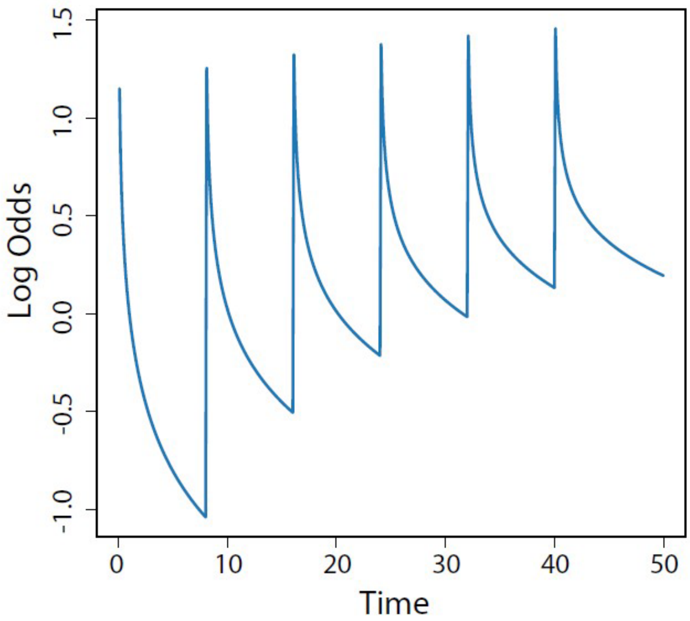


Figure 1: Base-level activation pattern with multiple repetitions, by (Borst, 2019a)

With time, the ’odds’, or the activation, of a chunk decreases when it is not used. With a new presentation of the chunk, the activation increases again.

Sometimes, a memory chunk is not used for such a long time, that it can not be remembered at all anymore. This is what happens when the activation of a chunk drops below the retrieval threshold. This retrieval threshold can be explained by looking at the recall probability of a chunk, i (Equation 4).

$$\text{Recall probability}_i = \frac{1}{1 + e^{\frac{\tau - A_i}{s}}} \tag{4}$$

Equation 4: Recall probability of chunks, by (Borst, 2019a)

The recall probability is described by the retrieval threshold, τ, the activation, A_i, and the noise variance, s. Figure 2 displays what the curve of the recall probability looks like for different noise variance values. In this plot, the retrieval threshold is displayed by a red, vertical, line. In ACT-R, the retrieval threshold is set to 0.5 (ACT-R, 2022b).

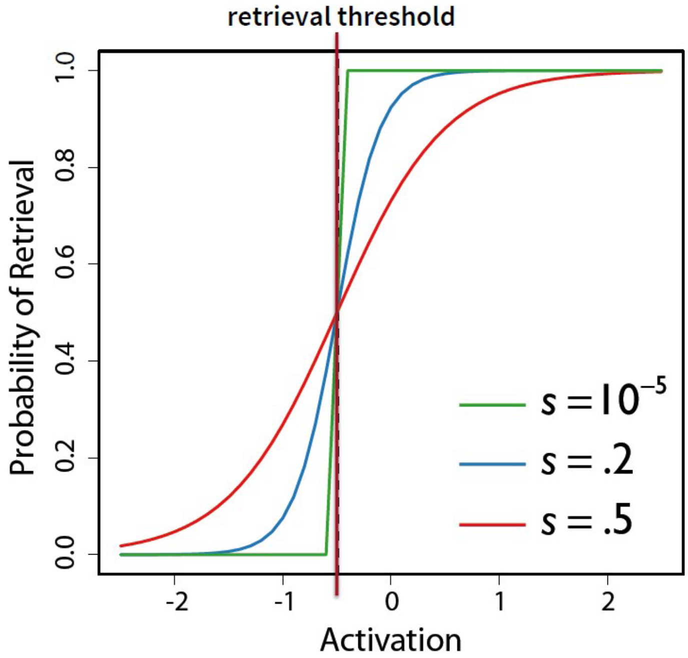


Figure 2: Retrieval threshold of chunks, by (Borst, 2019a)

Procedural knowledge

Procedural knowledge is made up of production rules. Each of these production rules can be seen as a modular piece of knowledge, because it represents a well-defined step of cognition (J. R. Anderson, 2014). Production rules are encoded as so-called 'IF → THEN', or 'Condition → Action' statements. An example of such an IF → THEN statement can be seen in Figure 3.

```
IF the goal is to classify a person
     and he is unmarried
THEN classify him as a bachelor
```

Figure 3: Example of IF → THEN statement (ACT-R, 2022a)

Procedural memory is thus a collection of actional forms of information. When a production rule is used, this is called 'firing'. Only one production rule can fire at a time. To determine which production rule should be fired, the buffers of ACT-R are used. These buffers function as the components between the procedural memory system and other components, referred to as modules, of the ACT-R architecture, such as the declarative memory module. The declarative module and procedural module are connected through the goal and retrieval buffers. One chunk at a time can be held by each buffer. The buffers' contents are matched with the content of production rules and if there is a correct match, that production is selected (ACT-R, 2022a).

Proceduralization

The cognitive intensity of declarative knowledge retrieval is represented in ACT-R by only allowing one item to be retrieved at a certain time (J. W. Kim et al., 2013). This is a time-intensive process, which becomes less time-intensive over time due to a concept that is referred to as 'proceduralization' (J. R. Anderson, 1982), or as 'knowledge compilation' (Neves and Anderson, 1981; J. R. Anderson and Lebiere, 2014; Jones, Ritter, and Wood, 2000 as cited in J. W. Kim et al., 2013). When declarative knowledge has an activation that is high enough, proceduralization occurs. After the same task is performed many times, new production rules can be created that contain the declarative knowledge values 'hard-coded' into them (Whitehill, 2013). A compiled rule is generated by ACT-R through elimination of the declarative knowledge retrieval request in the first rule, and of the retrieval condition in the second rule (J. W. Kim et al., 2013). Proceduralization allows for a faster process, by generating procedural knowledge that is task specific, where previously a string of declarative retrievals had to be used (J. W. Kim et al., 2013). The process of proceduralization plays an important role in the different stages of skill acquisition, as described in the following section.

2.1.3 Stages of skill acquisition

Learning a skill can be separated into different learning phases. Over the years, several theories have been proposed that explain these different phases. J. W. Kim et al. (2013) synthesized a theory of learning from earlier theories of Fitts (1964), J. R. Anderson (1982), Rasmussen (1987) and VanLehn (1996). The theory encompasses a three-stage process of learning, which is described in Figure 4. The theories that J. W. Kim et al. (2013) based their theory on also posited three learning phases. Fitts (1964) proposed the phases *'Cognitive', 'Associative'* and *'Autonomous'*. J. R. Anderson (1982) then based his theory on the theory by Fitts (1964), and proposed the phases *'Declarative', 'Transitional'* and *'Procedural'*. The framework that was created by Rasmussen (1987), splits up the phases in *'Knowledge-based', 'Rule-based'* and *'Skill-based'*. The latest model that J. W. Kim et al. (2013) took into account, by VanLehn (1996), described the three phases as *'Early', 'Intermediate'* and *'Late'*. J. W. Kim et al. (2013) chose to describe the different phases with terms that represent what underlying mechanisms play the largest role in each phase: *'Declarative', 'Declarative + Procedural* and *'Procedural'*.

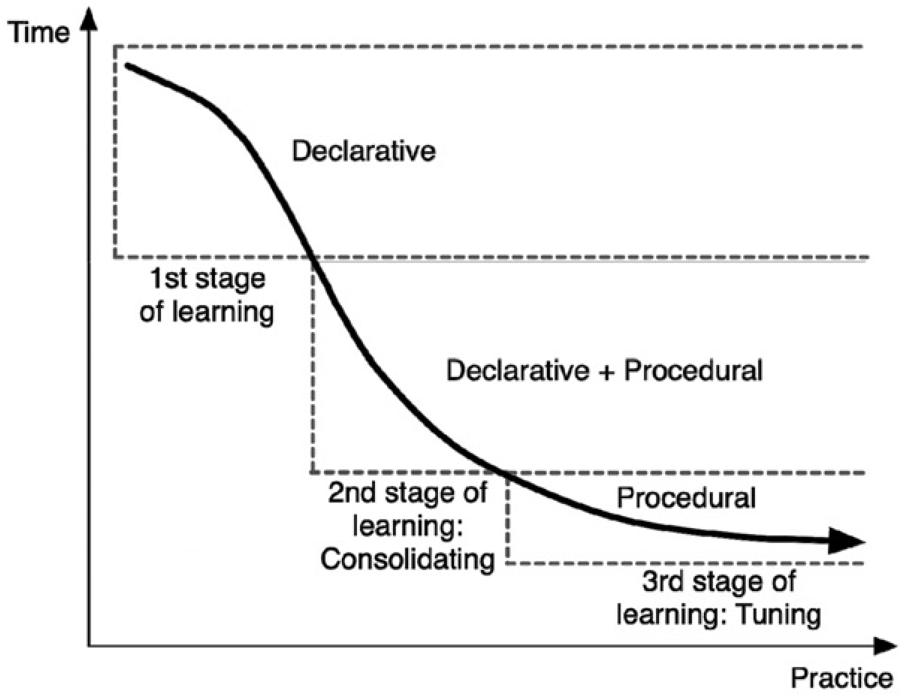


Figure 4: A plot describing the three different stages of learning by J. W. Kim et al. (2013), with the amount of practice plotted against the task completion time.

In the first, declarative, stage, the learner is mainly focused on acquiring the declarative knowledge of a skill. In this stage, the declarative knowledge decays when it is not used, leading to a potential inability to perform the task. A decreased memory strength leads to increases in response times and decreases in accuracy (J. W. Kim et al., 2013).

In the second, declarative + procedural, stage, the task knowledge consists of a mixture of declarative and procedural knowledge. Still, the declarative knowledge of the skill decays over time, but the procedural knowledge remains strong. Inability to perform the task can still occur when necessary declarative knowledge is not readily available.

In the third and final stage, procedural knowledge is primarily responsible for the performance of a skill. As indicated by the horizontal axis of Figure 4, with more practice, a learner moves from one phase to the next. As the learner progresses into the final stage, the declarative knowledge about tasks will transform into procedural knowledge. This happens because of the previously described process of proceduralization. With lack of use, the declarative knowledge will still degrade. However, J. W. Kim et al. (2013) state that the task can still be performed if all the knowledge is proceduralized, if the information is available in the environment, or if the task does not require declarative inputs. In a fully proceduralized skill, the declarative knowledge would no longer be necessary and would decay with time (Ericsson, Krampe, and Tesch-Römer, 1993, as cited by J. W. Kim et al., 2013). This is also substantiated by the cognitive architecture ACT-R and how it models the behaviour of declarative and procedural knowledge. Infrequently used aspects of a skill (such as the character '#' in the typing skill), however, can still require declarative retrieval.

2.2 Skill retention

This section describes the current situation in the emergency medical services in terms of skill retention, the concept of spacing and how it relates to skill retention and refresher interventions that support skill retention.

2.2.1 Emergency Medical Services

Cricothyroidotomy
The cricothyroidotomy procedure can be split up into two methods: the elective cricothyroidotomy and the emergency cricothyroidotomy (Walls, 1988). The elective cricothyroidotomy is usually performed in an operating room or intensive care unit. In this procedure, an incision is made in the cricothyroid membrane, through which a tube is placed. The emergency cricothyroidotomy procedure is usually performed in the emergency medical services and is performed with a special needle-kit, instead of through an incision with a scalpel. The emergency cricothyroidotomy generally has higher complication rates than the elective cricothyroidotomy (Walls, 1988). It is, however, also a life-or-death procedure, so performing the cricothyroidotomy is usually a better alternative to failing to secure the airway.

In the following patient circumstances the cricothyroidotomy is performed as emergency airway control:

- Patient with blunt trauma patient where oral or nasal endotracheal intubation are not possible or contraindicated;
- Patients for whom other methods, such as oral endotracheal intubation, fail;
- Patients where relief of obstruction of the upper airway is needed;
- Patients with severe facial trauma rendering other techniques impossible (Walls, 1988)

Rare procedures
As discussed by J. W. Kim et al. (2013), skill retention is especially significant in professions where important skills have to be performed that are rarely practiced. A study that J. W. Kim et al. (2013) describe, showed that only a quarter of EMS first responders could still proficiently execute cardiopulmonary resuscitation (CPR) six months after receiving their CPR training (McKenna and Glendon, 1985 as cited by J. W. Kim et al., 2013). Similar research shows that the same phenomenon is observed in respiratory procedures, such as tracheal intubation skills (Deakin, King, & Thompson, 2009). Advanced airway skills are generally used infrequently by EMS nurses. The cricothyroidotomy procedure is an example of such an infrequently performed respiratory procedure. The cricothyroidotomy, being a procedure of last resort, performed when conventional intubation is impossible (Chang et al., 1998), is even more rare than other intubation skills and therefore even more prone to skill decay. Because of the decaying nature of intubation skills, it is implied that the nurses have never reached the previously described third stage of learning (J. W. Kim et al., 2013), where most knowledge is proceduralized. If they would have reached the third stage of learning, skill decay should not or barely occur. Although the exact learning stages of nurses in different rare EMS procedures require further research, the decay of the skills is a clear indication that EMS nurses develop the skills corresponding to rare procedures up until the first or second stage of learning.

2.2.2 Spacing

Although contradicting findings have been presented in research, a general guideline is that spaced practice results in better learning outcomes than massed practice. Spaced practice means that the training of knowledge is spaced out over time, whereas massed practice means that the training is concentrated over a short period of time. Different studies that looked into spaced practice compared to massed practice worked with timespans of 24 hours (Shea, Lai, Black, & Park, 2000), 28 days (Dail & Christina, 2004) or even 8 weeks (Arthur Jr et al., 2010). All of these studies researched the acquisition of a new skill.

J. W. Kim et al. (2013) propose that the amount of spacing that is suitable in a specific learning context is dependent on the type of skill and the learning phase. Different types of skills (e.g. perceptual-motor skills or cognitive skills) differ in which strategy best supports the long-term retention of that knowledge (J. W. Kim et al., 2013). Also, the different phases a learner can be in are of influence when deciding on a proper spacing strategy. In the first stage, a learner might benefit most from a distributed practice. If a learner is about to move into the third stage, however, the performance of the learner might by optimized by massed practice, because the knowledge can then become strong enough to proceduralize (J. W. Kim et al., 2013). J. W. Kim et al. (2013), however, also mention that the literature on spacing of training is often contradicting and therefore no firm conclusions can be made at this point. Nevertheless, the above-described spacing strategies that correspond to different learning phases provide interesting suggestions. When looking at rare EMS skills, spaced practice might be the most optimal learning strategy. Skills that rely mostly on procedural knowledge do not profit from spaced practice sessions, as this knowledge does not show decay (Cecilio-Fernandes, Cnossen, Jaarsma, & Tio, 2018).

Spacing in ACT-R

The spacing effect refers to the beneficial learning effect that people experience when the duration between practice episodes is increased (Pavlik & Anderson, 2003). In a different study, (Pavlik & Anderson, 2005) proposed to relate the decay values to the activation of a particular item at the time of a particular encounter (as cited in Van Rijn et al., 2009). Van Rijn et al. (2009) describes that the second encounter of a chunk that is presented quickly after the first encounter will be associated with a higher decay value, as the activation of recently presented chunks is already high. Contrastingly, if the last encounter of a chunk was a longer time ago, the chunk has a lower decay value. The next encounter of that chunk will therefore have a longer-lasting impact on the activation of that chunk (Van Rijn et al., 2009). The following equation was described by Pavlik and Anderson, (2005, as cited by Van Rijn et al., 2009), to explain a variety of spacing-related learning concepts:

$$d_{ji} = ce^{A_i(t_j)} + \alpha \tag{5}$$

Equation 5: Decay Equation by (Pavlik & Anderson, 2005)

The decay, d, is calculated for encounter j of chunk i. The $e^{A_i(t_j)}$ represents the inverse of Equation 3 at the time of encounter j. α stands for the decay intercept, which is the minimum decay for an encounter (Van Rijn et al., 2009). The variable c, the decay scale parameter, corresponds to the relative contribution of the activation.

Optimized spacing

To optimize the retention of knowledge, the spacing of occurrences of chunks is important. The time between occurrences should not be too long (to prevent chunks from being forgotten) but also not too short (to ensure that each learning event attributes to the activation maximally). To ensure this optimal form of learning, a new occurrence of a chunk should be introduced just in time before a chunk is forgotten. This corresponds to the time at which the retrieval threshold is about to be crossed, as can be seen in Figure 5. This plot displays a solid line that represents optimal learning (where a chunk is reinforced just before it crosses the retrieval threshold, the red horizontal dashed line) and a dashed line that represents set spacing intervals. As can be seen, the number of occurrences to reach the same activation level for both spacing methods is higher for the set spacing than for the adaptive spacing. This concept forms the basis of the intelligent tutor model called SlimStampen, that was later developed by Van Rijn (2010).

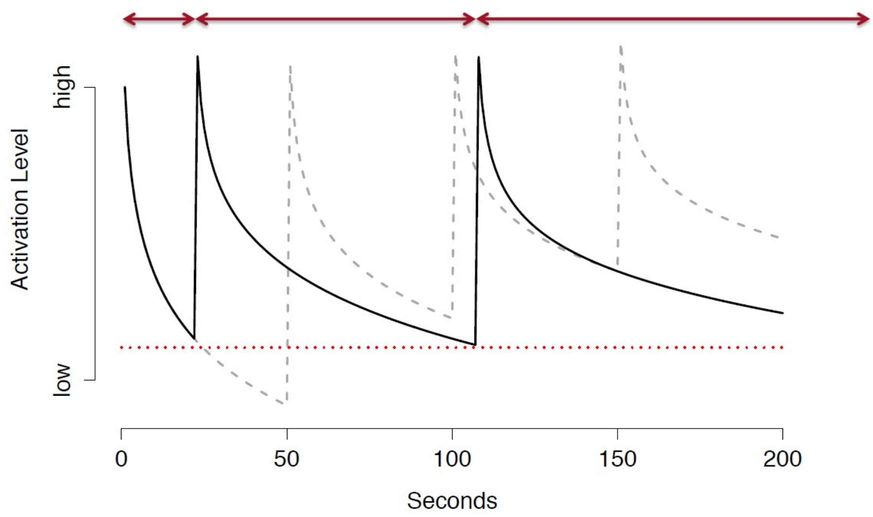


Figure 5: Optimized learning by Borst (2019b)

This model of adaptive spacing (Van Rijn, 2010), however, mainly looks at acquisition of new skills through spacing, whereas this paper is focused on knowledge that has already been acquired, but needs refreshing after a substantial period of time. The ACT-R architecture and the SlimStampen model (Van Rijn, 2010), however, provide a solid understanding of the mechanisms of declarative and procedural memory and the spacing effect.

Furthermore, behavioural research into the spacing effect (Shea et al., 2000; Dail and Christina, 2004; Arthur Jr et al., 2010) also only looks at skill acquisition of an entirely new skill, not at skill maintenance over a longer timespan.

2.2.3 Refresher interventions

Currently, medical education relies on refresher training to supplement initial training (Sullivan, Elshenawy, Ades, & Sawyer, 2019). Refresher interventions aim at re-establishing a certain level of a skill that was acquired at the end of initial training (Kluge, Burkolter, and Frank, 2012 as cited in Kluge and Frank, 2014). This training takes place a substantial time after the initial training. On top of the fact that the amount of time between the initial training and the refresher training has an influence on the success of a refresher training, this amount of time in between can also influence the willingness to attend the training (Woollard et al., 2006). A suitable amount of time between sessions is thus necessary for optimal learning. Research by Shahsavari et al. (2017) has shown that taking part in a refresher course reduces the anxiety levels of students, increases the level of clinical self-efficacy and leads to better clinical skills in medical students during their internships in their final year. A study by Salina et al. (2012) showed that the use of an educational video can be successful as an instrument to refresh a nursing skill. Students who viewed the video applied the techniques of moving a patient to a lateral position better than students who had not seen the video. Most studies that research refresher interventions, however, look mainly at infrequent refreshing, with long intervals between the initial training and the refresher training. This is also the approach of current medical education. Sullivan et al. (2019) therefore propose a framework in which skills are maintained over time, rather than refreshed over time, with less intensive re-training (displayed in Figure 6). They argue that this approach of maintaining a skill mitigates skill decay better than refresher training and prevents that the competency of students becomes insufficient. Refresher interventions would no longer be necessary if a skill is maintained over time. An interesting direction to look into as a tool for this skill maintenance is mobile applications. They allow for accessing information anywhere, rapidly, therefore providing convenience and efficiency in the day of, for example, a nurse.

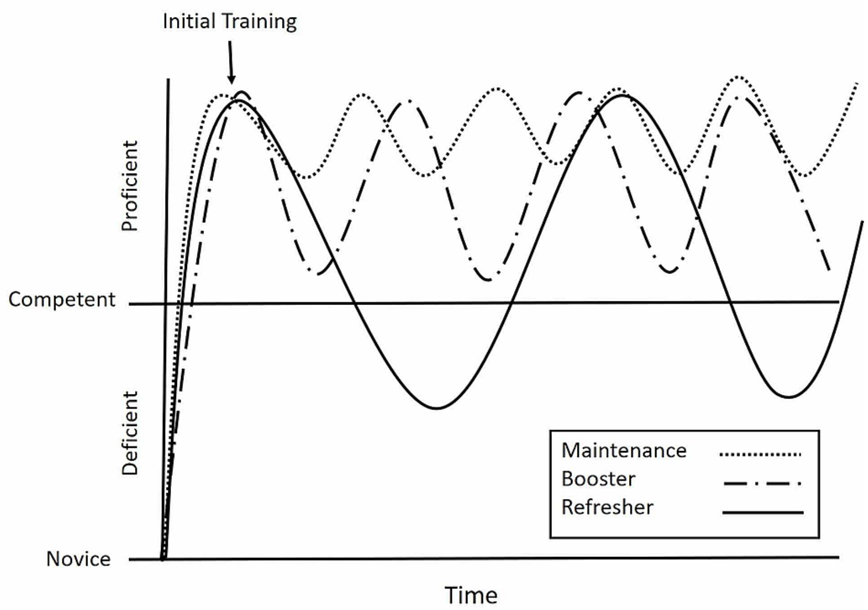


Figure 6: Maintenance, Booster and Refresher training and their corresponding competency over time, by Sullivan et al. (2019).

2.3 Mobile medical education

Nowadays, the variety of different types of media to practice or find information with is wide. For example, applications for phones and other mobile devices (i.e. tablets and smartwatches) have become mainstream in health care education over the years. In a study performed by Wallace et al. 2012, results showed that 55% of medical students, 75% of medical school faculty and 95% of residents report that using a mobile device for rapid access to educational resources had a positive educational effect for them. On top of this, it is shown that mobile applications offer invaluable tools to support clinical decision-making (Mickan et al., 2013; Mosa, Yoo, and Sheets, 2012 as cited by Ventola, 2014).
A different study created an application for advanced life support-trained doctors to practice cardiac arrest procedures (Low et al., 2011). Their findings show that the use of their application improved the doctors' scores during a standardized simulated cardiac arrest scenario, when compared with the doctors in their control group, who relied purely on their (unrehearsed) knowledge. This research is one of the ample examples that show the potential of educational medical educational applications.
Although these studies provide a clear indication that the use of medical educational applications can be beneficial, the example provided above and a considerable amount of other medical educational applications are not very immersive. The information in the applications is mainly in the form of text.

2.3.1 Media in education

A wide range of research has covered the use of different types of media in educational settings, such as virtual reality (Pottle, 2019) or simulator training (Al-Kadi et al., 2012 as cited in Cnossen, 2015). Both of these examples of educational media, however, require a physical element (VR glasses or i.e. medical phantoms).

Video-based education
Another example of a media being used in educational settings that is commonly used is video-based learning. It is widely accepted that visual media draw more attention during the learning process (Ahmet, Gamze, Rustem, & Sezen, 2018), in comparison to, for example, text-based learning materials. Video-based education has been proven to lead to significant improvements in assessment scores and acquisition of surgical skills and techniques, as compared to conventional education (Farquharson, Cresswell, Beard, and Chan, 2013; Van Det et al., 2011; Autry et al., 2013; Crawshaw et al., 2016, as cited in Ahmet et al., 2018). Ahmet et al. (2018) also mention drawbacks to video-based education, namely that some technical details might be missed because the students have to focus during the entire video. A suggestion that they do is that at important points, a static image might be more effective (2018).

J. Kim, Glassman, Monroy-Hernández, and Morris (2015) argue that using only videos in education, however, can provide a disconnected and passive learning experience. They looked into embedding interactive multimedia exercises in lecture videos. Their findings show that both teachers and students appreciate their developed concept, as it gives teachers an insight in the thought processes of their students and it allowed students to be able to interact more in the lesson, rather than just taking notes or listening (J. Kim et al., 2015). These findings are further substantiated by a meta-analysis performed by Chi (2009), where they found that as engagement with the learning material increases, a better learning-outcome is also observed. They also argue that answering in-video quizzes creates a more active and constructive learning environment (Chi, 2009).

On top of the increased engagement and decreased passivity of interactive-video education, another benefit of this type of education is the fact that it allows for self-paced learning, which is stated to improve learning (Zhang, Zhou, Briggs, & Nunamaker Jr, 2006) and task performance (Hannafin, 1985). Zhang et al. (2006) also show in their research that interactive video can have a positive effect on learner satisfaction, can result in reduced costs and time for educational institutions and allows for knowledge to be updated and maintained in a more timely and efficient manner.

Overall, interactive-video applications provide an interesting direction for skill maintenance, by decreasing the intensity of learning sessions and potentially the time between learning sessions.

3 Methods

3.1 App Development

3.1.1 Cognitive Task Analysis

Development of the interactive video application started with a structured analysis of the procedure to be learned, using a 'cognitive task analysis' (CTA). Usually, a CTA requires a large amount of research. However, due to the fact that the main aim of this research is creating a pilot application, the timeframe did not allow for an extensive CTA. In addition, the procedure under consideration, the cricothyroidotomy, does not require many complex steps. A small-scale CTA should suffice to obtain the necessary information. This CTA was performed by eliciting tacit knowledge, judgment and decision processes, environmental cues, critical incidents and goal structures of the experts performing the procedures through observation of video footage of cricothyroidotomy procedures (Canares and Valente, 2020; Bridgwater, 2018) and investigating literature about the procedure (Scrase and Woollard, 2006; Ambulancezorg Nederland, 2022).
These findings were then combined into a flowchart that describes the decision points and steps that need to be taken by the nurse. The steps can require procedural or declarative knowledge. Some steps require either a combination of declarative and procedural knowledge, or it is dependent on which learning phase a nurse is in whether it is declarative or procedural knowledge that is relied on (J. W. Kim et al., 2013). Such an example is the palpation of the thyroid cartilage with the index finger, while stabilizing the larynx with the thumb and middle finger. It could be that nurses are able to perform this step fully relying on procedural knowledge, thus simply knowing how to use their fingers. This example is similar to the previously described example of learning how to type on a keyboard: first, you have to memorize all positions of the keys, thus you rely on declarative knowledge. With practice you intuitively know where to click, thus the knowledge shifts from declarative to procedural. It could also be the case that nurses still have to remember the declarative knowledge to be able to perform the step (i.e. declarative knowledge of which fingers to use to stabilize, where to position them, that the larynx needs to be stabilized and the thyroid cartilage needs to be palpated). The steps that may require declarative knowledge are steps that are of interest to include in the interactive videos. For the cricothyroidotomy procedure, most of the steps that are involved in the execution potentially require declarative knowledge. The procedure is therefore very suitable for the concept of the application. The result of this CTA can be considered the blueprint for the interactive video analysis. The full CTA flowchart can be found in Figure 7.

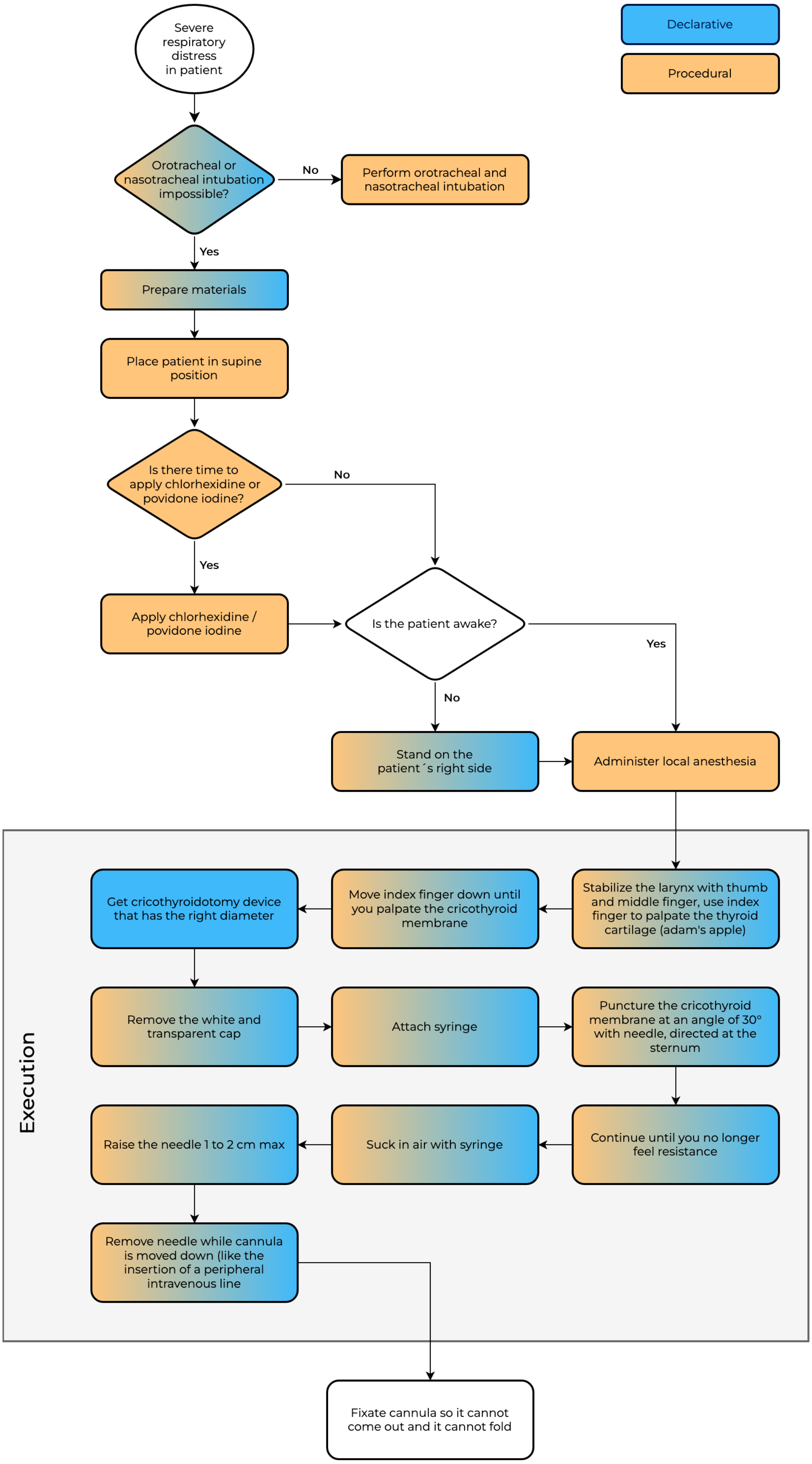


Figure 7: Flowchart of CTA

3.1.2 Videos

The videos are provided by medical professional and project supervisor Bert Dercksen. The videos are selected based on their medical accuracy. The videos are analyzed based on the CTA steps and edited with Adobe Premiere Pro. Different sections of the videos are created, such that after each section a question can appear. The cut-off point in the video for a question was set right before the answer would be provided in the video. Thus, after a participant answers the question, the following video section immediately displays and explains the correct answer. All questions are formulated in Dutch and can be found in Appendix A.

Ideally, the application would also contain personalized feedback based on the answer that was provided by the participant, explaining the reason why their answer was correct or incorrect. However, within the scope that was set for the development of this project, it was impossible to incorporate this.

3.1.3 App development

The development of the application is divided into three phases. First, the designs are created in Figma, an interface design and prototyping tool (*Figma. The collaborative interface design tool.*, 2022). Next, the designs are developed into a functioning application, using Flutter, which is an open-source UI software development kit (*Flutter. Build apps for any screen.*, 2021). Finally, the application is distributed through Apple TestFlight for iOS devices (*Testflight. Apple Developer*, 2022) and through Firebase for Android devices (*Firebase. Make your app the best it can be.*, 2022). The app is made for Dutch nurses and therefore consists of Dutch content.

Design

The scope of this research did not allow for the customary method of interface and user experience design, where end users are heavily involved. The design of the app is therefore mostly based on findings from similar applications for medical education. The color scheme of the app is deducted from those similar apps, since it is apparent that fresh-looking colors (i.e. light-blue, purple, mint-green) are popular for the purpose of medical education (see Figures 8 and 9). Also for the purpose of the app, nurses studying the material in their own time, these colors are suitable. The app should portray a calm, non-urgent, environment that invites a user to use it regularly. App environments that deal with more urgent situations, such as CPR information apps, generally utilize more pressing colors (such as bright red).

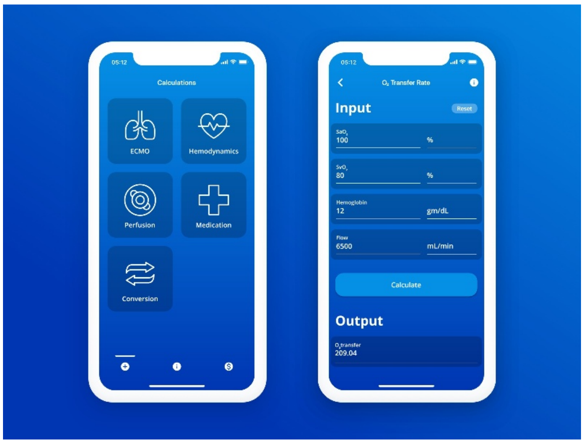


Figure 8: Mediweb — Medical Calculators by FiveDotTwelve (2020).

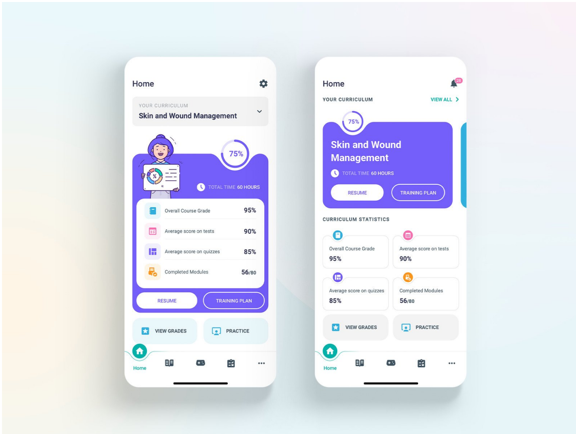
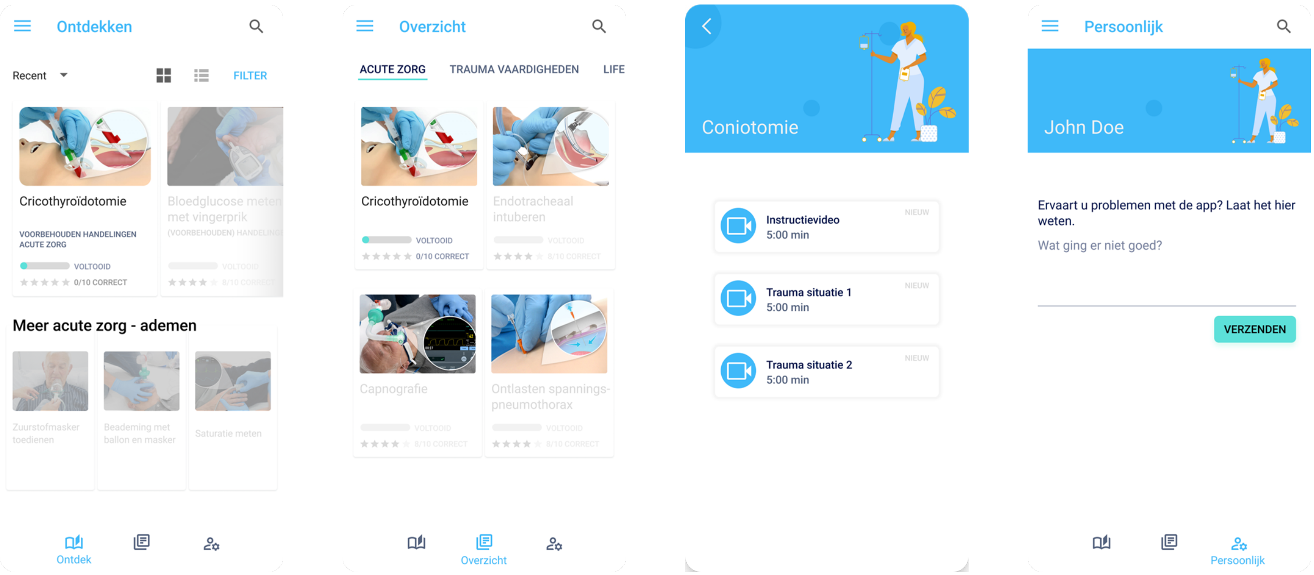


Figure 9: Learning App - Home screen by Lift Agency (Lift Agency, 2020).

Furthermore, it is important that the information that users want to view is findable (Costa, 2020), so other applications generally avoid showing too much irrelevant information in one place. This is often done by using Tile or Card designs, thereby sectioning the different blocks of information. This design pattern is also used in the current application.

Finally, the different necessary menu items are brought to a minimum, so the user flow is kept as simple as possible. This results in four main components: the *Home/Discovery* screen, the *Overview* screen, the *Video* screens and the *Personal* screen. The *Home* screen offers a quick access to the most recently viewed topics and their related topics. The *Overview* screen allows users to navigate through the different types of topics (e.g. respiratory topics or trauma skills topics) and offers a search function. The *Video* pages display an overview of the different training videos and the actual interactive videos. The *Personal* page displays the user's name and the option to give feedback on potential bugs. This addition was made because within the scope of this research, the application is likely to only be developed as a beta application. This usually brings about sporadic errors and bugs.
The Walk-Through of the complete design of the app can be found in Appendix B.



(a) *Home/Discovery* screen (b) *Overview* screen (c) *Video* screen (d) *Personal* screen

Figure 10: *Home/Discovery*, *Overview*, *Video* and *Personal* screen.

Responsiveness

The distribution of different devices of the participants and end-users is unclear, so the application design has to work for both Android phones and iPhones. A design language that is widely accepted is the Material Design system by Google (*Material Design*, 2022). The design patterns used in the application are therefore strongly based on Material Design patterns (i.e. button shapes, scale and proportion, typography, etc.)

The application also has to be responsive to the orientation of the device. The designs work for both horizontal and portrait mode. Ideally, the application would finally be responsive when switching from a phone to a tablet size. However, since the aim of this research is to create a mobile phone application, this feature is not designed for now.

Accessibility
In line with the design choices in terms of color, a color scheme is created that is also accessible for visually impaired users, for instance users who are color blind. A scheme with contrast ratios was created for all four types of colorblindness (Deuteranopia, protanopia, tritanopia and achromatopsia), which can be viewed in Figure 11.

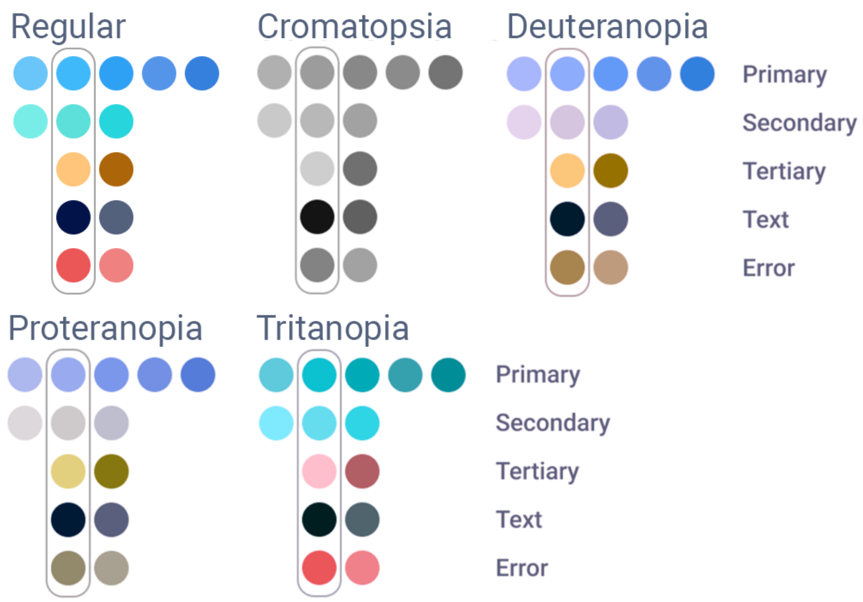


Figure 11: Color scheme accessible for color blindness

Development
For the development of the design, the open-source UI software development kit Flutter is used *Flutter. Build apps for any screen.* (2021). This kit allows for native app development for both iOS and Android, therefore ensuring an efficient process of development. When planning the development phase, it became apparent that not every designed feature can be included in the beta version of the app. To ensure proper prioritization, the MoSCoW method of prioritization is used (acronym for Must, Should, Could and Won't have) (Clegg & Barker, 1994).

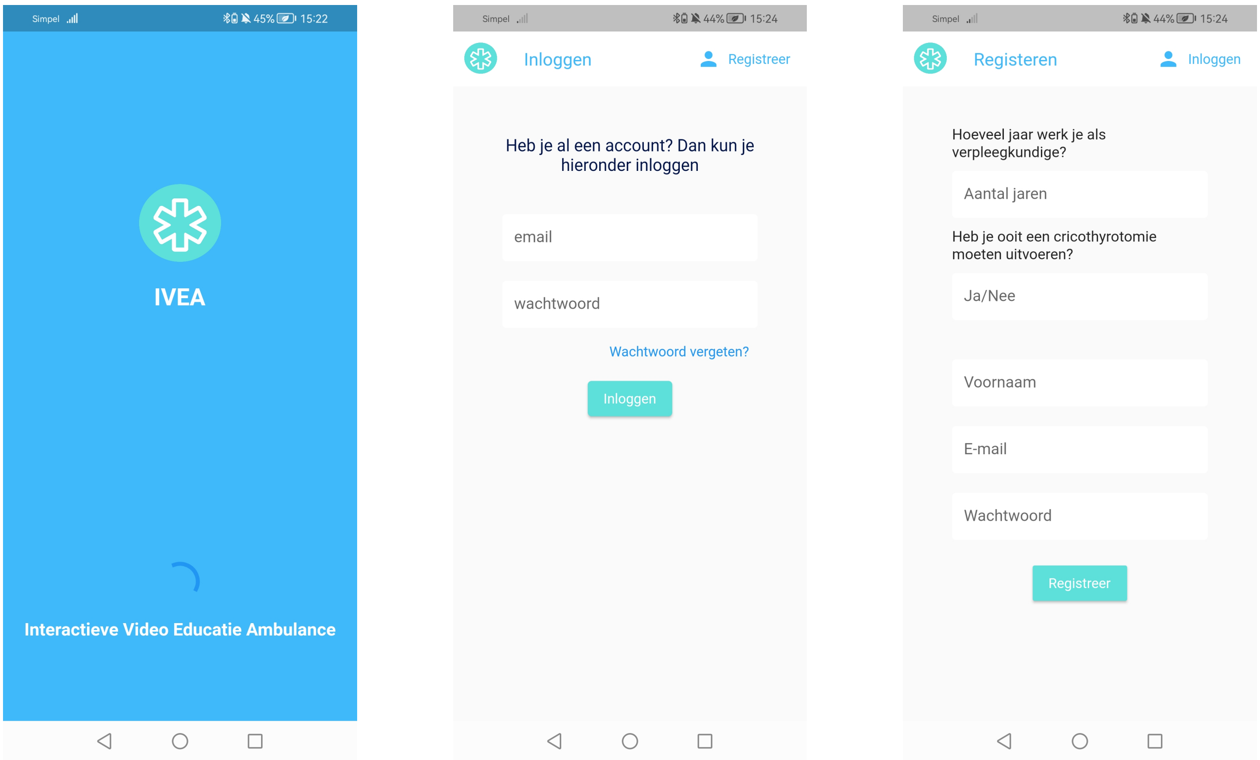
MoSCoW
The MoSCoW method is a prioritizing technique that is commonly used in software development. The method is developed by Clegg and Barker (1994) for their Rapid Application Development (RAD) approach. The first category, Must have, describes the components that are absolutely essential for the success of the product. The next category, Should have, often describes components that are similar in importance, but that are less time-critical. Those components could, for instance, also be implemented later or in a less well- developed manner. The Could have category describes the components that are desirable, but not necessary. In the final product, these components would likely improve the user experience, but they should only be included if the time and resource constraints allow it. The Won't have category, the final category, includes the components that are least critical to a good end product.
The final MoSCoW table can be found in Appendix C.

Walk-Through of app

LAUNCHING

The first screens that the user encounters are the *Loading, Welcome, Login and Register* screens. The *Loading* screen, also called a 'Splash screen', was added to introduce the user to the application before showing the actual content. This screen enhances the look and feel of the application and indicates what the content of the application is (*IVEA* - Interactieve Video Educatie Ambulance).

Next, the *Information* Walk-Through is shown to the user, as described in Figure 13.
Finally, the *Log-in* and *Register* screens serve as regular log-in and register pages, where the *Register* page also requires the input of several demographics of the user. When opening the application for the first time, the *Register* page is opened, since the user does not yet have an account to use the app with. Upon further usage (if the user is logged out), the *Log-in* screen is presented first. The user can toggle between two screens using the button in the top-right corner of the screen. When exiting the application, the log-in information remains intact. Users, therefore, only have to log-in once (unless they log out of the application themselves).



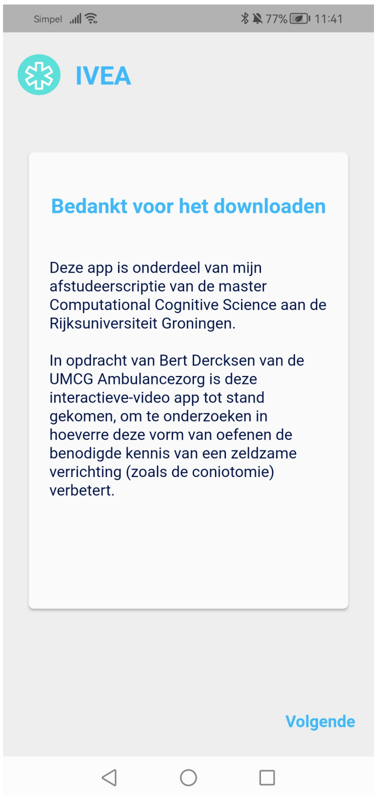
(a) *Loading* screen (b) *Login* screen (c) *Register* screen

Figure 12: *Loading, Welcome, Login* and *Register* screens.

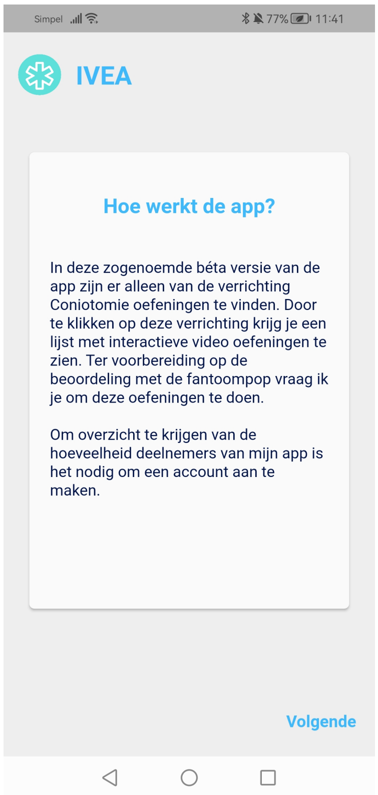
INTRODUCTION

Since the application is still in the beta stage, where not all functionalities are operational yet, some background information was deemed necessary. This background information was split up into four categories. First, a description of the master thesis is given so that the background of the development of the application is clear. Also, the participant is thanked for downloading the application. Secondly, clarification is provided for the fact that the app still only contains a low amount of content. Only the cricothyroidotomy section is available. It is also explained how the interactive training videos can be accessed. Thirdly, an explanation is provided about the use of data in the application. It is described here that the scores of the different questions of the videos are locally stored and removed from the database when the experiments have been performed. It is also stated here that an internet connection is necessary for the use of this app. Finally, information is provided on what to do when the app is not functioning in the way that it should, which is often referred to as a 'bug'. If a participant experiences a bug, contact information is provided so that they can reach out. Also, the participant is pointed to the *Personal* page, where the 'bug'-form can be found.

The information Walk-Through was added to the app design to prevent frequently occurring issues or questions to interfere with the use of the application. Upon opening the application for the first time, the participant has to click through the four screens before being able to open the actual application. This way, it is ensured that all relevant information is read before starting to use the app. If a participant wishes to read the information again, they can open the information Walk-Through again when using the app, through the *Hamburger Menu*.

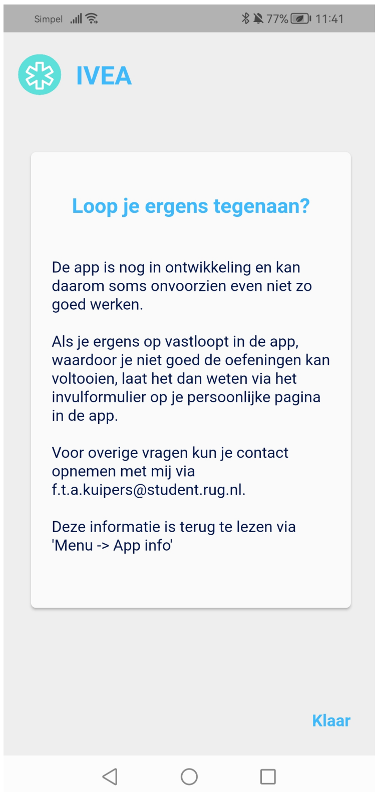


(a) First *Information* screen



(b) Second *Information* screen

(c) Third *Information* screen

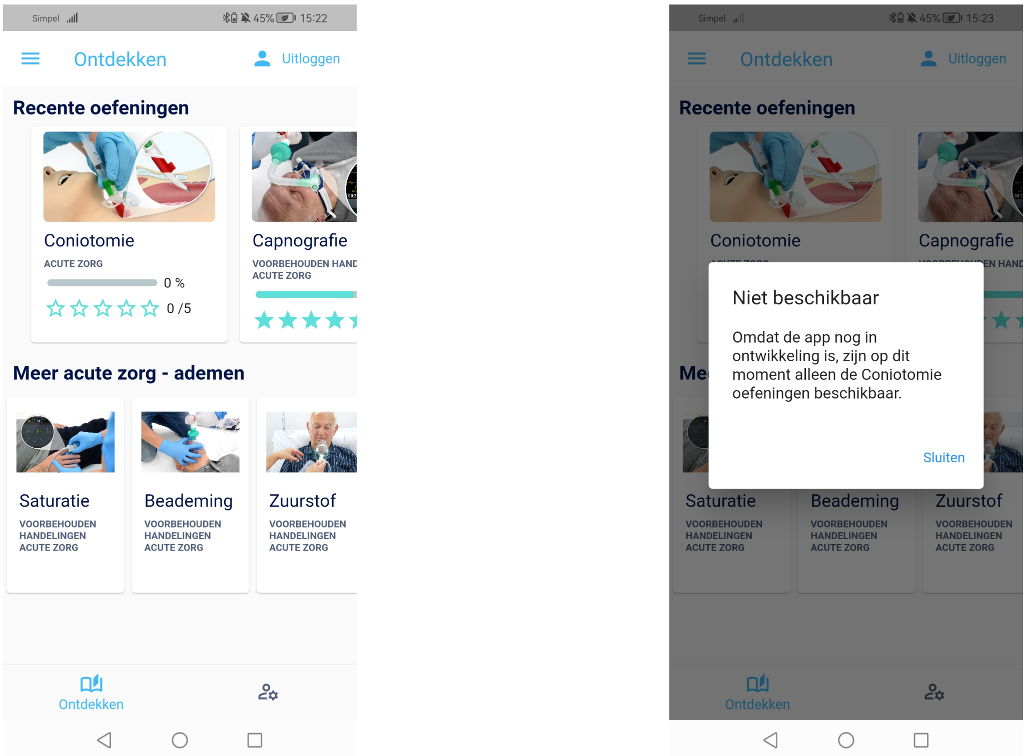


(d) Final *Information* screen

Figure 13: *Information* Walk-Through screens.

HOME

The first page that the user sees when they are logged into the application is the *Homepage*. This page contains an overview of the different topics in the application that the user can practice with. Because of the scope of this research, only the cricothyroidotomy procedure was available in the application. Therefore, when a user clicks on an unavailable tile, they see a pop-up that explains to them that the training that they tried to open is not yet available. The reason that also the tiles are displayed which have no content yet, is that this gives the application a more coherent feel. It appears more trustworthy and 'real' if more content is shown. It is also easier to imagine what the final product would look like and how it would work.

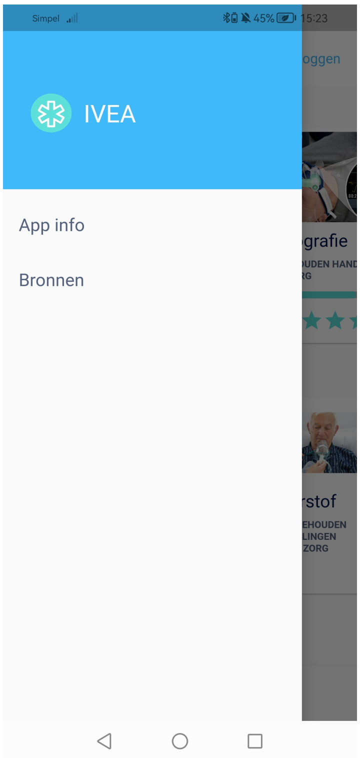


(a) *Homepage* (b) *Not Available* screen

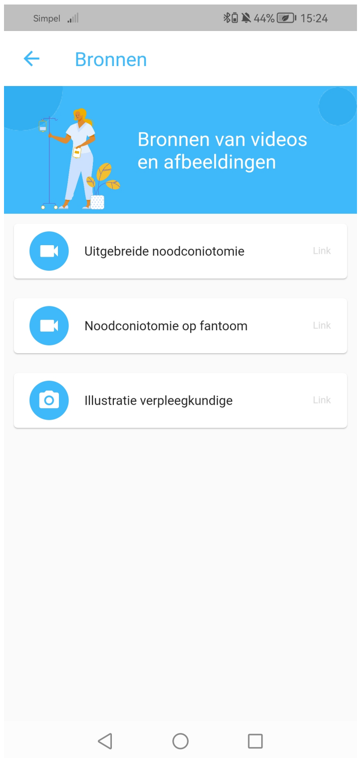
Figure 14: *Homepage* and *Not Available* screen.

INFORMATION

Additional information about the application can be accessed through the *Hamburger Menu*. On this menu, the initial *Information* Walk-Through can be found, as well as the sources of the different videos in the application.



(a) *Hamburger Menu*

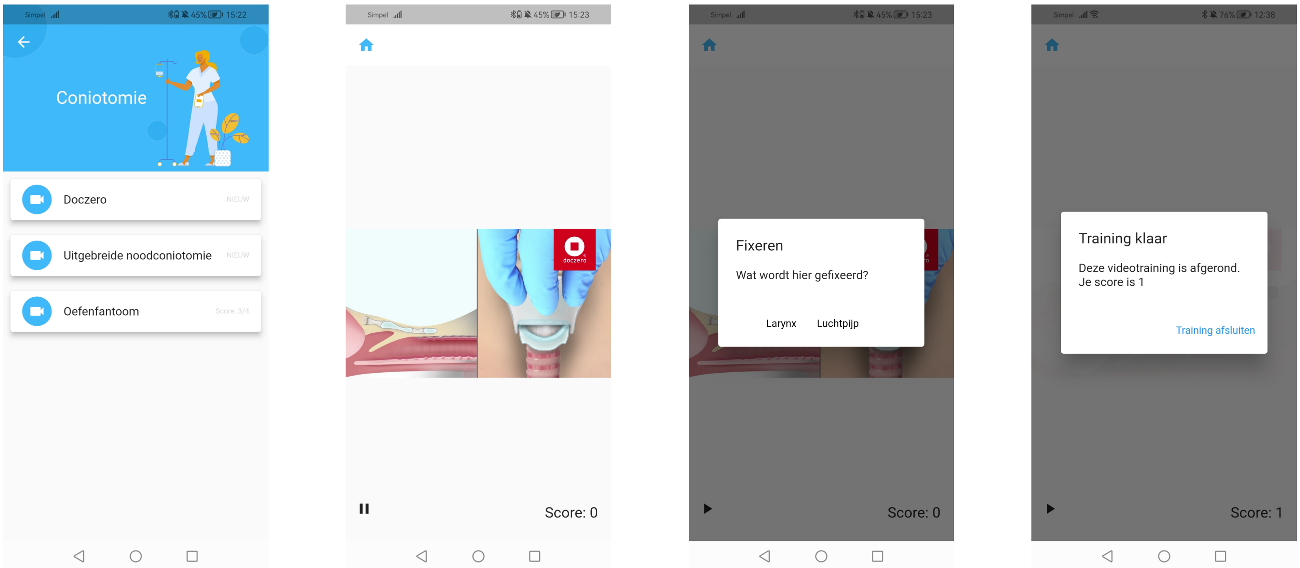


(b) *Sources Menu*

Figure 15: *Hamburger Menu* and *Sources Menu*.

Videos

The most important component of the application is the interactive video training. These videos can be accessed through the *Video Overview* screen. In this screen, the title of the page indicates the topic of the training. The sections below the title show the different available interactive videos to practice with. Behind the title of the different videos, there is a line of text that indicates whether or not the video has been viewed before. When clicking on one of these available videos, the user gets redirected to the specific video. The training starts immediately with the first section of the video. After each video segment, a question pops up that prompts the user on what the next step in the video would be. In the example in Figure 16, the question is aimed at the position that has to be fixated (either the larynx or the trachea). When answering the question through one of the buttons below the question, the user gets redirected to the next video segment of the training video. This continues until the video is completed, after which the user views a new pop-up that states that the training is finalized and what their score is.



(a) *Video Overview* screen

(b) *Video* screen

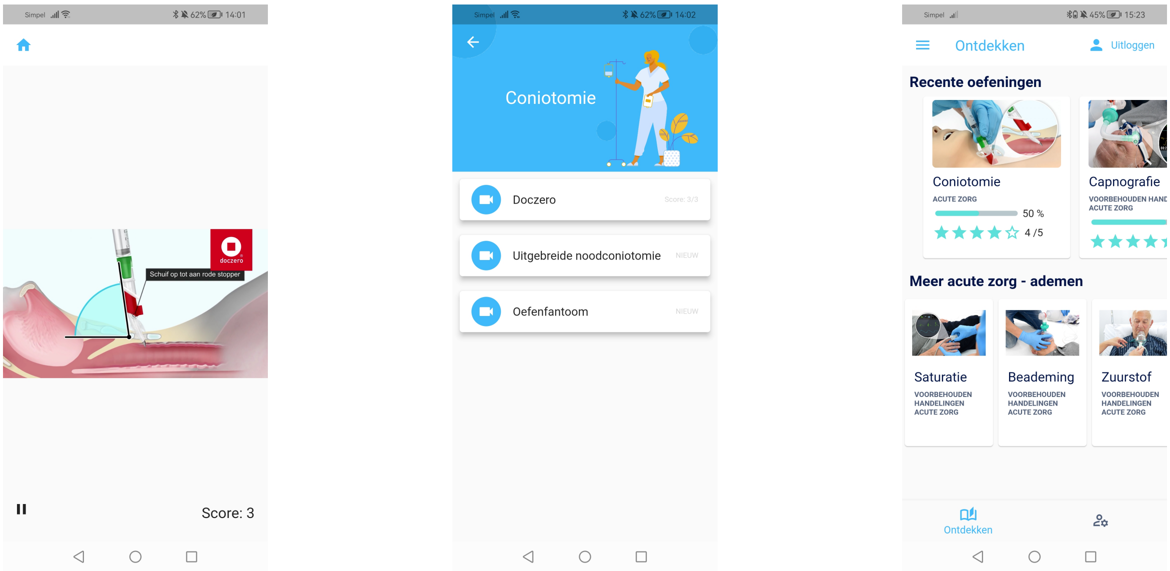
(c) *Video Questions* screen

(d) *Finalized Training* screen

Figure 16: *Video overview* and *Video* screens, accompanying *Video Questions* screen and *Finalized Training* screen.

Progress feedback

The scores that users get in the different training videos are shown first in the bottom-right corner of the *Video* screen. After the training is completed and the user returns to the training page, the 'NIEUW' text after the training video is changed to the score of the user (i.e. '3/3'). When the user then would return to the *Homepage*, the average score of all videos of that topic is displayed through stars (i.e. '4/5') and the progress is shown through a progress bar and a percentage (i.e. '50%'). These two metrics together visualize the progress of a user.



(a) *Video* screen with score (b) *Video Overview* screen (c) *Homepage*

Figure 17: Updated *Video Overview* screen, *Video Feedback* and updated *Homepage*.

CONTACT

The final page that can be accessed by the user is the *Personal* page. On this page, the user sees the name that they are logged in with, to create a personal feeling with the application (i.e. 'Floor'). Below this header section, a form is displayed through which the user can enter feedback or issues that they are experiencing with the app. By including this feedback functionality inside the app, the user flow does not get interrupted because the user no longer has to leave the app to send an e-mail about the issue.

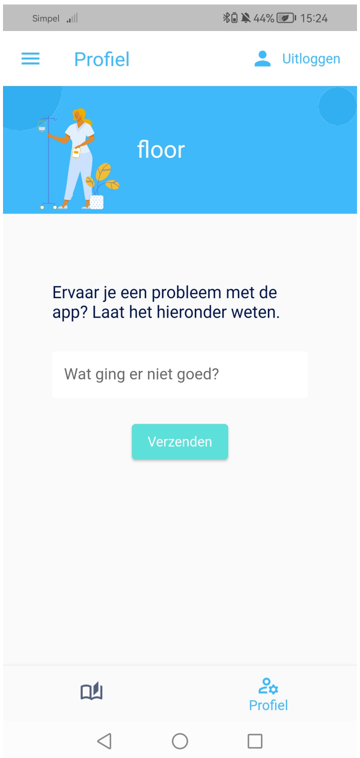
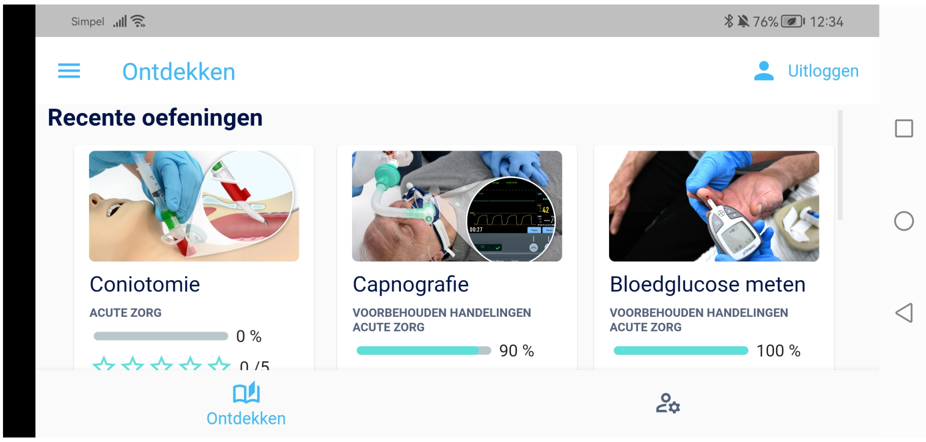


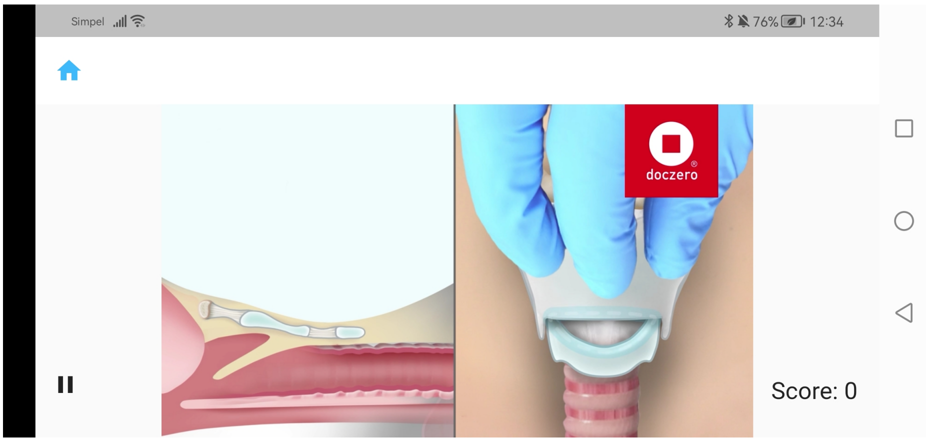
Figure 18: *Profile* page with feedback option.

RESPONSIVENESS

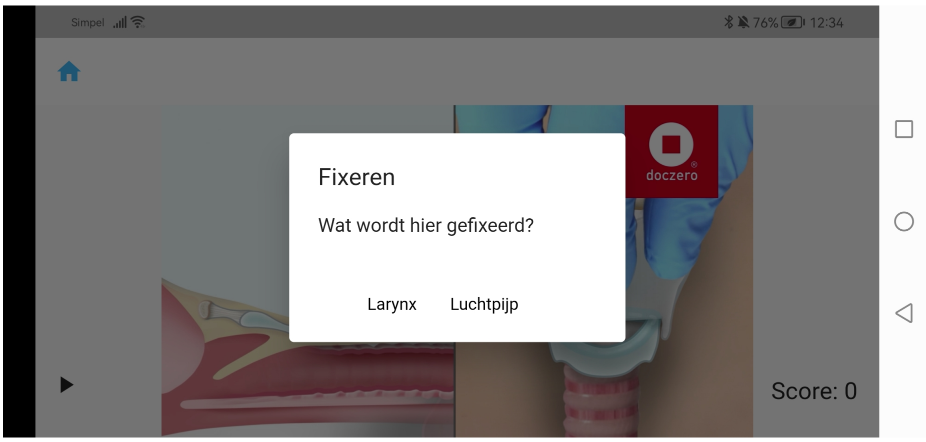
As previously described in Section *Development - Moscow*, the responsiveness of the application was deemed important to include in the end product. The images below show examples of what the landscape mode looks like.



(a) Landscape *Homepage*



(b) Landscape *Video* page



(c) Landscape *Video Answer*

Figure 19: Updated *Video Overview* screen, *Video Feedback* and updated *Homepage*.

Distribution

The most important aspect of the distribution process is to make it as accessible and easy as possible. Because of the fact that the app was merely in a beta state, it could not be published to the public domain. The easiest download methods (Android Playstore and Apple Appstore) are therefore impossible. The alternative options that are used are called Google Firebase (*Firebase. Make your app the best it can be.*, 2022) and Apple TestFlight (*Testflight. Apple Developer*, 2022). Because the download protocols are more complex this way, a clear overview of the steps that have to be taken is sent to the participants after they have given their consent to participate in the experiment.

3.2 Experiment

The experiment was aimed at answering the research question of this paper, namely:

> *"To what extent does an interactive video practice application support EMS nurses in refreshing declarative knowledge of rare EMS procedures, in particular the cricothyroidotomy?"*

. This research question was split up into the following sub-questions:

1. *"To what extent does the use of an interactive video application improve the performance of EMS nurses in the cricothyroidotomy skill?"*
2. *"To what extent is an interactive video application useful in supporting the learning process of nurses"*

The experiment comprised nurses using *IVEA* for two weeks, after which they performed a cricothyroidotomy procedure on a phantom. Their performance of this procedure was compared to the performance of nurses who did not use the application. To investigate the extent to which nurses find IVEA useful and usable, usability data was also gathered. The following sections describe the experiment in more detail.

3.2.1 Participants

Two groups of participants were recruited, namely participants that were in the control group and participants that were in the experimental group. The participants did not know of the existence of the other group or about the fact that both groups were being compared based on their performance.

Recruitment
The participants of both groups were recruited through the training coordinators of two specific ambulance posts (in Leeuwarden and in Emmen, both ambulance posts are a part of the UMCG). The training coordinators sent out the information about the experiment (see Appendix E) as well as the information that consent of participation is voluntary and can be withdrawn at any moment.
When participants consented to participate in the experiment, the researcher contacted them directly with further information on what is expected from them.

Demographics
In total, 8 participants consented to participate in the experiment. Four participants were part of the experimental group and four were part of the control group. One participant in the experimental group did not manage to successfully download the application and therefore could not take part in the experiment. All participants are Emergency Medical Service nurses in the Netherlands, at the University Medical Center Groningen (UMCG). The ages of all participants ranged from 29 to 53, with an average age of 40 years. The ages of the experimental group participants ranged from 29-31 with an average age of 30. The number of years of experience as an EMS nurse of participants in the experimental group ranged from 1-7 years, with an average of 4.3. The ages of the control group ranged from 44-53, with an average age of 47. The number of years of experience as an EMS nurse of participants in the control group ranged from 3 years to 21.5 years, with an average of 10.4 years.

3.2.2 Materials

The developed application in Section 3.1 was used as a learning tool in the experimental group.

For the experiment itself, a phantom was placed in both ambulance posts by the two training coordinators of the ambulance posts. Both ambulance posts used a slightly different phantom model, provided by the training coordinators of the ambulance posts in Emmen and Leeuwarden. Also the materials necessary for the cricothyroidotomy procedure were laid out at the ambulance posts by the training coordinators.

The execution of the cricothyroidotomy procedure by the participants was recorded through video glasses (iVue, 2022). These glasses recorded both the audio of the participant as well as the view of the participants.

The quantitative part of the usability questionnaire contained the System Usability Scale questions (SUS) by Brooke (1986)). The SUS is an industry standard when researching usability of a product. A Dutch translation of the original SUS was used (*System Usability Scale voor meten gebruiksvriendelijkheid*, 2021).

3.2.3 Procedure

This section describes the experiment protocol, the data collection methods and the data analysis methods.

Protocol

Figure 20 describes the components of the experiment protocol and how long each stage took. The sub-sections below describe these stages in further detail.

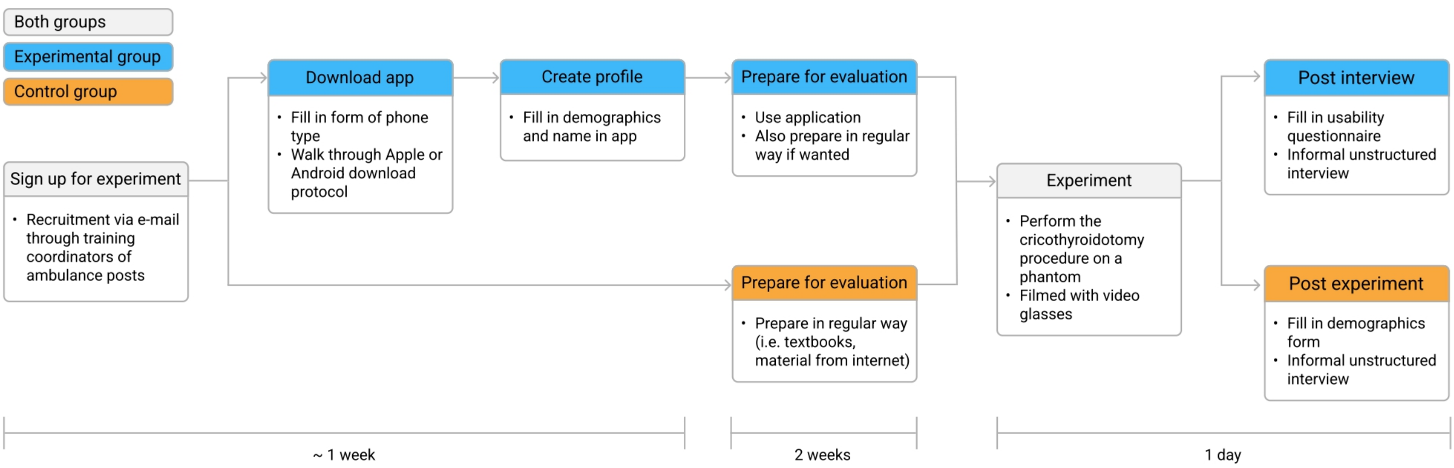


Figure 20: Experimental procedure

App download and configuration
The participants in the experimental group first received an e-mail inquiring about the type of phone that they will be downloading the application onto through an online form. Next, the download instructions were sent to the participants of the phone-type that was filled in, in the form of a pdf that includes a step-by-step, visual, instruction. This process of phone type inquiry and assisting the participants in downloading the application took approximately one week.
Once the application was downloaded on the participants' phones, they could enter the relevant demographics when creating an account for the application. These demographic questions included their name, the number of years that they have worked as a nurse and whether or not they have had to perform the cricothyroidotomy procedure before.

Experiment preparation
The participants in the control group only received one e-mail that explained to them that they could prepare in their regular manner for an interim evaluation of a cricothyroidotomy that will take place in two weeks. The experimental group participants received the request to use the application together with their regular study material to refresh their knowledge for the interim evaluation of the cricothyroidotomy that took place two weeks after receiving the email. This information was sent to them in the same e-mail as the download protocol.

Experiment
First, the ambulance post of the experimental group was visited twice in one week. The schedule of the participants that consented to participate was shared with the researcher, so the moments for the experiments could be chosen carefully. To minimize the risk of participants having to leave mid-interview because they were paged to an emergency care setting, the experiments were planned 30 minutes before the start of their shifts (where possible). The experiment took place in a meeting room in the ambulance post.

No training coordinators were present, only the experimenter and the participants. First, an explanation was provided to the participant about the structure of the experiment. Both groups performed the cricothyroidotomy procedure on the phantom, while wearing video glasses that record their moves and comments. The participants were encouraged to talk about the steps that they took and why. The researcher did not offer guidance when participants appear to struggle with the tasks.

Post-interview
A usability questionnaire was filled out by the experimental group participants after completion of the cricothyroidotomy procedure. This questionnaire was included because the usability of the product could influence the learning experience with the app. The questionnaire contains both qualitative questions as well as open, quantitative, questions. After this questionnaire there was room for questions from the participant and a short informal, unstructured, interview. The control group filled out the demographic questions (that the experimental group participants already filled out in the app). Also, there was room for questions and a short informal, unstructured, interview.
Altogether, the experiment took approximately 20 minutes.

Data collection

Video footage

The previously mentioned video glasses were used to record the actions of the participants. No personal details were displayed this way, aside from the hands of the participants.

App usage

Through the Firebase database, the usage of the app could be viewed. This usage included which videos have been viewed and which answers were correct and incorrect. Furthermore, if a participant experienced a bug, they could notify the researcher through a special page in the application. These notifications would also appear in this Firebase database. Finally, the demographic answers (age, years of occupation and whether or not they have had to execute a cricothyroidotomy before) could be found in this database.

Demographics

The demographics of participants were gathered differently for the control group than for the experimental group. The experimental group filled in their demographics inside the application, to keep the actual experiment as brief as possible (since they also have to fill in a usability questionnaire). The control group filled in their demographics after the experiment, through an online form.

Usability

To investigate the extent to which *IVEA* was useful and usable, the usability of the product was investigated.

QUALITATIVE

Qualitative data on the usability of the product was gathered through open questions in the usability questionnaire, as well as through informal unstructured interviews after the procedure was executed. The findings from the informal unstructured interviews were recorded by noting down the most significant comments from participants during and immediately after the interview.

QUALITATIVE

Quantitative usability data were gathered through the SUS questions.

Data analysis

Performance measures

EVALUATION BY MEDICAL EXPERTS

The video footage that was recorded with the video glasses was evaluated by three EMS educational experts. The videos were qualitatively evaluated, describing per participant to what extent the appropriate steps were taken in the correct manner during the procedure. All three experts evaluated the video footage separately without discussion of the results. This way, the findings were not influenced by each other. All videos were anonymized by cutting the video precisely so that it started with the beginning of the procedure and ended when the procedure is finalized. The 'small talk' before and afterward were removed, to ensure that no personal details were revealed to the evaluators.

Usability

QUANTITATIVE

The SUS questionnaire results were described through descriptive statistics as well as a SUS score, which indicates the overall usability of the product (Brooke, 1986). The SUS score can be interpreted using the scale in Figure 21.

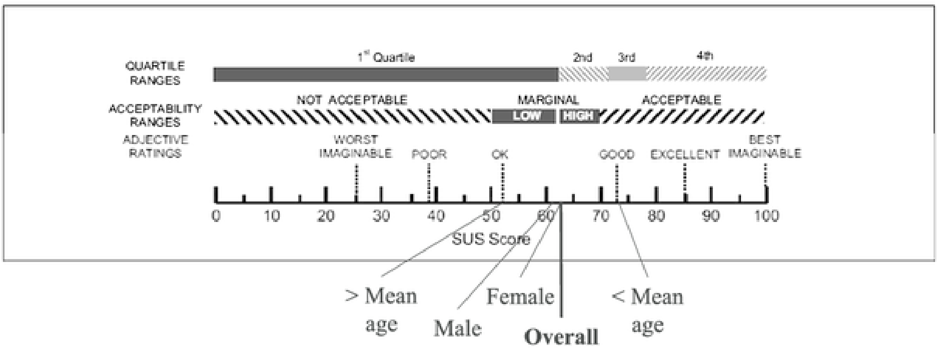


Figure 21: SUS score scale by Vaziri et al., 2016

OBSERVATION OF VIDEOFOOTAGE

The video footage was also observed with a focus on the comparison between comments by the participants from the experimental group and the questions and answers in the application. The specific phrasing of the content of the application (e.g. questions, answers, spoken text in the videos) was compared to the literal comments of participants during the execution of the procedure. Also, the results from the app usage (e.g. which questions were answered correctly and incorrectly) were compared to the performance in the video footage.

QUALITATIVE

The analysis of the qualitative usability findings was performed through a thematic content analysis (R. Anderson, 2007). The qualitative findings of the open questionnaire questions and the informal unstructured interview were first combined. Relevant quotes and descriptions were then highlighted. The criteria for 'relevant' descriptions and quotes were that they were only focused on the use of the app, not on the download process (since this was not representative of how the actual download process would be). Furthermore, the focus was on finding indications of usability issues and the underlying opinions and motivations of participants. Furthermore, suggestions and positive comments were considered 'relevant'. This process is referred to as 'coding' (R. Anderson, 2007).

Next, the coded findings are combined into groups of similar themes. Suitable themes were then drafted for these different groups. It was critically assessed whether all categories describe the coded findings fully, or whether a category could potentially be split up into multiple categories.

4 Results

4.1 Performance measure

4.1.1 Evaluation by medical experts

Two out of three participants in the experimental group did not perform well in the opinion of the three medical experts. The general consensus was that both participants took too long to perform the overall procedure, did not perform the procedure in a sterile manner, palpation was not performed efficiently and the insertion of the needle was done incorrectly (i.e. the angle of insertion). The other experimental group participant, however, executed the procedure perfectly. This participant took all the appropriate steps in a swift and correct manner.
The control group generally performed well, with, again, one participant who executed the procedure perfectly and one participant who executed the procedure nearly perfectly. The other two participants, however, made a few mistakes. Most of these mistakes were made during the palpation of the phantom. Also, both participants did not perform the procedure in a sterile manner.
The entire synthesis of the evaluations by the medical experts can be found in Appendix G.

4.2 Usability

4.2.1 App usage

All three users completed all videos and all corresponding questions. Generally, the questions were filled in correctly, apart from two questions: “At what angle do you insert the needle?”, which two participants answered incorrectly, and “When removing the needle, what should you pay attention to?”, which one participant answered incorrectly. One of these participant answered that the needle should be inserted at a 45-degree angle in the app twice. One participant answered all questions correctly. This is also the participant who executed the procedure perfectly. Two participants only used the application one day, the other participant used *IVEA* multiple days.

4.2.2 Observation of videofootage

Two participants mentioned most of the steps out loud that were also highlighted in the interactive videos in the questions. All participants mentioned that they raised the needle until air comes into the syringe. Two participants mention that they struggle with the material of the phantom (the rubber material of the cricath did not move smoothly into the rubber patch on the phantom). Two participants mentioned the angle at which the needle is supposed to be inserted. One of these participants states that the needle should be inserted at a 45-degree angle. This is the same participant that also incorrectly answered twice that the needle should be inserted at a 45-degree angle in the application. One participant mentioned that he had to ’peek’ under the rubber of the phantom, because he was unsure of the position of needle insertion. Other than this comment, no participant specifically mentioned that the failure of a step in the procedure was due to lacking knowledge.

4.2.3 Quantitative usability

The averaged results of the SUS questions can be found in Figure 22. The score that resulted from this is 80.83, as calculated through the formula that was created by Brooke (1986). This score corresponds to a ’good’ to ’excellent’ on the adjective rating scale (Vaziri et al., 2016). In terms of usability, this indicates that the product is acceptable to use.

Figure 22: SUS analysis

4.2.4 Qualitative usability

Notable comments from participants were written down during the conversation. Afterwards, these notes were later combined with the findings from the open-ended questions from the questionnaire and structured using a thematic coding approach. From this thematic coding, it became apparent that four main themes could be established: the current situation of studying for an evaluation, potential improvements for the application, positive notes of participants and technical difficulties experienced with the application. The suggestions for improvements and comments about the current situation can provide interesting directions for future development and research. Also indications were found of limitations of the application, namely the fact that the videos were considered somewhat inconsistent and incoherent. The different videos used different materials for the procedure. Overall, the positive notes give indications that the concept of IVEA would be useful in the everyday lives of the nurses. The complete results can be viewed in Table 1.

Current situation
Generally, there is not much usable video footage out there to practice with
Even the information of the actual manufacturer is not informative in terms of procedure steps (Ventrain)
The usual methods of preparing of the evaluation do not fully correspond to how the nurses actually practice the cricothyroidotomy
The usual methods of preparation include: - Reading through the protocol - Watching YouTube videos - ’Zenya DOC’- an internal database with all protocols

Improvements
The app contains no videos of the actual material that’s used in their ambulance post (Ventrain)
The videos are a bit inconsistent and incoherent
The different videos used different kits and materials for the cricothyroidotomy
More videos would make the app more realistic
More theoretical information aside from the videos would be a nice addition. For example, the right angle of the needle and general hick-ups of the procedure and how to avoid them.
Participant was unsure whether he maybe missed something, since “There were only three videos”.

Positive
All three participants state that it would definitely be useful to have an app to practice
Two out of three participants did not deem it necessary to use other learning materials, aside from IVEA
Two participants mentioned that it would be nice to use such an app in your free time, rather than only when an evaluation is coming up.
Direct quote: “Every once in a while I think about ‘Oh, right, that procedure. If I had to perform that one right now, I would not fully remember how to do this. I should look into that”

Technical
Process of downloading was not what the participant was used to and she mentioned that it was difficult for her.
The participant experienced a bug once (the screen kept refreshing), but she mentioned that it did not hinder her use of the app. She managed to close and reopen the app, which fixed the problem.

Table 1: Thematic content analysis

5 Discussion

This research aimed at answering the research question:

> *"To what extent does an interactive video practice application support EMS nurses in refreshing declarative knowledge of rare EMS procedures, in particular the cricothyroidotomy?"*

This research question was split up into the following sub-questions:

1. *"To what extent does the use of an interactive video application improve the performance of EMS nurses in the cricothyroidotomy skill?"*
2. *"To what extent is an interactive video application useful in supporting the learning process of nurses"*

To answer these question, the application '*IVEA*' was developed. An experiment was performed where two groups of nurses were compared based on their performance of a cricothyroidotomy procedure. Before this procedure was executed by the nurses, one group prepared by learning with IVEA and, if they deemed it necessary, their regular method of preparation. The other group only used their regular method of preparation.

The results showed that participants generally appreciated the concept of *IVEA* and that they indicate that they would use such an application. Furthermore, *IVEA* received a high usability score, of 83%, indicating that the product can be considered 'good' to 'excellent', in terms of its usability. Although the product is deemed both useful and usable, the matter of performance improvements remains debatable.

5.1 Did *IVEA* improve performance?

The results from expert evaluation of the video footage of the experiment showed that generally, the control group participants tended to perform better at the cricothyroidotomy procedure than the experimental group participants. From these results, however, no firm, significant, conclusion can be drawn about the performance due to the most confounding factor of the experiment, the low number of participants.

5.1.1 Age and experience as confounding factors

Another important limitations to note here is that the participants in the experimental group were on average a lot younger and had fewer years of experience than the participants in the control group. Where the control group ranged in years of experience from 3 to 21,5 years, the experimental group ranged between 1 and 7 years. The experimental group showed the most mistakes that can be traced back to incorrect declarative knowledge (e.g. the incorrect angle of insertion or palpation of the incorrect location). The steps that went wrong in the execution of the procedure by these participants, were also the steps in the CTA that were marked as 'Declarative knowledge' (see Figure 7. It is likely that over the years, after several re-evaluations of the cricothyroidotomy skill, the more experienced nurses have moved further along the learning phases than the less experienced nurses. Although no firm statements can be made on this regard due to the low number of participants, it is possible that most of the experimental group nurses have developed the skill up until learning stage one, where the control group nurses might be further along (e.g. in stage two, see Figure 4). This would mean it is easier for more experienced nurses to recall the necessary knowledge (with less preparation) than it would be for less experienced nurses. Overall, both groups showed instances where necessary declarative knowledge was either incorrect or no longer retrievable, indicating that the assumption that nurses generally learn rare skills only up to learning stage one or two is realistic.

5.2 Is *IVEA* useful?

Even though no significant results were found that show that the use of an interactive video application improves performance, all three nurses from the experimental group describe that they would find such an application useful as a supporting educational tool. They state that it would be nice to use such an app in their free time, rather than only when they have to go through an evaluation. One nurse even mentioned the fact that she often thinks about procedures that she has not practiced in a while and that she then looks into that procedure. This is a perfect example of where an app that contains learning materials for all rarely practiced procedures could be useful. Furthermore, many nurses from both the experimental and the control group mention the use of YouTube videos to refresh their knowledge of a procedure when an evaluation is coming up. Also the use of an internal database that contains all protocols is mentioned, indicating that a product that contains all information about different procedures is useful.

Overall, the participants indicate that the usability of the product is somewhere between 'good' and 'excellent'. Generally, in usability research, the number of participants that are needed to reach a firm conclusion is rather low, namely 3-7 participants (Nielsen & Landauer, 1993). If there were clear usability issues in the application, they would likely have been found, even with only three participants. Furthermore, the findings from the thematic content analysis substantiate the fact that an application that would help nurses practice in their free time would benefit the nurses.

5.3 Limitations

Several limitations can be assumed that affect the results of this research and potential conclusions that are drawn from it. These limitations can be split up into limitations in development and experimental limitations.

5.3.1 Development

The main limitation was the lack of medical background of the researcher. This complicated the development of proper medical questions in the interactive videos, as well as finding more suitable video material. It also led to a less critical view on the coherence of video material and the medical accuracy of the content of those videos. The researcher had to fully rely on the suggestions of a medical professional for which videos to use. From the results of the thematic content analysis, it became clear that the videos were considered somewhat inconsistent and incoherent, due to, for instance, the different kits and materials used for the procedures in the videos.

5.3.2 Experiment

The lack of medical background of the researcher also affected the experiments, both positively and negatively. On the one hand, the researcher was not able to influence the participants by giving hints on what to do. On the other hand, it was difficult to support the participants properly when materials were not cooperating. When participants struggled to insert the needle, the researcher did not know to what extent this was due to the materials (rubber materials on rubber materials generally tend to move stiffly) or to the participant's actions. There were also other limitations caused by the materials. The most important limitation had to do with the fact that the rubber that was placed over the phantom had already had holes poked in it from previous needle insertions. This implies that part of the declarative knowledge that is needed for the task is now present in the environment and therefore no longer has to be present in the participant's knowledge base. To truly test whether a participant knows all necessary declarative knowledge, environmental cues that would normally not be present should be eliminated as much as possible. The second limitation of the material was due to the fact that two different phantoms were used and one of the two received more complaints on the material of the phantom. The rubber of the phantom of the experimental group seemed to cause more friction with the Cricath (Ventinova Medical, 2022) than the phantom of the other group. Although the different phantoms should not have had an effect on the results, as stated by the two training coordinators of the ambulance posts, the possibility can not be fully ruled out that they did have an effect. A final limitation caused by the material is that the video glasses had a tendency to get foggy, leading to reduced sight of the participant and a potentially negative impact on their performance. Furthermore, the variation of the age range of the participants in the experimental group was too small (age range 29-31). The participants that chose to take part were all relatively young, therefore may be more interested in using technology as a learning tool than nurses in an older age group would be. As mentioned, a side effect from this young average age was that nurses were on average less experienced in the experimental group than in the control group. This likely influenced the performance of the nurses. As mentioned, however, the most important limitation was that the number of participants was insufficient for statistically significant findings. To be able to draw generalizable conclusions from the research, a much higher number of participants is needed.

5.4 Recommendations

5.4.1 Feedback in the app

An interesting observation from the video footage was the fact that the participant who answered the question about the angle of insertion incorrectly twice in the application, also inserted the needle in the incorrect angle at the cricothyroidotomy evaluation. This participant has stored a piece of declarative knowledge of the procedure incorrectly, namely the degree at which the needle should be inserted. The participant continued to retrieve the wrong information (the angle of insertion is 45 degrees, instead of 30 degrees), even after studying with the application.
Even though the application provided feedback to participants in the form of a score and through information in the video directly after a question was answered, it can be inferred from these results that not enough feedback was provided for the participant to recognize their mistake. The scope of this project did not allow for extensive development of feedback, as the topic of feedback in learning can be seen as a whole research area on its own. It therefore provides an interesting direction for future development of this application.

5.4.2 Content of the app

Another finding from the qualitative usability data was that the video material was not consistent and did not contain the materials that are actually used on their ambulance posts. The *Ventrain and Cricath* (Ventinova Medical, 2022), the material kits used by the nurses that took part in the experiment, require a different method of inserting the catheter than one of the videos showed. The videos therefore did not always fully adhere to the protocol that is followed at the ambulance posts. Furthermore, it appeared that the nurses would find it useful if the app did not only contain video material, but also other educational resources (such as text and images). In the next version of the application, the video content should be better aligned with the used protocols and complemented with other educational resources. To ensure proper alignment of the videos with medical standards, a medical professional should be more actively involved in the video analysis. Another important addition for future development would be to incorporate a wider variety of procedures in the application, since this would encourage users to practice with the application in a more natural and immersive manner.

5.4.3 Future research

As previously mentioned, maintenance training might be favorable to refresher training, to ensure that competency of skills never becomes insufficient (Sullivan et al., 2019). This research only looked at the use of *IVEA* as a refresher tool, however. To investigate to what extent *IVEA* would also be useful for maintenance training, an experiment should be conducted over an extended period of time, with well worked out spacing periods between the training sessions. Within the scope of this research, in terms of time, it was unfortunately impossible to perform an experiment where nurses would use the application over an extended period of time. In a real-life situation, this way of using *IVEA* would be more realistic, however. In the performed experiment, participants were free to choose where and when to use the application, as long as it was within the period of use of two weeks, to ensure usage in the most natural way for users. Two participants used the application during one day, where the other participant used the application on multiple days. In the future, it would be interesting to look into more specific short-term intervals (in the order of minutes) and long-term intervals (in the order of days, weeks or even months). It could be examined to what extent adapting these intervals to the specific learner, through the use of an intelligent tutor, such as SlimStampen (Van Rijn, 2010), would benefit the learning experience. This would also shed more light onto the learning stage at which most nurses are in for the cricothyroidotomy skill (or another rare EMS skill).

6 Conclusion

This research investigated the effectiveness of an interactive video application named '*IVEA*' as a tool for EMS nurses to practice their skills in rare procedures. It has been shown that the nurses would appreciate the use of such an app in their daily lives. It is indicated by the nurses that the application '*IVEA*' is both useful and usable. However, to what extent *IVEA* actually improved the nurses performance in a cricothyroidotomy procedure could not be established due to the low number of participants.

Beyond the application that was developed during this project, this research has shown a clear indication that there is a need for a more coherent application for refresher education. It was apparent that nurses searched for the materials themselves on the internet, leading to sources that would sometimes claim different things. This can lead to misconceptions and confusion.

Overall, *IVEA* shows promise in supporting nurses in practicing skills that rarely have to be used, in an interactive and accessible way. This project can be considered a step towards better, more coherent refresher education for nurses of the emergency medical services of the University Medical Center Groningen. An application such as *IVEA* would give nurses the tools to refresh knowledge when they deem it necessary, at their own pace and in an interactive and immersive way. This could, in the future, also reduce the workload of trainers that have to re-learn skills to nurses.

References

ACT-R. (2022a). *Unit 1: Understanding production systems.* Retrieved from http://act-r.psy.cmu.edu/wordpress/wp-content/themes/ACT-R/tutorials/unit1.htm

ACT-R. (2022b). *Unit 4: Activation of chunks and base-level learning.* Retrieved from http://act-r.psy.cmu.edu/wordpress/wp-content/themes/ACT-R/tutorials/unit4.htm

Ahmet, A., Gamze, K., Rustem, M., & Sezen, K. A. (2018). Is video-based education an effective method in surgical education? a systematic review. *Journal of surgical education*, *75*(5), 1150–1158.

Ambulancezorg Nederland. (2022). *Azn (version 2.2.0).* [Mobile App]. Retrieved from [https://apps.apple.com/nl/app/ambulancezorg/id888657032]

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological review*, *89*(4), 369.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* Oxford University Press.

Anderson, J. R. (2014). *Rules of the mind.* Psychology Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological review*, *111*(4), 1036.

Anderson, J. R., & Lebiere, C. J. (2014). *The atomic components of thought.* Psychology Press.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological science*, *2*(6), 396–408.

Anderson, R. (2007). Thematic content analysis (tca). *Descriptive presentation of qualitative data*, 1–4.

Arthur Jr, W., Day, E. A., Villado, A. J., Boatman, P. R., Kowollik, V., Bennett Jr, W., & Bhupatkar, A. (2010). The effect of distributed practice on immediate posttraining, and long-term performance on a complex command-and-control simulation task. *Human Performance*, *23*(5), 428–445.

Autry, A., Knight, S., Lester, F., Dubowitz, G., Byamugisha, J., Nsubuga, Y., ... Korn, A. (2013). Teaching surgical skills using video internet communication in a resource-limited setting. *Obstetrics & Gynecology*, *122*(1), 127–131.

Benneck, J. C., & Bremer, A. (2019). Registered nurses' experiences of near misses in ambulance care–a critical incident technique study. *International emergency nursing*, *47*, 100776.

Borst, J. (2019a, December). *Lecture slides in architectures of intelligence - lecture 4.* Faculty of Science and Engineering, University of Groningen.

Borst, J. (2019b, December). *Lecture slides in architectures of intelligence - lecture 5.* Faculty of Science and Engineering, University of Groningen.

Bridgwater, C. (2018). *Emergency needle cricothyrotomy procedure explained.* [Video]. Northern Indiana. Retrieved from https://www.youtube.com/watch?v=a4I0MEwtXFY

Brooke, J. (1986). System usability scale (sus): a quick-and-dirty method of system evaluation user information. *Reading, UK: Digital equipment co ltd*, *43*, 1–7.

Canares, T., & Valente, J. (2020). *Needle cricothyroidotomy.* [Video]. Rhode Island Hospital, Brown University. Retrieved from https://youtu.be/F_PV7N2c2pQ

Carlson, J. N., Karns, C., Mann, N. C., Jacobson, K. E., Dai, M., Colleran, C., & Wang, H. E. (2015, August). Procedures performed by emergency medical services in the united states. *Prehospital Emergency Care*, *20*(1), 15–21. doi: 10.3109/10903127.2015.1051682

Cecilio-Fernandes, D., Cnossen, F., Jaarsma, D. A., & Tio, R. A. (2018). Avoiding surgical skill decay: a systematic review on the spacing of training sessions. *Journal of surgical education*, *75*(2), 471–480.

Chang, R. S., Hamilton, R. J., & Carter, W. A. (1998). Declining rate of cricothyrotomy in trauma patients with an emergency medicine residency: implications for skills training. *Academic Emergency Medicine*, *5*(3), 247–251.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in cognitive science*, *1*(1), 73–105.

Clegg, D., & Barker, R. (1994). *Case Method Fast-Track: A Rad Approach (Computer Aided System Engineering).*

Cnossen, F. (2015). Cognitive skills in medicine: an introduction. *Panvascular medicine*, 4719–4752.

Costa, C. D. (2020, Jan). Retrieved from https://uxdesign.cc/the-7-factors-that-influence-user-experience-2805282616f9

Crawshaw, B. P., Steele, S. R., Lee, E. C., Delaney, C. P., Mustain, W. C., Russ, A. J., ... Champagne, B. J. (2016). Failing to prepare is preparing to fail. *Diseases of the Colon & Rectum*, *59*(1), 28–34.

Dail, T. K., & Christina, R. W. (2004). Distribution of practice and metacognition in learning and long-term retention of a discrete motor task. *Research quarterly for exercise and sport*, *75*(2), 148–155.

Deakin, C., King, P., & Thompson, F. (2009). *Prehospital advanced airway management by ambulance technicians and paramedics: is clinical practice sufficient to maintain skills?* British Association for Accident and Emergency Medicine.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review*, *100*(3), 363.

Farquharson, A., Cresswell, A., Beard, J., & Chan, P. (2013). Randomized trial of the effect of video feedback on the acquisition of surgical skills. *Journal of British Surgery*, *100*(11), 1448–1453.

Fernandes, D. C. (2018). Knowledge and skills acquisition in medical students: exploring aspects of the curriculum.

Figma. the collaborative interface design tool. (2022). Retrieved from https://www.figma.com/

Firebase. make your app the best it can be. (2022). Retrieved from https://firebase.google.com/

Fitts, P. (1964). Perceptual-motor skill learning. In *Categories of human learning* (pp. 243–285). Elsevier.

FiveDotTwelve. (2020). *Mediweb — medical calculators.* Retrieved April 10th 2022, from https://dribbble.com/shots/12120419-MediWeb-Medical-Calculators

Flutter. build apps for any screen. (2021). Retrieved from https://www.https://flutter.dev/

Hannafin, M. J. (1985). Empirical issues in the study of computer-assisted interactive video. *ECTJ*, *33*(4), 235–247.

Herault, R., Lincke, A., Milrad, M., Forsgärde, E.-S., & Elmqvist, C. (2018). Using 360-degrees interactive videos in patient trauma treatment education: design, development and evaluation aspects. *Smart Learning Environments*, *5*(1), 1–15.

iVue. (2022). *ivue videoglasses.* Retrieved from [https://ivuecamera.com/]

Jones, G., Ritter, F. E., & Wood, D. J. (2000). Using a cognitive architecture to examine what develops. *Psychological Science*, *11*, 93-100.

Kettanurak, V., Ramamurthy, K., & Haseman, W. (2001). User attitude as a mediator of learning performance improvement in an interactive multimedia environment: An empirical investigation of the degree of interactivity and learning styles. *International Journal of Human-Computer Studies*, *54*(4), 541–583.

Kim, J., Glassman, E. L., Monroy-Hernández, A., & Morris, M. R. (2015). Rimes: Embedding interactive multimedia exercises in lecture videos. In *Proceedings of the 33rd annual acm* *conference on human factors in computing systems* (pp. 1535–1544).

Kim, J. W., Ritter, F. E., & Koubek, R. J. (2013). An integrated theory for improved skill acquisition and retention in the three stages of learning. *Theoretical Issues in Ergonomics Science*, *14*(1), 22–37.

Kluge, A., Burkolter, D., & Frank, B. (2012). “being prepared for the infrequent”: A comparative study of two refresher training approaches and their effects on temporal and adaptive transfer in a process control task. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 56, pp. 2437–2441).

Kluge, A., & Frank, B. (2014). Counteracting skill decay: four refresher interventions and their effect on skill and knowledge retention in a simulated process control task. *Ergonomics*, *57*(2), 175–190.

Lift Agency. (2020). *Learning app home screen.* Retrieved April 10th 2022, from `https://dribbble.com/shots/10410012-Learning-App-Home-screen`

Low, D., Clark, N., Soar, J., Padkin, A., Stoneham, A., Perkins, G. D., & Nolan, J. (2011). A randomised control trial to determine if use of the iresus© application on a smart phone improves the performance of an advanced life support provider in a simulated medical emergency. *Anaesthesia*, *66*(4), 255–262.

Material design. (2022). Retrieved April 10th 2022, from `https://m2.material.io/design`

McKenna, S. P., & Glendon, A. I. (1985). Occupational first aid training: Decay in cardiopulmonary resuscitation (cpr) skills. *Journal of occupational psychology*, *58*(2), 109–117.

Mickan, S., Tilson, J. K., Atherton, H., Roberts, N. W., Heneghan, C., et al. (2013). Evidence of effectiveness of health care professionals using handheld computers: a scoping review of systematic reviews. *Journal of medical Internet research*, *15*(10), e2530.

Mosa, A. S. M., Yoo, I., & Sheets, L. (2012). A systematic review of healthcare applications for smartphones. *BMC medical informatics and decision making*, *12*(1), 1–31.

Neves, D., & Anderson, J. R. (1981). *Knowledge compilation: mechanisms for the automatization of cognitive skills.* (Tech. Rep.). CARNEGIE-MELLON UNIV PITTSBURGH PA DEPT OF PSYCHOLOGY.

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. In *Proceedings of the interact’93 and chi’93 conference on human factors in computing systems* (pp. 206–213).

Pavlik, P. I., & Anderson, J. R. (2003). An act-r model of the spacing effect.

Pavlik, P. I., & Anderson, J. R. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive science*, *29*(4), 559–586.

Petrosoniak, A., Ryzynski, A., Lebovic, G., & Woolfrey, K. (2017, April). Cricothyroidotomy in situ simulation curriculum (CRIC study). *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, *12*(2), 76–82. doi: 10.1097/sih.0000000000000206

Rasmussen, J. (1987). *Information processing and human-machine interaction. an approach to cognitive engineering*. North-Holland.

Salina, L., Ruffinengo, C., Garrino, L., Massariello, P., Charrier, L., Martin, B., ... Dimonte, V. (2012). Effectiveness of an educational video as an instrument to refresh and reinforce the learning of a nursing technique: a randomized controlled trial. *Perspectives on medical education*, *1*(2), 67–75.

Scrase, I., & Woollard, M. (2006). Needle vs surgical cricothyroidotomy: a short cut to effective ventilation. *Anaesthesia*, *61*(10), 962–974.

Shahsavari, H., Ghiyasvandian, S., Houser, M. L., Zakerimoghadam, M., Kermanshahi, S. S. N., & Torabi, S. (2017). Effect of a clinical skills refresher course on the clinical performance, anxiety and self-efficacy of the final year undergraduate nursing students. *Nurse Education in Practice*, *27*, 151–156.

Shea, C. H., Lai, Q., Black, C., & Park, J. (2000). Spacing practice sessions across days benefits the learning of motor skills. *Human movement science*, *19*(5), 737–760.

Sullivan, A., Elshenawy, S., Ades, A., & Sawyer, T. (2019). Acquiring and maintaining technical skills using simulation: initial, maintenance, booster, and refresher training. *Cureus*, *11*(9).

System usability scale voor meten gebruiksvriendelijkheid. (2021). Retrieved from `https://www.usersense.nl/usability-testing/system-usability-scale-sus`

Taatgen, N. (1999). The atomic components of thought: by john r. anderson and christian lebiere. *Trends in Cognitive Sciences*, *3*(2), 82.

Testflight. apple developer. (2022). Retrieved from `https://developer.apple.com/testflight/`

Van Det, M., Meijerink, W., Hoff, C., Middel, L., Koopal, S., & Pierie, J. (2011). The learning effect of intraoperative video-enhanced surgical procedure training. *Surgical endoscopy*, *25*(7), 2261–2267.

VanLehn, K. (1996). Cognitive skill acquisition. *Annual review of psychology*, *47*(1), 513–539.

Van Rijn, H. (2010). Slimstampen. nl: Optimaal leren van feitelijke informatie in nederlands als tweede taal (nt2) lessen door individuele aanpassing van leerregimes. final report to kennisnet.

Van Rijn, H., van Maanen, L., & van Woudenberg, M. (2009). Passing the test: Improving learning gains by balancing spacing and testing effects. In *Proceedings of the 9th international conference of cognitive modeling* (Vol. 2, pp. 7–6).

Vaziri, D. D., Aal, K., Ogonowski, C., Von Rekowski, T., Kroll, M., Marston, H. R., … others (2016). Exploring user experience and technology acceptance for a fall prevention system: results from a randomized clinical trial and a living lab. *European Review of Aging and Physical Activity*, *13*(1), 1–9.

Ventinova Medical. (2022). *Ventrain kit.* Retrieved from `[https://www.ventinovamedical.com/cricath/]`

Ventola, C. L. (2014). Mobile devices and apps for health care professionals: uses and benefits. *Pharmacy and Therapeutics*, *39*(5), 356.

Wallace, S., Clark, M., & White, J. (2012). 'it's on my iphone': attitudes to the use of mobile computing devices in medical education, a mixed-methods study. *BMJ open*, *2*(4), e001099.

Walls, R. M. (1988). Cricothyroidotomy. *Emergency medicine clinics of North America*, *6*(4), 725–736.

Whitehill, J. (2013). Understanding act-r-an outsider's perspective. *arXiv preprint arXiv:1306.0125*.

Woollard, M., Whitfield, R., Newcombe, R., Colquhoun, M., Vetter, N., & Chamberlain, D. (2006). Optimal refresher training intervals for aed and cpr skills: a randomised controlled trial. *Resuscitation*, *71*(2), 237–247.

Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker Jr, J. F. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & management*, *43*(1), 15–27.

A Questions of interactive videos

The bold answers represent the correct answers. *Vraag 1. Moet de naald craniaal of caudaal gericht worden bij het inbrengen?'*
Question 1. Should the needle be inserted cranial or caudal ?
a. Craniaal
a. Cranial
b. **Caudaal**
b. **Caudal**

Vraag 2. Hoe weet je dat je de naald diep genoeg ingebracht hebt?
Question 2. How do you know that you have inserted the needle deep enough? a. Je voelt weerstand
a. You feel resistance
b. **Er komt lucht in de naald**
b. **Air flows into the needle**
c. De naald zit 2cm diep
c. The needle is 2 cm deep

Vraag 3. Waar moet je op letten bij het verwijderen van de naald?
Question 3. When removing the needle, what should you pay attention to?
a. **Katheter op zijn plek houden**
a. **Keep the catheter in place**
b. De naald op 45° houden
b. Keep the needle in a 45 degree angle

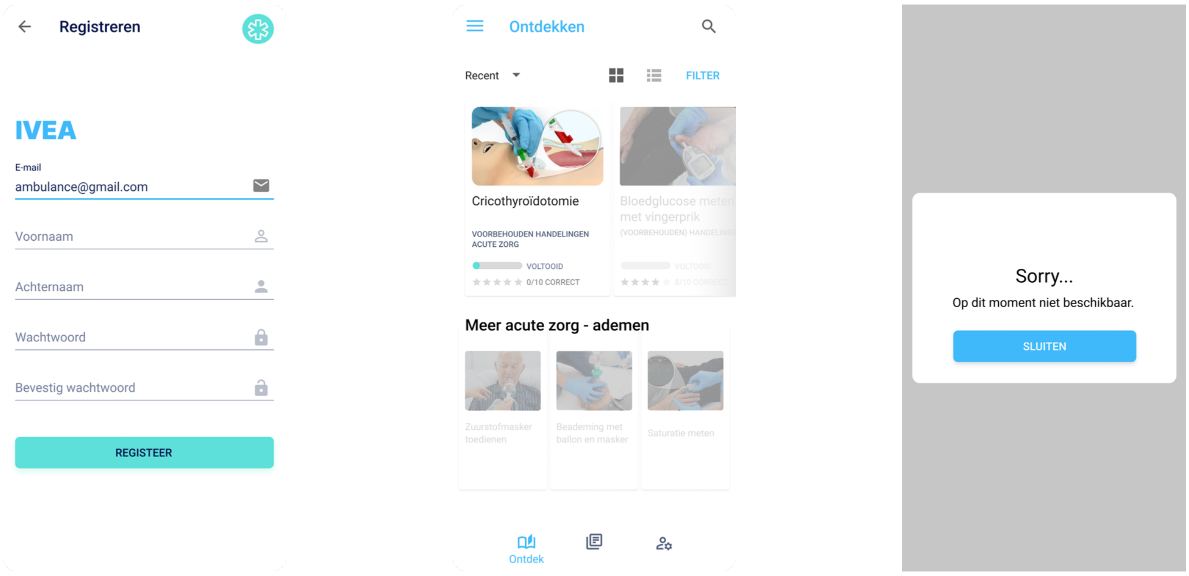
Vraag 4. Met welke vingers immobilizeer je de larynx?
Question 4. With what fingers do you immobilize the larynx?
a. **Duim en middelvinger**
a. **Thumb and middle finger**
b. Wijsvinger en middelvinger
b. Index finger and middle finger
c. Wijsvinger en ringvinger
c. Index finger and ring finger

Vraag 5. Welke locatie moet je palperen?
Question 5. What location should you palpate?
a. **Adamsappel**
a. **Adams apple**
b. Trachea
b. Trachea

Vraag 6. Op welke hoek breng je de naald in?
Question 6. At what angle do you insert the needle?
a. 90°
b. 45°
c. **30°**

B Walkthrough of Figma designs

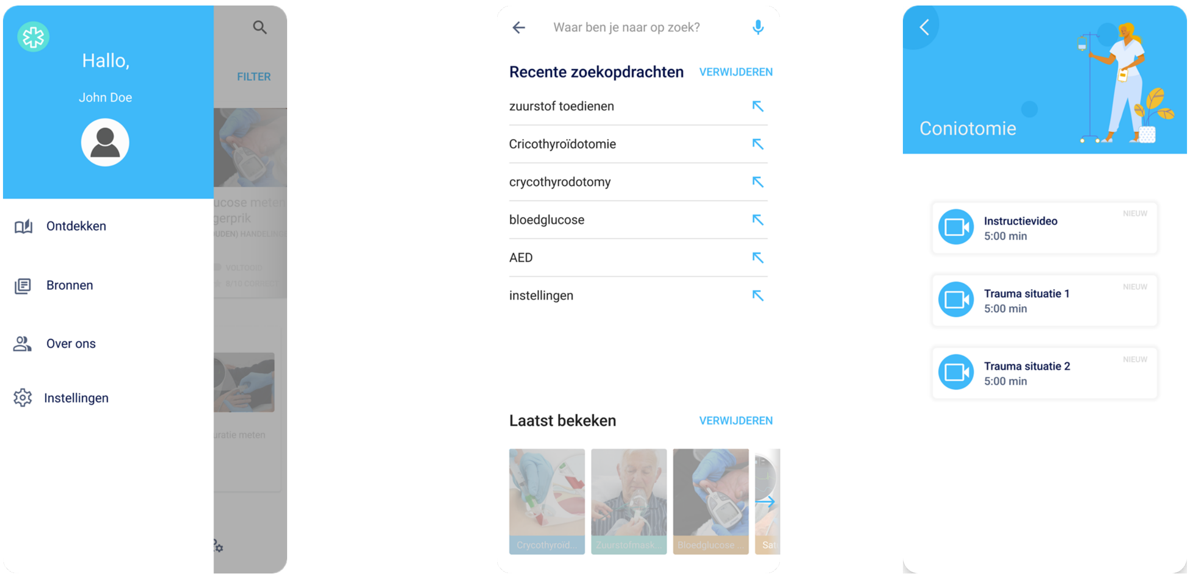
Introduction and Home



(a) Login screen (b) Homepage (c) Not available screen

Figure 23: Login, homepage and not available screens.

Information



(a) Hamburger menu 3.5cm (b) Search menu (c) Video page

Figure 24: Hamburger menu, search menu and video page.

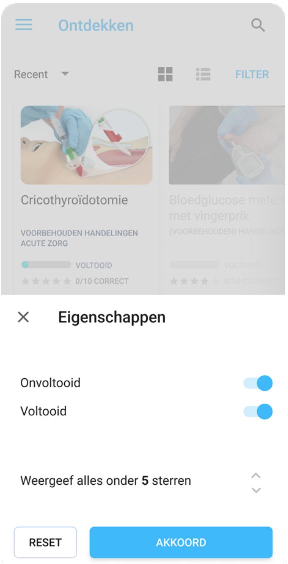
Videos

(a) Video

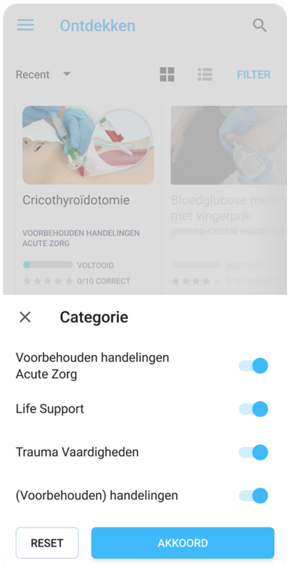
(b) Video question

Figure 25: Interactive video player.

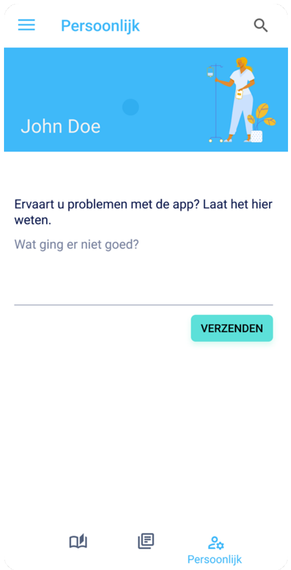
Personalisation and contact



(a) Properties



(b) Category



(c) Personal/contact

Figure 26: Personalisation and contact screens.

C MoSCoW table

Must have	Should have	Could have	Won't have
Home page: Clickable tiles that lead to practice pages of specific procedures	**Video practice/ practice page:** Using the number of correct answers to set the score and progress	**Overview page**	**Animations:** Instructional videos created with animations
Practice page: Only for Coniotomie page; containing different clickable video practices	**Practice page:** Also a practice page for all other procedures	**Filter function**	**Video page:** Other question types aside from multiple choice (i.e. drag and drop or draw)
Video practice: Interactive video with corresponding questions	**Video page:** Adjustable audio volume level	**Video page:** Subtitles for accessibility	**Accessibility and responsiveness:** Light & mode
Video practice: At least 3 videos	**Responsiveness:** Possibility to turn screen	**Search function**	**Responsiveness/ accessibility:** Possibility to use on tablet vs. phone
Responsiveness/ accessibility: Also available for iPhone	**Personal page:** Feedback field		
Login screen: Link to database (so that progress stays up to date)	**Personalized feedback**		
Navigation: Hamburger menu that directs to less frequently used pages (sources, additional information about the app)	**Sources:** Page that displays sources of videos and information in		

Table 2: MoSCow Table

D Source code

The source code of the IVEA application can be found in the following GitHub repository:
https://github.com/floorkuipers/ambu.git

E Information for participants

E.1 General information e-mail (in Dutch)

Geachte meneer/mevrouw,

Middels deze mail en bijgevoegde informatiebrief willen we u vragen of u wilt meedoen aan wetenschappelijk onderzoek. Meedoen is vrijwillig. U ontvangt deze mail omdat u werkzaam bent als ambulance-verpleegkundige op AMBULANCEPOST op DATUM. Mijn naam is Floor Kuipers, ik ben een master student aan de RUG (Computational Cognitive Science). Ik ben op dit moment mijn scriptie aan het afronden, waarvoor ik een experiment uit kom voeren op uw ambulancepost. Mijn onderzoek draait om het onderhouden van vaardigheden die bijna nooit uitgevoerd moeten worden als ambulanceverpleegkundige.
Op DATUM bezoek ik de ambulancepost, met een fantoom. Als deelnemer zult u een zeldzame verrichting uitvoeren op deze fantoom (de precieze verrichting wordt verder toegelicht na toezegging van deelname aan het experiment). Deze oefening wordt opgenomen ter analyse. Ter voorbereiding op dit oefenmoment neem ik twee weken voor DATUM contact met u op over de benodigde voorbereiding.
Alvast heel erg bedankt voor de tijd, moeite en (hopelijke) deelname.
Voor eventuele vragen ben ik bereikbaar via mijn e-mailadres: f.t.a.kuipers@student.rug.nl.

Met vriendelijke groet, Floor Kuipers

E.2 Extensive experiment information (in Dutch)

Doel van het onderzoek
U als ambulanceverpleegkundige moet gereed zijn om op ieder moment veel verschillende verrichtingen te kunnen uitvoeren. Niet al deze verrichtingen komen echter vaak voor. Zoals met alle kennis die niet vaak toegepast hoeft te worden, vergaat deze over tijd. Omdat het in veel situaties echter essentieel is dat u de nodige kennis meteen paraat heeft, is het belangrijk om deze kennis te onderhouden. Mijn onderzoek houdt zich dus bezig met het verval van benodigde kennis van een zeldzame verrichting en hoe dit verval geminimaliseerd kan worden.

Beschrijving van het experiment en verwachtingen van deelnemers
Om mijn onderzoeksvraag te beantwoorden voer ik een experiment uit. Op DATUM bezoek ik ambulancepost LOCATIE. Op het meegenomen fantoom wordt door u een zeldzame verrichting uitgevoerd. De precieze informatie over de specifieke verrichting e.d. wordt na toezegging van deelname verstrekt. Deze uitoefening wordt opgenomen (middels een videobril) om later te analyseren. Twee weken voor deze oefening ontvangt u van mij de informatie over uw benodigde voorbereiding op de oefening. Deelname en beëindiging De deelname aan dit onderzoek vindt plaats op vrijwillige basis. Uw deelname kan helpen met het krijgen van inzicht in mogelijke nieuwe mediatoepassingen bij het ondersteunen van vaardigheden van (onder andere) ambulance verpleegkundigen. U kunt weigeren om deel te nemen aan het onderzoek en u kunt zich op elk moment terugtrekken uit het onderzoek

zonder dat u hiervoor een reden hoeft op te geven.

Risico's, voordelen en kosten
Er zijn geen risico's verbonden aan deelname aan dit onderzoek. Voordelen die verbonden zijn aan dit onderzoek zijn het opfrissen van de nodige kennis van een zeldzame verrichting en het bijdragen aan potentieel vernieuwend onderwijs. Er is geen vergoeding verbonden aan deelname aan dit onderzoek.

Vertrouwelijkheid
Tijdens de voorbereiding op en tijdens het experiment worden gegevens verzameld die nodig zijn voor de analyse van de resultaten van het experiment. Deze gegevens worden geanonimiseerd bewaard en zijn niet te herleiden naar uw persoonsgegevens. Na de uitvoering van het experiment worden de gegevens verwijderd uit de database en alleen nog lokaal geanalyseerd.

Contactpersoon
Als er schade optreedt ten gevolge van het onderzoek, of als u aanvullende informatie wenst over het onderzoek of over uw rechten en plichten, kunt u in de loop van het onderzoek op elk ogenblik contact opnemen met: Naam: Floor Kuipers E-mail: f.t.a.kuipers@student.rug.nl

F Usability questionnaire

Wat is uw leeftijd? *

Jouw antwoord

Ik denk dat ik dit product frequent zou willen gebruiken. *

Sterk mee oneens 1 2 3 4 5 Sterk mee eens

Ik vond het onnodig ingewikkeld. *

Sterk mee oneens 1 2 3 4 5 Sterk mee eens

Ik vond het product makkelijk te gebruiken. *

Sterk mee oneens 1 2 3 4 5 Sterk mee eens

Ik denk dat ik technische support nodig heb om het product te gebruiken. *

Sterk mee oneens 1 2 3 4 5 Sterk mee eens

5. Ik vond de verschillende functies van het product goed met elkaar geïntegreerd *

Sterk mee oneens 1 2 3 4 5 Sterk mee eens

Ik vond dat er te veel tegenstrijdigheden in het product zaten. *

Sterk mee oneens 1 2 3 4 5 Sterk mee eens

Ik kan me voorstellen dat de meeste mensen snel met het product overweg kunnen. *

Sterk mee oneens 1 2 3 4 5 Sterk mee eens

Ik vond het product omslachtig in gebruik. *

Sterk mee oneens 1 2 3 4 5 Sterk mee eens

Ik voelde me zelfverzekerd tijdens het gebruik van het product. *

Sterk mee oneens 1 2 3 4 5 Sterk mee eens

Ik moest veel over het product leren voordat ik het goed kon gebruiken. *

Sterk mee oneens 1 2 3 4 5 Sterk mee eens

Hoe vaak heeft u de app gebruikt? *

- 1 keer
- 1 dag
- Meerdere dagen
- Option 4

Op welke manier heeft u zich, naast het gebruik van de app, voorbereid op de evaluatie? *

Jouw antwoord

Heeft u nog overige opmerkingen over de app of ervaringen die u wilt delen?

Jouw antwoord

G Synthesis of expert evaluations

Participant 1
Summary:
- The overall procedure takes too long
- The procedure was not performed in a sterile manner

Palpation
- The palpation is not efficient, the larynx is not fixated
- Wrong insertion position is used. The participant relies more on the available needle holes than his own reasoning

Needle insertion
+ The direction of the needle is correct (caudal)
- Angle of insertion is incorrect

Raise of the needle
- Participant already raises the needle before loss of resistance is felt
- The needle is raised too much, making it impossible for the catheter to move
- No air is aspirated after insertion

Participant 2
Summary:
- The procedure takes a long time
- The procedure was not performed in a sterile manner

Palpation
+ Palpation is at the correct position
- The palpation is not efficient
- Trachea is not fixated

Needle Insertion
- Angle of insertion also seems correct, but after insertion in the skin the participant changes the angle of insertion. Because of this, the insertion fails in this attempt. The participant moves the needle to 30 degrees too soon

Raise of the needle
+ Aspirates after insertion
- The participant has not yet experienced loss of resistance. The needle is raised too soon. Because of this, the ‘venflon’ bent
- Enters the final part of the cricath with the needle already fully removed

Participant 3
Summary:
+ Fast and correct
+ Stands next to the patient
+ Sterile procedure

Palpation
+ Good technique of palpation and determination of correct position

Needle insertion
+ Angle of insertion is correct

Raise of the needle
+ Aspirates after insertion
+ Raises catheter in the correct manner

Participant 4
Summary:
+ Thorough preparation
- Procedure is not performed in a sterile manner

Palpation
+ Palpation is at the correct position
- The palpation is not efficient
- Larynx is not fixated

Needle insertion
+ Angle of insertion is correct

Raise of the needle
+ Aspirates after insertion
+ Raises the cannula correctly
+ Mentions the loss of resistance

Participant 5
Summary:
+ Fast and correct
+ Explains reasoning of standing behind patient
- Procedure is not performed in a sterile manner
Palpation
+ Palpation is at the correct position
- Larynx is not fixated

Needle insertion
- Angle of insertion is not mentioned

Raise of the needle
+ Aspirates after insertion
- Mentions that they raise the needle ‘until they feel resistance’ – they probably mean ‘until they lose resistance’.

Participant 6
Summary:
+ Fast
+ Good preparation
+ Stands on the side of the patient
+ Procedure is performed as sterile as possible

Palpation
+ Palpation is at the correct position
+ Larynx is fixated

Needle insertion
- Mentions different degree of insertion than is executed

Raise of the needle
+ Aspirates after insertion
+ Raises catheter in the correct manner

Participant 7
Summary:
+ Excellent

Palpation
+ Palpation is at the correct position
+ Larynx is fixated

Needle insertion
+ Insertion is correctly done
+ Angle of insertion is correct

Raise of the needle
+ Aspirates after insertion
+ Raises catheter in the correct manner