Checklist completed?

An examination of checklist design in shipping

Diploma thesis in the Master Mariner Programme

ANDERS ÖHRN

DAVID HEDLUND

Department of Shipping and marine technology
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2014

REPORT NO. SK-14/168

Checklist completed?

An examination of checklist design in shipping

ANDERS ÖHRN

DAVID HEDLUND

Department of Shipping and Marine Technology

CHALMERS UNIVERSITY OF TECHNOLOGY

Gothenburg, Sweden, 2014

Checklist completed?
An examination of checklist design in shipping

ANDERS ÖHRN
DAVID HEDLUND

Report no. SK-14/168
Department of Shipping and Marine Technology
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Printed by Chalmers
Gothenburg, Sweden, 2014

Checklist completed?
An examination of checklist design in shipping

Anders Öhrn
David Hedlund
Department of Shipping and Marine Technology
Chalmers University of Technology

Abstract

The checklist is a widely used tool for improving safety in aviation and maritime operations as well as improving surgical care. In aviation checklists have been used with great success for decades. While they have become more common in the maritime industry as well, there is still some resistance towards the procedure. The purpose of this review was to find whether or not there is a systematic approach to creating checklists in shipping or in other fields, and if there is knowledge in other fields that could be utilized to improve the design of maritime checklists.

To answer these questions, a systematic literature study was conducted. It investigated earlier research related to checklists that have been made in various areas. To make the study as systematic as possible the PRISMA method was used combined with conversations with certain persons with insight in the shipping industry and how checklists implemented.

The review did not find any research published on maritime checklists, however, that does not disprove that there is a systematic approach to the design of checklists in shipping. It did find a significant amount of research concerning aviation and medicine, indicating that there is a systematic approach in those fields. Design recommendations based on human factors research were presented.

Keywords: Checklist, safety, aviation, medicine, surgery, shipping, maritime, marine, design, layout

Sammanfattning

Checklistor används idag inom många olika områden, allt från att förbättra säkerheten inom sjöfart och luftfart till att få operationsrutiner inom kirurgi att fungera. Inom luftfarten har checklistor använts med stor framgång i årtionden. Även trots att de har blivit vanligare inom den maritima industrin, finns fortfarande ett visst motstånd mot användandet av checklistor. Syftet med denna studie var att finna huruvida det finns ett systematiskt tillvägagångssätt för att skapa checklistor inom sjöfarten eller inom andra områden, och om det finns kunskap inom andra områden som skulle gå att implementera för att förbättra utformningen av maritima checklistor.

För att kunna svara på frågorna gjordes en systematisk litteraturstudie. Genom den undersöktes tidigare forskning relaterad till checklistor som gjorts inom olika områden. För att göra studien så systematisk som möjligt användes PRISMA-metoden, i kombination med kontakt med särskilda personer som har insikt inom sjöfart och hur checklistor implementeras.

Studien har inte hittat någon forskning som publicerats om checklistor inom sjöfarten, dock bevisar inte det att det finns inte ett systematiskt tillvägagångssätt för utformningen av checklistor inom sjöfarten. Den hittade en betydande mängd forskning kring luftfart och medicin, vilket tyder på att det finns en systematik i dessa områden. Rekommendationer för design av checklistor baserade på forskning presenterades.

Nyckelord: Checklistor, säkerhet, flyg, medicin, kirurgi, sjöfart, maritim, marin, design, layout

Acknowledgments

Firstly, the authors would like to express their sincere gratitude to their advisor, Lars Axvi, for his support and guidance throughout the whole process.

Additionally, the authors wish to thank Mr John Murray, Marine Director at the International Chamber of Shipping, and Captain Hans Hederström, Managing Director at CSMART, for their correspondence.

Table of contents

List of figures

List of tables

Definitions

ICS	**“**International Chamber of Shipping” is a voluntary organisation for ship owners around the world. They are the principal international trade association for the shipping industry. They are also the author of the Bridge Procedures Guide.
IMO	**“**International Maritime Organisation” is an organisation under the United Nations working with important things regarding pollution, safety and navigation at sea.
FAA	**“**Federal Aviation Administration” is the national aviation authority of the United States. FAA is a part of the U.S Department of Transportation. It regulates and oversees all American civil aviation.
OCIMF	**“**Oil Companies International Marine Forum” is a voluntary organisation of oil companies whose interests lay in shipping and stocking different kinds of oil and gas products.
WHO	**“**World Health Organisation” is a special organisation under the United Nations and their work concerns international public health.
BPG	**“**Bridge Procedure Guide” is published by ICS and is considered a best practice for watch keeping in the merchant fleet. It follows the requirements of IMO’s STCW convention.
ISGOTT	“International Safety guide for oil Tankers and Terminals” is a guide for handling oil products in a safe way. It is published by ICS and OCIMF.
ISM	**“**International Safety Management” system is published by IMO and its purpose is to prevent pollution from ship operations and set an international standard for safety at sea.
MARPOL	**“**International Convention for the Prevention of Pollution from Ships” is the main convention for prevention of pollution at sea.

SMS	**"**Safety Management System" is a mandatory part of the ISM code that requires every shipping company to develop an SMS manual which should contain instructions and procedures to ensure safe operation of ships and protection of the environment.
SOLAS	**"**Safety Of Life At Sea" is a convention that stipulates the minimum requirements for safety equipment and how ships should be constructed. The convention was formed after the sinking of the RMS Titanic.
ICU	**"**Intensive Care Unit", is a special department in a hospital that provides intensive care for patients in critical condition.

1 Introduction

On board most ships today there are checklists for most emergencies, but also for regular operations like departures, arrivals, loading or unloading. The Federal Aviation Administration (2007) defines a checklist as “a formal list used to identify, schedule, compare, or verify a group of elements or actions. A checklist is used as a visual or audible aid that helps the user overcome the limitations of short-term human memory …”. In simple terms, a checklist is a memory aid designed to reduce human error.

Today checklists are used in many different fields involving complex procedures, including aviation, product manufacturing, medicine and shipping, and have been shown to significantly increase safety when properly used (Hales & Pronovost, 2006; Weiser et al., 2010).

However, there is some resistance in shipping towards written procedures. Reliance on documents such as checklists and risk assessments are thought by some to counteract the use of common sense, or what is known by sailors as “seamanship” (Knudsen, 2009). In other fields dealing with complex operations, such as aviation and medicine, checklists are a central part of the everyday routine, preventing simple mistakes caused by human error (Hales & Pronovost, 2006). Although checklists are very common on ships, they are not always properly used. For example, in the case of the grounding of M/V Maersk Kendal, a contributing factor may have been the improper voyage planning prior to starting the voyage. There was a checklist for voyage planning available on board and filled out for the voyage, but all the steps had not actually been carried out (Marine Accident Investigation Branch, 2009). An example of a checklist for creating a voyage plan can be found in Appendix 1.

For a checklist to be effective the contents must be relevant to the task, and the design must be user friendly (Degani & Weiner, 1990). It is crucial that the user feels that the checklist is not just another procedure, but instead a tool at his disposal (Knudsen, 2009). The purpose of the checklist should be to free the mind from having to think about the things that are routine, leaving more room to think about things that are not.

We believe that the first step towards changing the general opinion of checklists at sea is to provide checklists that are properly designed to meet the needs of the sailor. In order to find out what that means we have conducted a systematic literature review, focusing on research made on the subject in aviation, medicine and shipping.

1.1 Purpose

To examine the design of checklists from a scientific perspective, on board ships as well as in other fields, in order to find out if any improvements can be made to the way checklists are designed for ships today.

1.2 Questions

- Is the design of checklists on board ships today based on a systematic approach?
- Is there a systematic approach behind the design of checklists used in other fields that could be utilized to make improvements to checklists on ships?
- If so, on what principles is that approach based?
- What are the design criteria proposed by the current research?

1.3 Delimitations

The thesis does not examine whether or not checklists aid or impair safety, and merely focuses on the design of checklists as they are today. It has been quite extensively researched and proven that checklists aid safety in other fields where they are used, and so for the purposes of this thesis we assume that this will be the case for shipping as well (Weiser et al., 2010).

The thesis examines checklist design under the assumption that they are actually being used. There are several reasons for checklists not being used. For example the checklist may be skipped or cut short because the user got distracted by other event, or because other matters seem more pressing (Boorman, 2001).

The thesis does not test the viability of any alterations we suggest. Testing and evaluating will be beyond the scope of this review.

The thesis focuses mainly on paper checklists, but includes other types that are available today briefly for awareness.

The thesis focuses on procedural checklists, but includes other types briefly for awareness.

2 Method

The examination was started by defining the purpose and research questions. To make sure that the result was based on scientific data, a systematic qualitative literature review method was used. In addition to that, e-mail conversations with specifically selected persons provided good insights into the shipping community.

In order to be certain that most or all of the relevant literature is found, a good and well elaborated method is required. A systematic literature review following the guidelines outlined by the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) was conducted. The advantage of the PRISMA method is that it provides clear guidelines and steps for the process, which makes sure that the collection, selection and implementation of information is systematic and thorough throughout the thesis (Moher, 2009).

The original guidelines were slightly modified to fit our purpose. While the data collection process was performed according to the guidelines of the method, it was not documented according to the specifications because of lack of time. Some parts of the method were not applicable to this thesis, and were not included, such as “additional analysis” (described in the PRISMA guidelines) of the collected data (Moher, 2009).

To find, select and verify the validity of the research we followed a specific procedure described below.

First, sources were identified using different databases. The criteria for selection was set to only include studies that: (1) examined checklist design or the use of checklists from a design perspective, and (2) were made with aviation, medicine or shipping in mind and (3) were published in English. During the search process no restriction was put on the publication date of studies. Studies were screened by reading titles and abstracts, and only those matching the criteria were included.

A combination of the following keywords were used: Checklist, checklists, safety, aviation, pilot, pilots, medicine, surgery, surgical, shipping, maritime, marine, design, designing, layout, human error. To find as many relevant studies as possible different combinations of search terms were used depending on the database and its requirements and limitations. In addition to this, the reference lists of relevant studies were used to find additional sources.

The following electronic databases were used: Scopus, Web of Science, Google and Google Scholar. Searches were conducted from the beginning of September until the beginning of October 2014.

The selected material was read in full text and some additional studies were excluded based on lack of relevance or eligibility. In this second examination process the selection was narrowed further by strictly enforcing the same criteria used when searching the databases. During this process any duplicates were excluded. Those that matched the criteria were then included in our reference list. In the end a total of 17 relevant studies were included.

In order to categorize the selected studies, a reference material database was created. In the database information such as the study’s name, authors, publication journal and year or publication was recorded. Studies were rated by relevance to further illustrate their respective pertinence to the subject. Each study was checked for validity and reliability by verifying the credibility of the author and were the material was published. Cross-references were used when possible in order to verify the information further. Studies that were not deemed to meet the requirements in this regard were used exclusively for cross-referencing (Denscombe, 2009).

In some cases the results of studies were conflicting. In those cases attempts were made to find cross-referencing information. If no cross-referencing information could be found, the credibility of the author was the deciding factor.

3 Background

The first checklists were created as a solution to a problem. In 1935 the US Army held a competition for airplane manufacturers to decide who was going to build their next-generation bombers. The Boeing Corporation's Model 299, more commonly known as the B-17 Flying Fortress, was thought to be the only contender for the first price, being that it was significantly more advanced than all the other aircraft that entered the contest (Meilinger, 2004).

As it turned out, it was better in all ways but one. Shortly after taking off the plane stalled and subsequently crashed. An investigation into the cause of the accident revealed that human error was to blame. The plane was so advanced that the pilot, with his limited cognitive capabilities, had forgotten one vital step in configuring the aircraft. It was deemed "too much airplane for one man to fly". Even though the Model 299 did not win the competition the army still purchased a few of them as test planes. Some pilots remained convinced that the airplane was not impossible to fly, and began working on a solution to overcome the overwhelming complexity of operating this machine. Their solution did not lie in changing the way pilots trained. Their solution was a checklist (Meilinger, 2004).

The American army realized the B-17 bombers' great value during the Second World War and began cooperating with Boeing to build more planes (Gawande, 2009). The Boeing factories built in total 6'981 B-17 bombers in different models, and additional 5'745 bombers together with Douglas and Lockheed in an international cooperation. The B-17 started a long and successful career for the US Army and aided the allied forces greatly in their devastating bombings over Nazi-Germany (Boeing, 2014).

3.1 Checklists in medicine

In medicine checklists has started to become an important tool to help surgeons and nurses. In western hospitals as early as the nineteen-sixties, nurses redesigned their patient charts and forms to include the four vital signs, essentially making it a checklist for themselves. Even with the overwhelming amount of tasks each nurse had to perform during her shift, this "vital chart" was there to prevent neglect and remind them to do the most important thing – make sure their patient was alive. This made sure that every six hours, or more often when required, the nurses would not forget to check their patient's pulse, blood pressure, temperature, and respiration to see exactly how the patient was doing (Gawande, 2009).

At the Johns Hopkins Hospital in Baltimore a critical care specialist named Peter Pronovost gave checklists a try in 2001. The idea was not to make a checklist that covered every procedure performed in a hospital, but to try and reduce the risks associated with one of them. The objective of the checklist was to prevent central line infections. He wrote down five steps for the staff to perform before starting the procedure. They were:

1. Wash their hands with soap
2. Clean the patient's skin with chlorhexidine antiseptic
3. Put sterile drapes over the entire patient
4. Wear a mask, sterile gown, and gloves
5. Put a sterile dressing over the insertion site once the line is in

This was a very simple checklist, and things like this were supposed to be common practice. Still, when Pronovost asked a nurse in his ICU team to observe the doctors and their routines during one month before the list was in use. Their findings showed undeniable proof that there was a problem. In more than one third of all cases the doctors skipped at least one step (Gawande, 2009).

The situation was so unfamiliar in medicine that Pronovost realized there was a need for change. He asked the hospital administration to authorize nurses to stop a doctor immediately if they skipped a step in the procedure. With this change in place, Pronovost and his colleagues observed and documented all cases of line-infections in the hospital during one year. The result was stunning. The rate of line-infections went down from eleven to zero percent. During this short observation period, this checklist had prevented at least forty-three infections and eight deaths, saving around two million in extra costs, in one hospital alone (Gawande, 2009).

Today the use of checklists has become widespread in the field of medicine. In 2006 the World Health Organization became aware of the high mortality rate during surgical care and after treatment. In 2004 there were 230 million surgeries being performed worldwide. They needed a solution to lower those figures that would work in every part of the world, the rich as well as the poor. Dr. Atul Gawande's idea was to create a checklist to aid surgical teams that would improve care before and after surgeries. They turned their focus to the aviation industry where checklists had been used for a long time. Gawande came in contact with Daniel Boorman, a veteran pilot who had been developing checklists for the Boeing Company for more than twenty years. They used knowledge from several different fields in order to find a type of design for their checklist that would suit the workflow of the surgical theatre. The finished checklist was tested in Boston during a simulated surgery with a complete surgical team (Gawande, 2009).

The checklist went out for trials in eight different hospitals all around the world, in rich countries as well as in poor. When the trial was a success, hospitals around the world started using the checklist on their own. The reactions amongst the surgeons and medical staff were mixed. At first there was a lot of resistance to it, but after a while many changed their minds.

A survey consisting of over 250 anonymous participants from different medical professions that was made some months after the checklist was introduced showed over 80 percent were positive towards it and now found it to be a resource rather than an obstacle. Still there were 20 percent that did not like the idea. Still, when asked if they would like the checklist to be used if the surgery was to be performed on them, 93 percent answered yes. Today the checklist is used in many hospitals all around the world (Gawande, 2009). The published study of the implementation of the checklist showed a decrease in death during surgery from 1.5% to 0.8%, meaning thousands of lives saved every year (Ziewacz et al., 2012).

Nevertheless, according to Hales & Pronovost (2006) there is still some antagonism towards the introduction of checklist procedures. In a similar way to the opinion of many seafarers, many medical practitioners feel that the use of checklists is a sign of weakness and lack of knowledge, as well as a limitation on their use of common sense and judgment. They argue that this is probably the reason why these checklists are still not present in every hospital today (Hales & Pronovost, 2006; Knudsen, 2009).

3.2 Checklists in shipping

The checklist in shipping, just like in many other sectors, is there to prevent and prevail emergency situations. One such event was the capsizing of the English ferry M/S Herald of Free Enterprises. She departed on the 6th March 1987 from the inner harbor of Zeebrügge. The weather was calm with only a light breeze from the east. At 18.24 she passed the outer mole, and capsized four minutes later. She came to rest lying on her port side on a shallow sandbank, which prevented her from sinking completely and left the starboard side above the surface (Great Britain Department of Transport, 1987).

The investigation showed that her bow door had been unintentionally left open and this caused her car deck to be flooded. It also mentioned several critical points that might have triggered the accident. One of them was that there was no clear routine in how to make sure the stern door was closed before departure. Before the accident this had happened on no less than five occasions without incident. In the “ship standing orders” issued by the company, there was no reference to the opening and closing of the bow and stern door. This is an accident that could probably have been prevented by using a checklist (Great Britain Department of Transport, 1987).

In shipping today there are several different regulations that require the use of checklists or similar documents. The International Maritime Organization (IMO) has published the International Safety Management (ISM) code that every member state has to follow. The purpose of the ISM code is to prevent pollution from ship operations and set an international standard for safety at sea. The ISM code requires that ship operators provide procedures for safe practices in ship operations, a safe working environment, and establish safeguards against all identified risks. It also requires that they continuously improve the safety management skills of personnel ashore and on board ships, including preparing for emergencies related to safety and environmental protection. The code requires a shipping company to develop a Safety Management System (SMS) which should contain special requirements such as instructions and procedures to ensure safe operation of ships and protection of the environment. Additionally, recommendations by ISGOTT require that tankers to which the ISM code does not apply provide a system of equivalent standard of safe operations (International Chamber of Shipping, Oil Companies International Marine Forum & International Association of Ports and Harbors, 2006; International Maritime Organization, 2010).

In the maritime sector one of the main developers of checklists is the International Chamber of Shipping. They are the principal international trade association for the shipping industry, and their membership states add up to 80% of the world's merchant tonnage. This means that they have a great influence in the shipping community. For our purpose, their Bridge Procedures Guide (BPG) is the most interesting of their publications. Established as the best practice manual for watch keeping, it includes several example checklists that may be used, in original or modified versions, by shipping companies (International Chamber of Shipping, 2013).

3.3 Checklist design

The airline industry has been developing, using and improving their checklists for over 70 years. Despite the simplicity of the concept, there are a lot of aspects that should be considered when making sure a checklist is as optimally designed as possible. Even though not every aspect is of equal importance, everything from typography to the selection of items chosen has some manner of effect on the checklists performance. Most of the research on checklists has been made with aviation in mind, however in recent years checklists have become a hot topic in medicine. Nevertheless, by drawing from those years of researching, evaluating and modifying, it may be possible to apply some of that knowledge in other areas as well (Burain, 2006; Weiser et al., 2010).

3.4 Paper checklists

The visual impression of a printed document may have a large impact on the user's willingness and motivation to use it. Some factors that affect this are the quality of the paper, the quality of the print, color, and even typography (Degani, 1992). Furthermore, the environment and conditions in which the checklist is to be used must be considered (Wilson, 2013). A paper checklist has some advantages and disadvantages compared to other types. They are not very expensive to produce, or update. They are easy to keep and stow away when the checklist is completed. However, they can be easily worn out and are easy to mark, and they can be removed from their normal location and hard to find. They can also be hard to read under poor lighting conditions (Gross, 1995).

Figure 1 Pilot using a laminated paper checklist in pre-flight check (Hedlund, 2014).

- ### Paper quality

According to Degani (1992), the thickness of the paper could have an effect on the legibility of print. If there is print on both sides of the checklist and the paper is thin, the text on the other side of the checklist may show through in some situations. If the paper is held between the eyes and a light source the light may also shine through making the list harder to read.

- ### Print quality

It is important to make sure that the print is clear and sharp. Each character must be easily distinguishable. A checklist that has been copied and re-copied too many times may lose quality of print and become harder to read (Degani, 1992).

- ### Surface glare

Many checklists are laminated in order to make them last longer. They are also harder to mark and they normally last much longer then paper checklists. Laminating it may also make it easier to stow. However, laminating the paper may increase the amount of light reflected by the paper. This may be a problem in conditions with low luminance levels or when reading under a flashlight or similar. Experiments have shown significantly slower reading speed when using a laminated paper with high reflection. 75 percent of subjects in the experiment preferred using a non-glossy paper. The laminated version is also more expensive than paper checklists and when needed they can be difficult to fold (Degani, 1992; Gross, 1995).Scroll checklists

3.5 Scroll checklist

A scroll checklist is a checklist that consists of a roll of paper connected to two wheels. The paper roll is located inside a small box with a window. The two wheels can be used to scroll the checklist up or down. On the window there is a marking line that acts as a pointer, which makes it easy to see what item you are currently on. This type of checklist is most commonly used in older military transport aircraft (Degani & Wiener, 1990; Gross, 1995).

One of the benefits to using scroll checklist is that they are easy to correct and make updates to. It is also easy to mark the tasks, it cannot be lost since it is mounted, and it shows a head-up picture. Since this type of checklist is of an older design, and due to the small size of the scroll-box and the text, it can be hard to read the text, especially if it is mounted far away from the user. Some of them are not lit up and can be hard to read in poor lightning conditions. Due to the simple design there is no electronic memory that can help the pilot go back and remember un-finished items. Still, many military pilots prefer this system (Degani & Wiener, 1990; Gross, 1995).

3.6 Vocal checklists

A vocal checklist is a preprogrammed device that generates an audible checklist for the pilot. On the device there is a switch that can be rotated depending on what type of checklist is needed. There is one button on the yoke for “acknowledge” and one button for “proceed” that generates the next point on the list. If the pilot proceeds to the next point and intentionally skips one item on the list, it will be moved to the bottom of the checklist and read again before the checklist is completed. This type of checklist, like the scroll checklist, is most common in the aviation industry (Degani & Wiener, 1990).

There is one disadvantage to the checklist being audible. When the pilot goes through the checklist there may be other communication in the cockpit or flight deck, and that communication can interfere with the run-through of the checklist. Moreover, the checklist could potentially interfere with important communication from the flight control tower or similar (Degani & Wiener, 1990).

3.7 Electronic and computer-based checklists

The aviation industry has been using electronic checklists since the 1980s, to various extent. Today they can be found in many different areas and in different formats. They may exist as anything from inbuilt devices in the cockpit on modern airplanes to handheld devices, or in laptops in a hospital (Boorman, 2001; Degani & Wiener, 1990).

Over time the checklists format has changed and in the aviation industry it has become increasingly common to have electronic checklists. In 1980 Rouse and Rouse began to experiment with an on-board based computer to present checklists. The computer showed the checklists that were stored in a database. When a task was completed it would be dimmed down to show the user is had been checked off. A second experiment by Rouse, Rouse and Hammer was conducted in 1982. This time they compared a computer based checklist with a paper checklist in a simulator. The result showed that the computer based checklist was lower in errors than the paper checklist, but the paper checklist was faster to perform. It is possible that training or a different layout of the computer's keyboard could have lowered the completion time (Degani & Wiener, 1990).

There are some advantages to using electronic checklists. It can be connected to a network, making it possible to synchronize it between different devices. The computer behind an electronic checklist usually has a great storage memory that can store and handle many checklists, and it is easy to configure and set-up. The user can easily choose which checklist it should show. On the other hand, this type of checklist often brings higher initial costs and there is always a risk that electronic devices stop working or has different kinds of technical problems (Verdaasdonk, Stassen, Widhiasmara & Dankelman, 2009).

In the early 1990s the flight company Boeing started developing what would become their first electronic checklist system for the Boeing 777. There were some requirements that had to be met. The checklists should be of help to the crew instead of distracting them. They should not be too hard to handle, and compared to paper checklists they had to be easy to use. After the introduction of electronic checklists the pilot errors were decreased by another 46% as compared to paper checklists (Degani & Wiener, 1990; Boorman, 2001).

Usually an electronic checklist in aviation is a display and pointer list. On the screen there is a cursor that can move around for the different tasks on the checklists. In some systems a task will be lit up when the cursor is moved over it. When an item is executed the task changes color to make it easier to overview which tasks have been completed. If points are skipped they will be highlighted to make sure they are not completely forgotten. One manufacturer decided to design their checklists so that it would be possible to switch between different checklists before they were completed. The color design on the screen can vary from different manufacturers (Degani & Wiener, 1990).

In aviation checklists are divided into normal and non-normal checklists. The normal checklist is used in everyday operations such as starting and landing. Non-normal checklists are used in emergencies or other extraordinary situations (Boorman, 2001)

Designing an HCI (Human-Computer Interface) including both normal and non-normal checklists may provide designers with a challenge, since non-normal checklist are often used in high-stress, high workload situations. One of the primary purposes for the electronic checklists was to provide the crew only with the most vital information in order to help reducing errors during the most critical phases of an emergency. Boeing used a human centered automation design, which made it error-tolerant, reliable and anticipating. Before it was introduced it was important for them to make sure that it did not introduce any new, previously unknown errors (Boorman, 2001).

Another common system on aircraft today is the Flight Management Computer (FMC). It features checklists that are partially automated, and includes both normal and non-normal checklists. Automating parts of the checklists relieves the pilots of some of the workload during stressful and high-workload situations (Boorman, 2001).

The designers of the checklists in the FMC were able to make them in two ways. They could make the entire sequence fully automatized, or split it up in more sequences and let the crew go through the list. The first option ran the risk of making the process over-automated. If the FMC is totally automatized there might be situations and multiple failures that it has not been programed for. A modern airliner pilot is trained to handle multiple failures using the paper checklists, and they also have the ability to handle far more complex and variable situations that can occur. By using their knowledge and previous experience they can make better decisions than a machine could (Boorman, 2001).

3.8 Creating a new checklist

There is no one methodology for creating a checklist. Different researchers have provided different methods depending on what they found to be the most important aspects of the design process. However, there are some similarities. Many of the aspects discussed in this chapter will be described in more detail under their own respective headings.

Many studies stated that the first step of creating a checklist should be to clearly define the checklist objective (Degani & Wiener, 1993; Verdaasdonk, 2009; Ziewacz, 2012; Wilson, 2013). Ideally the checklist should fit on a single page, which may in some cases mean that the objective has to be more clearly defined, or the checklist divided into smaller parts (Degani & Wiener, 1993; Weiser et al., 2010).

Reviewing checklist design guidelines is also important. Many decisions regarding design can and should be made in the beginning of the development process. Examples are type of checklist, color, typography, response format, among others (Ziewacz, 2012; Wilson, 2013). Hales, Terblanche, Fowler & Sibbald (2008) stressed the importance of using peer-reviewed guidelines and evidence-based best practices when creating a new checklists.

The timing of the checklist should be considered. In order to determine what should be included in the checklist and in what order it is important to decide when the checklist should be initiated (Weiser et al., 2010).

When a first version of the checklist is completed it should be submitted for testing. In aviation this is usually done in a simulator. Changes are made based on feedback from the tests, which will usually reveal if something is unclear or if there are other problems (Verdaasdonk et al., 2009; Weiser et al., 2010). Additionally, all stakeholders should be included in the process and be given a chance to provide feedback (Verdaasdonk et al., 2009; Ziewacz; 2012; Wilson, 2013).

The checklist should then be formally field-tested and evaluated. This will make sure that the checklist is working as intended. Through observation and measurement of the process this will also make sure that the checklist is actually improving safety (Verdaasdonk et al., 2009; Weiser et al., 2010; Wilson, 2013).

If applicable the checklist should also be modified locally at the workplace where it is to be used. The checklist must be adapted to their specific procedures and workflow. This will make it easier to adjust to including the checklist into everyday operations (Hales et al., 2008; Weiser et al., 2010).

In the maritime sector one of the biggest developers of checklists is the ICS. In an e-mail conversation with John Murray, Marine Director of ICS, he suggested that their checklists are created mainly by utilizing the experience of a team of operational experts. However, he also stated that they use experience from other fields and take that into account when creating their checklists. Hans Hederström, Managing Director of CSMART stated that they have studied several books as well as used an experienced airline captain as a consultant when developing checklists.

3.9 Types of checklists

There are several different ways of categorizing checklists. The type of situation the checklist is meant to be used in should directly influence the design to accommodate for the difficulties and requirements of that specific situation and the limitations of human cognition and capabilities (Turner & Huntley, 1991; Mauro, Degani, Loukopoulos & Borshi, 2012; Wilson, 2013).

Two fundamental types of checklists are the read-do and the do-confirm checklists. The do-confirm checklist is intended to be used as a redundant procedure to improve safety. The idea is that the crew performs the task required from memory, and then verifies that no mistakes or omissions were made by reading the checklist, and confirming that each item has been performed correctly. In aircraft this is usually performed by having the pilot not flying read the checklist out loud, while the pilot flying will check the status of that item. The pilot not flying will also cross check, creating a mutual redundancy between the two (Degani & Weiner, 1990; Gawande, 2009; Wilson, 2013).

The read-do checklist is much like a recipe. Each item is read from the checklist and the required action is performed. The read-do checklists are usually longer and more time consuming than the do-confirm checklist. They also lack redundancy that the do-confirm list offers (Degani & Weiner, 1990; Gawande, 2009; Wilson, 2013).

Checklists can also be categorized according to what their purpose is. Example of such types of checklists are procedural checklists, evaluation checklists, feature checklists, behavior sampling checklists, entry/exit checklists and research checklists (Wilson, 2013).

The procedure checklist is a list that includes the different items that are required to complete a task or procedure successfully, while avoiding mistakes and omission of items. For example if the task is to plan an event, examples of items required would be booking a conference area, setting the budget and making sure there is enough food. It may contain all the steps or only the most critical ones depending on the objective of the checklist (Wilson, 2013). An example of a procedure checklist in shipping can be found in Appendix 1.

An evaluation checklist can be used to evaluate a specific product, to make sure that it meets the required standards. The checklist must be adapted to the experience of the user. Accomplished users may have a deeper understanding of the specific subject compared to a person with little or no previous experience (Wilson, 2013).

A feature checklist is a checklist that can be used in an interview to gather information about a feature such as frequency of use or importance of features in a system. Examples of questions could be “How often do you use it?”, “How do you use it?” and “With what frequency do you use it?”. This gives the interviewer an opportunity to gather qualitative data that can in other way be very hard to collect (Wilson, 2013).

The behavior sampling checklist can be used to record a specific type of behavior of an individual during a specific period of time. This method is used both in laboratory and as a field study method. Examples of behaviors that can be recorded are how often a person smokes, drives or listens to the radio. It is important to choose the correct level of granularity for the items on the checklist. If the objective is to examine a person’s mobile phone use during one day, the user of the mobile phone may use it for many different things, such as making phone calls, texting, browsing the internet or taking photos. Depending on what the research goals and delimitations are it may be necessary to find a method to sift the information to find what is relevant and what it not (Wilson, 2013).

An entry/exit checklist can be used to evaluate a product that is close to being released.
When developing software an entry checklist may be used to decide whether or not a product is ready to be submitted for a usability inspection. The entry checklist can be used for physical products but can also be used for a service or a process. An exit checklist may be used to determine if the product is ready to be released to beta customers. It can also be used when leaving an employment (Wilson, 2013).

A research checklist can be used when evaluating the methodology used in scientific research studies. It can also be used as a guide when writing research papers. The checklist may include items that deal with things such as the motivation for the research, the limitations of the research, or the quality of the used references (Wilson, 2013).
Furthermore, checklists in aviation are commonly categorized into normal, and non-normal checklist. Non-normal checklists are sometimes further categorized into abnormal and emergency checklists (Gross, 1995).

Normal checklists are used in foreseeable circumstances, and their main goal is to ensure that ordinary operational procedures such as configuring an airplane for takeoff or preparing an operational theatre for surgery are performed in a standardized manner, in order to prevent mistakes and omissions (Degani & Weiner, 1993; Boorman, 2001).

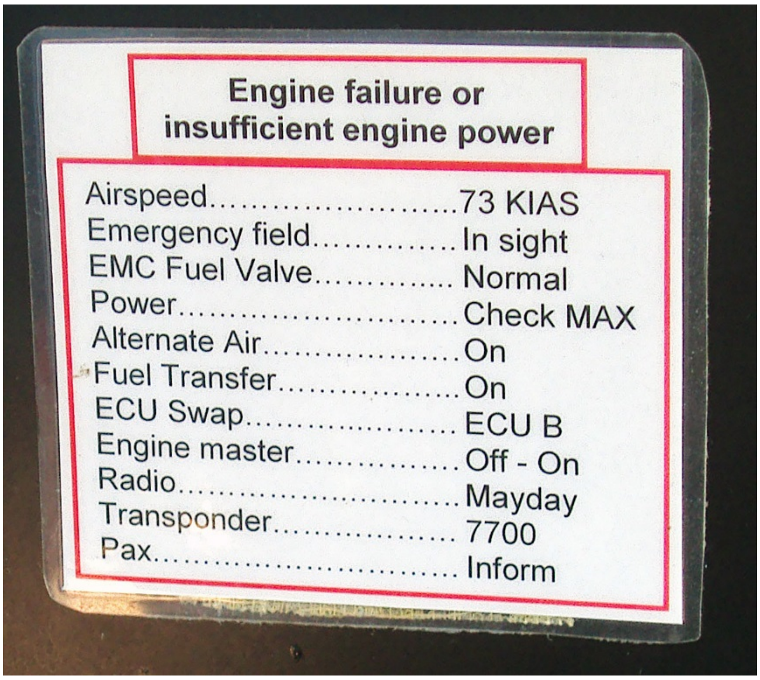

Figure 2 Emergency checklist in aviation (Hedlund, 2014).

The differences between an abnormal and an emergency checklist are a bit more subtle. Abnormal procedures are defined by the Civil Aviation Authority (2006) as “Procedures that require actions to maintain safe flight, and prevent further incidents from occurring”, while emergency procedures are defined by the same as “Procedures that require immediate action in relation to situations that threaten physical danger to people, and/or damage to the aircraft”. In either case, the pilot has to be guided through a complex setup of tasks that is highly variable depending on what the problem is (Boorman, 2001; Burian, 2006).

It is not always possible to perform all the steps of the checklist in chronological order from top to bottom. In airplanes there seems to be almost an endless amount of combinations of errors that can occur. When designing a checklist it may be important to consider what type of checklist is needed and adapt the design to the specific requirements of the situation (Boorman, 2001; Burian, 2006).

Both the normal and the non-normal checklists are taught and trained in modern flight schools, though the non-normal checklist are usually not used much after flight school. This can lead to uncertainty and decreased skill amongst the crew in how to use it. The normal checklist is used every day and the crew is therefore very familiar with it. For this reason manufacturers try to make the non-normal checklists look much like the normal checklists in both design, color and layout. Everything in order to minimize the errors that can occur in a stressed and unfamiliar situations (Boorman, 2001).

3.10 Typography

Typography is defined by Merriam-Webster's dictionary (2014) as "the style, arrangement, or appearance of printed letters on a page". There are two important concepts to be aware of when discussing typography – legibility of print (discriminability) and readability (Degani, 1992). Legibility of print refers to the properties of an alphanumeric that enable the reader to easily and positively distinguish it from others. Readability refers to how quickly words, symbols or abbreviations can be recognized. It should be noted that most of the research on typography has been done in laboratory studies. It also focuses mainly on paper checklists (Degani, 1992).

- **Typeface**

Typefaces, also known as fonts, refer to the style of alphanumerics in a document. It is important to make sure that each individual character is easily distinguishable and unambiguous. Researchers have been able to show that sans-serif fonts provide greater legibility to roman fonts, as long as other typographical factors are controlled. Sans-serif fonts are fonts that do not include the serifs, which can usually make them look simpler and cleaner. Serifs are the small strokes that project horizontally from a large stroke, for example at the top and bottom of an upper case letter "I" in Times New Roman, as in this text. However, serifs may aid the horizontal movement of the eye by preventing it from accidentally slipping to the next or previous line. This could possibly be helpful in very long lists (Degani, 1992).

In 1965 researchers compared several different fonts, and found significant increases in reading comprehension when using a font called Gill-Medium, now called Gill Sans MT. Many other sans-serif fonts include characters that resemble each other, which may decrease legibility (Degani, 1992). Degani (1992) lists some of the properties that may impede the reader's ability to distinguish between characters as:

Figure 3 Example of Gill Sans MT font (Hedlund, 2014).

1. The standardized or modular appearance of the character ("P," and "R")

2. The effect of mirror image between the upper and the lower part of the character ("E," "B," and "D").

3. The use of equal radius for different letters ("G," "O," and "C").

- **Upper-case versus lower-case**

Researchers in the human factors community mostly agree that lower-case letters are more legible than upper-case characters. Studies have reported that texts written in lower-case are read faster and found more pleasurable by the reader. This is believed to be due to a few different factors. Most of the reading that we do every day is in lower-case. This means that most of us are much better accustomed to looking at and reading lower case-words. The human memory will store the shape of a word, the "total word form", making us able to recognize it faster than we would if reading each individual character. Lower-case words generally have a much more unique shape because of ascenders and descenders. Ascenders are the vertical strokes in for example the letters "d" or "b", while descenders are the vertical strokes in for example the letters "p" or "q". Upper-case words lack this attribute, and look more like a rectangular box. This forces the mind to read each individual letter, significantly slowing down the speed of reading and readability (Degani, 1992).

If a paragraph of text is written with only capital letters, the boxy shape of each word could also make the text seem like a pattern of vertical lines or "stripes". This may induce discomfort and unusual visual effects (Degani, 1992).

It has also been shown that visual emphasis on the first letter of a word will significantly increase legibility. When writing in lower case it is natural to start a sentence or proper name with an upper-case letter. However, the same effect can be shown when writing in upper-case if the first letter of a sentence is made bigger than the following (Degani, 1992).

- **Italics and boldface**

Italics have been shown in experiments to slow down reading by 2.7 percent, and 96 percent of the participants in the study found the text less legible than the same text with just a regular roman font (Degani, 1992).

Boldface showed no reduction in the subjects' reading speed, however, 70 percent of the subjects found it unpleasing. Boldface has also been shown in other experiments to be no less legible than normal face in low illumination conditions. This means, although it may not be optimal to print a whole checklist in boldface, it may be used to emphasize important words or phrases without compromising legibility (Degani, 1992). This is under condition that the bold text is formatted clearly and letters are equally spaced (Hales, 2008).

- **Font height**

When measuring font height there is some confusion about the methods of measurements and the scales that are used. Firstly the height of a "point", a unit for measuring font height, has been a different height on different platforms. On traditional printers a point used to be 1/100 of an inch (Degani, 1992). However, according to the software company Microsoft (2014), a point in their word processor program Microsoft Word is approximately 1/72 of an inch, which is equal to 0.014 inches or 0.035 cm.

Furthermore, when measuring font height you must also differentiate between "overall height" and "x height". "Overall height" meaning the height of characters from the top of the ascender to the bottom of the descender, while "x height" refers to the height of a character that lacks ascenders or descenders (Degani, 1992).

The font size recommended by the literature depends on a few different factors such as illumination level, viewing distance, stroke width and visual acuity. Most of this research was made with aviation in mind and as such may not always be appropriate for maritime application (Degani, 1992).

Illumination of the reading surface is an important factor when deciding the font height. Luminance is defined as the amount of light reflected by a surface, and is measured in foot-lambert (fL). A normal viewing distance is 40-60 cm, although this may vary depending on the situation and conditions. If you know beforehand that the checklist will need to be used in dark conditions the font height will need to be larger to make characters easier to distinguish. A study suggested that for a viewing distance of 71 cm or less and a luminance level of 1.0 fL or less, a font height of at least 0.5 cm should be used, although this study had a few flaws such as using only upper-case letters and paying no attention to contrast or color, among others (Degani, 1992).

Also, depending on the length of the list, font height may have to be compromised in favor of saving space. Nevertheless, as a general rule, the “x height” of the font should be between 0.35 cm and 0.5 cm. If there is not enough room on the page, the font should at least not be smaller than 0.25 cm. Naturally, the font height chosen must be combined with other typographical factors in a suitable way (Degani, 1992).

- **Stroke width and height-to-width ratio**

The stroke width is a function of the height of the character. The stroke width affects the eye’s ability to separate the vertical stroke of a character (“I”), from the space within a character (“E”). Most recommendations by human factors data books suggest a height-to-width ratio of 5:3. It should be noted that this recommendation is only for when the viewing angle is approximately 90 degrees to the document surface. With paper checklists it will seem natural for the reader to turn the checklist toward oneself when reading, however, if using a fixed platform for the checklist or an electronic checklist, depending on the positioning the stroke width may have to be adjusted to make the perceived width appear as if the height-to-width ratio was 5:3 (Degani, 1992).

- **Horizontal and vertical spacing**

The effect on horizontal and vertical spacing is most apparent when the font is small, however, using correct spacing could have several different benefits. For one, it makes the text clearer and easier to read. This is most dependent on vertical spacing between lines, and it may even be possible to decrease the horizontal spacing between characters and word if the vertical spacing is increased. This would in effect mean increased legibility and readability without sacrificing space. Also, increased vertical spacing decreases the risk of the eye slipping to an adjacent line while reading, which is critical when reading important documents such as checklists. Skipping a line on a checklist could mean omitting an important item from the procedure. It also decreases the probability of discomfort or unusual visual effects that may occur if the text forms a pattern of “stripes”, which was discussed previously (Degani, 1992).

The recommended vertical spacing is 25-33 percent of the overall size of the font. Note that this is a measurement different from the “x height” of the font. The horizontal spacing between characters should not be less than one stroke width. The spacing between words should be 25 percent of the overall height (Degani, 1992).

- **Cumulative effect of typographical features**

While some typographically sub-optimal choices may not make a checklist completely lose its value, a combination of them can make a checklist significantly harder to read, as well as reduce the comfort of the user and their willingness to use it. It could be worth spending a few extra minutes in the beginning of the development phase to make sure that the typography is sound. *To illustrate, this text is not as legible as it could be*, while this text is significantly easier to read (Degani, 1992).

3.11 Contrast and color coding

Researchers have tested using black print on a white background, as well as white print on a black background. Although there have been conflicting studies, most results point toward that black print on white background is superior. In tests where speed of reading, subject's preference, eye movement measurements, recognizability in the peripheral vision and discriminability at a distance were tested, black print on white background was proven to be better. It was also found to be easier to recognize at an angle (Degani, 1992).

In aviation it is common that checklists are color coded depending on the situation the checklist is meant to be used. Normal checklists are usually black print on a white background, while abnormal checklist are black print on a yellow background and emergency checklists are black print on a bright red background. Humans usually associate colors with emotions or types of situations. For example, green is usually associated with normal, while red is associated with danger or emergency (Degani, 1992).

Studies have reported that luminance differences have a larger impact on the reader's ability to visually distinguish characters than color differences. For example, red and blue have a good color contrast, but significantly less luminance contrast (Degani, 1992).

Black print on a white or yellow background has been shown to be the best choice for luminance contrast. If another color is to be used for the background, it should have a reflection percentage of at least 70 percent, and the luminance ratio between the characters and the background should be around 1:8. It should be noted older people will have a harder time differentiating between color, especially shades of blue-green and red. Black print over dark red, blue or green should always be avoided (Degani, 1992).

3.12 Phraseology and abbreviations

Words and sentences should be kept simple, clear, and phonetically balanced, while maintaining the technical language of the field (Degani & Wiener, 1990; Weiser et al., 2010). Phonetic balance is a property of a word which means that the various phonemes (phonetic units) occur at approximately the same frequency as they do in speech in the used language

(Farlex Partner Medical Dictionary, 2014). They should also be words that are used frequently (Degani & Wiener, 1990; Weiser et al., 2010). As few words as possible should be used without compromising the message. An active voice should be used, for example, “configure radar” rather than “radar is configured” (Gross, 1995). Special care must be taken when phrasing checklist items. Words such as “set”, “checked”, “completed”, “configured”, and similar, can be ambiguous. The same goes for responding to a checklist call when working in a team, in many cases instead of saying for example “checked”, the actual value or setting could be called out instead (Degani & Wiener, 1990).

Users should refrain from using non-standard phraseology when reading the checklist to another crewmember. A lot of miscommunication is caused by assuming that others will understand. This is an especially prominent issue in users that use radio communication frequently. The user may be tempted to use non-standard phraseology to be unique, show humor, show a high level of competency or if he finds the checklist inadequate. Beyond the obvious risk of misconceptions or mishearing, this belittles the checklist procedure in the eyes of the other crewmembers and may reduce the efficiency of the checklist in detecting errors (Degani & Wiener, 1990).

If checklist items include acronyms or abbreviations it must be made certain that there is no risk of misinterpretation. In some cases these may have several different meanings or their meaning may not be clear to all users (Wilson, 2013).

3.13 Item specificity

When creating a checklist it must be ensured that items are not too general. An item in a maritime departure checklist prompting the user to “prepare bridge for departure” is probably too general. It may be unclear to the user what checks are required. On the other hand, items should not be excessively specific. An item reminding the user to “set radar’s range to 6 Nm” could be too specific (Wilson, 2013).

3.14 Deciding what items to include

Deciding what items to include may be one of the hardest decisions to make when developing a checklist, as well as one of the most important. There are two main schools of thought. The first is the human performance approach, which aims at making the checklist in a way that accommodates human capabilities and limitations to make sure that the checklist catches the most critical errors while still avoiding to become a nuisance task. The second is the engineering approach, which includes every step of the procedure into the checklist (Degani & Wiener, 1990; Degani & Weiner, 1993; Verdaasdonk et al., 2009).

Advocates of the human performance approach believe that since the checklist in itself is a redundant task, only the most critically important items should be included. Such items are sometimes called “killer items”. Basically, a killer item is an item that singlehandedly, if omitted, could result in an accident occurring. They should also be items that are likely to be missed. Some tasks may be considered second nature to the user, and virtually impossible to forget, while others may be less intuitive. Another clear advantage of this method is that even though it may not cover everything, it is more likely to be used because of the shorter completion time. However, accident research has shown that things that are seemingly insignificant could in some cases have devastating consequences. This makes it difficult to decide what exactly should be defined as a killer item. Operational experience will play a large role in being able to determine the relevance and important of tasks (Degani & Wiener, 1990; Degani & Weiner, 1993; Gawande, 2009; Verdaasdonk et al., 2009).

Some argue that using the engineering approach may have an adverse effect on the user’s attitude towards the checklist, and increase the risk of it being cut short or skipped altogether. On the other hand, a long checklist including everything relevant to the procedure may, if used correctly, be more effective in verifying that every single item is correctly performed and checked (Degani & Wiener, 1990; Degani & Weiner, 1993; Gawande, 2009). Furthermore, what constitutes if a checklist is to be considered “short” or “long” may vary depending on the objective. The complexity and inherent risks of the tasks must be considered, and critical thinking must always be employed when making these decisions (Gross, 1995; Wilson, 2013).

In some cases it may be a good idea to include duplication of items. However, while this can further decrease the chance of something being missed, it can also have a negative effect on actual safety by diminishing the overall checklist performance. The practice of using duplication of items is most common in long and detailed checklists. If duplication is to be considered, it should only be used on “killer items”, that are in great danger of being missed, misinterpreted, or for which the circumstances may have changed during the course of completing the checklist (Degani & Wiener, 1990; Degani & Wiener, 1993).

3.15 Order of items

When deciding the order in which to put the checklist items, the first thing that must be taken into account in the operational sequence of the systems. For example, in an airplane it is inappropriate to check the hydraulic pressure before starting the hydraulic pumps. Next, the geographical locations of systems should be taken into account. An example from shipping could be configuring instruments in sequence from the port bridge wing to the starboard bridge wing. Some form of logical “flow pattern” such as this will make it harder to accidentally omit steps from the procedure (Degani & Wiener, 1990; Degani & Wiener, 1993). Additionally, since the probability of an item being completed is highest at the top of the checklist, the checklist should start with the most critical items, if possible. If the checklist is cut short for some reason, it will be less likely that essential items have been missed (Degani & Wiener, 1990; Degani & Wiener, 1993; Winters et al., 2009).

External circumstances should also be considered. If a certain step has to be skipped and deferred for later, for example the fueling of a plane, this may increase the risk of it being forgotten completely. During a flight from Denver in 1987, this was exactly what happened. Since the fueling of the plane was not finished when they were going through the pre-flight checklist, the pilots postponed the item for later. Approximately ten minutes after departure the plane was forced to return to the airport, after the second officer had noticed that the plane did not have enough fuel to reach its destination. The limitations of human memory must be taken into account when creating a checklist (Degani & Wiener, 1990; Degani & Wiener, 1993).

The seven plus-or-minus two rule, also known as Miller's Law, states that the human working memory can normally store between five and nine items. However, it also states that information can be stored in "chunks". For example, when remembering a random set of single syllable words, the average person can remember five. Each word consists of about 3 phonemes, but as a result of learning they are clustered together in our memory as words (Miller, 1956). When creating a checklist, especially if it is a long checklist, this property of human memory can be utilized by dividing the list into groups of similar tasks in logical "chunks". Different groups of tasks should then be separated physically to further illustrate the division (Degani & Wiener, 1990).

3.16 Response format

According Wilson (2013), it may be important to have some kind of validation in a checklist showing that a specific task has been accomplished. If the checklist allows for putting a mark or a signature, this will serve as confirmation that the task has been completed and unchecked items will be easy to recognize and redo at a later stage. The response format has to be appropriate in relation to the question. Some flexibility may be required. Some examples of response formats in checklists are:

- Checkbox ☑ / ☐: Indicates in a clear way that a task on the checklist has been done.
- Yes / No: Is used as response when the item is a question.
- "Always", "Most of the Time", "Sometimes", and "Never": Is used when the question refers to the frequency of an occurrence.
- Tick marks (𝍸) : Is used to count how many times an activity is performed. Can be used for observational or testing studies. An example can be when observing a person performing a task and recording how many times someone is doing something deviant related to that, or when testing a product and counting how many errors it makes during a specific time. It is important to make sure that there are enough space for the tick marks in the form, there can be more ticks than expected.
- "Done", "Done incorrectly" or "Not done": Can be used when the item is a task. When dealing with complex tasks, sometimes additional response options may be required (Schmutz et al., 2014).

Wilson (2013) gives an example of a Yes/No checklist that can be used to evaluate the usability of questions when creating a checklist:

Use this checklist to evaluate the usability of each question on your questionnaire	**Circle YES or NO for each question**	
Is the question a "double question"?	YES	NO
Does the question have a technical term that won't be understood by respondents?	YES	NO
Is the reference period missing (e.g., "during the last month")?	YES	NO
Does the question require respondents to perform complex mental operations?	YES	NO
Can the question be interpreted in multiple ways?	YES	NO

Table 1 Example of a Yes/No checklist (Öhrn, 2014).

3.17 Initiation

The timing of the initiation of a checklist can have a large effect on its effectiveness. A checklist is only effective if it is initiated and performed while it is still possible to remedy any error, inconsistency or problem that is found, and prevent them from causing an accident. It should also be initiated when it does not interfere too much with other tasks (Weiser et al, 2010).

In an airplane cockpit there are usually two pilots. The captain is always the pilot in command, but it is common to rotate the duties of the pilot flying between the two pilots, both in order to share the workload and for training purposes. It is the job of the pilot flying to judge when to call for a checklist. The pilot not flying will then read the checklists out loud, using a call and response system, and cross check the checks that the pilot flying performs for redundancy. Many pilots use external cues to tell them when to initiate a checklist. Initiation errors can occur when those cues are not present, or missed due to stress, high workload or other circumstances (Degani & Wiener, 1993).

In medicine, more often than not the surgical teams are not static, but are put together out of people who have not previously worked together. When using the WHO Surgical Safety Checklist it is usually the operating nurse that initiates the checklist. The reasoning is that if the person lowest on the hierarchy initiates the checklist, this opens up communication channels and makes it more likely that any errors detected will be brought to the surgeons attention. The checklist is initiated when there is a natural pause point in the work flow. This is important in order to avoid the checklist being viewed as an undue interruption (Gawande, 2009; Weiser et al., 2010).

3.18 Interruptions, distractions and shortcutting

In a study of cockpit distraction reports, 22 out of 169 were labeled as distractions caused by checklists. Common denominators of these reports were that checklist were given priority over other tasks. They were also performed simultaneously with other tasks. From the perspective of designing a checklist, the only way to avoid this is to make the checklist shorter, but it is also the duty of the person initiating the checklist to make sure that the checklist is initiated at an appropriate time (Degani & Wiener, 1990).

It is important to the integrity of the checklist process that steps are not skipped. “This has never been a problem before” is a common, but not very useful, way of thinking. This way of thinking is most common when working only from memory. Checklists provide the necessary steps in an explicit format, and may help to reduce “shortcutting” the procedure if used correctly (Gawande, 2009).

3.19 Completion

The last item on the checklist should be a call that the checklist is completed. When working in a team, this lets each member know that the checklist has been completed. In many cases crewmembers will already be familiar enough with procedures to know that the checklist is completed, nevertheless this can be a useful tool to allow them to mentally move on to other tasks (Degani & Wiener, 1990; Degani & Wiener, 1993).

3.20 Local modification

It is important to note that it is generally not possible to create a single checklist that will be optimal for every workplace. Although recommendations can be made for design and important items, there will always be differences in procedure between different organizations. In aviation as well as in medicine users are encouraged to revise checklists to fit their particular practices and work flow. This makes is easier to fit the checklist into the organization and make it part of everyday operations (Gawande, 2009; Weiser et al., 2010).

3.21 Testing and evaluation

Many researchers advocate testing the checklists in simulators before fully implementing it in actual operations. This may further improve the quality of safety improvement that the checklist can provide. It can also help discover any problems with for example the included or excluded items, clarity due to phraseology or otherwise, or order of items. Those using the checklists in the tests should preferably be the intended end users (Verdaasdonk et al., 2009; Winters et al., 2009).

Evaluation of the efficacy and functionality of a new checklist procedure is a vital part of creating an effective checklist. Feedback from crewmembers working with a checklist is essential in order to make improvements. It will also, such as the case with the WHO Surgical Safety Checklist mentioned previously, be useful to see if there are any actual improvements to safety. Without evaluating there is no way to know whether or not there has been any real improvement (Gawande, 2009; Weiser et al., 2010).

3.22 Process for revising the checklist

In shipping, the ISM code states that "the Company" is required to establish procedures, checklists among others, for key shipboard operations. It also states that "the Company" should control all such documents, and any changes must be reviewed and approved by authorized personnel. Those documents should also be found in the SMS (Safety Management Manual) on board (International Maritime Organization, 2010).

In many organizations checklists are used for many different kinds of operations. For example in aviation the checklist is not only a good reminder but a versatile tool critical to a successful operation. To make sure the checklists stays useful it needs to be regularly updated. This might be a problem in many types of organizations as the original author of the checklist may have left the organization a long time ago. It is possible that nobody knows who is responsible for updating the checklist anymore. It is important to have clear guidelines for this type of procedure (Wilson, 2013).

The checklists can be updated due to many reasons, among them are making sure the checklist reflects current procedure, new rules and regulations but also to improve the checklist's usability and performance. When the checklist has been used for a while Wilson (2013) describes the following steps for gathering feedback and keeping it up-to-date:

1. Gather feedback from your users on the usefulness and usability of the checklist.

2. Update the checklist based on feedback from users or changes to the target checklist. If you had a usability defect checklist for graphical user interface (GUI) applications, you would update that checklist when you move from GUI applications to web or mobile applications.

3. Provide a contact and mechanism to collect feedback from future users of the checklist.

4 Results

In this review no evidence has been found to suggest there has been any research made regarding checklists with maritime applications specifically in mind. While a few studies that mention maritime operations could benefit from their research, most of the studies have had aviation or medicine as their main focus. Although hardly comprehensive evidence reflecting the whole industry, our conversations both with the ICS and CSMART suggested that they rely more upon operational experience than scientific research when designing and reviewing their checklists. When searching for literature nothing published was found from either organization which suggests that their approach is not scientific. This is does not prove that their approach is not systematic, but suggests that it is not based on science. Although there is a lot of research on the subject of checklists, we have not found that the shipping industry in general is utilizing that knowledge.

In aviation and medicine it there is definitely a systematic approach to the development of checklists. They use a combination of scientific research and empirical evidence. Although the amount of published research is not monumental, it is apparent that nothing is left to chance. Most conceivable design aspects have been researched in one way or another, but while it is important to back decisions with scientific data, researchers also stress the importance of testing and evaluating the finished product. After all, if no improvement of safety can be shown, a checklist is simply another procedure increasing the workload for no good reason. Operational experience as well as testing will be important factors in being able to determine what type of checklist items are actually needed and what type of items that are going to be superfluous.

What constitutes a well-designed checklist depends on what type of checklist is required. There are some general guidelines that are proposed by the current research, most of which could apply to almost any checklist. Each point is more thoroughly explained in the background chapter of this thesis.

The physical form will have some effect on the design of the checklist. While this thesis mainly deals with design of paper checklists, it should be noted that there are several other types of checklists available. Each should be considered depending on the needs of the end user.

The situation the checklist is supposed to be used will have a large impact on whether a read-do or a do-confirm checklist is superior. Each has their advantages and disadvantages. Again, the needs of the end used must be considered when making this decision.

The font should be a sans-serif. Gill Sans MT is a good example of a legible sans-serif font. Lower case should be used, except on the first letter in the first word in a sentence or in proper names. Italics should be avoided. Boldface can be used to emphasize individual important words.

The height-to-width ratio of character though be 5:3, assuming that the checklist will be read at a 90 degree angle. The vertical spacing should be 25-33 percent of the overall size of the font. The horizontal spacing between characters should not be less than one stroke width, and the spacing between words should be 25 percent of the overall height.

When choosing colours, black print on a white or yellow background has been shown to be superior in many aspects. Black print over dark red, blue or green should always be avoided. If another color is preferred, it should have a reflection percentage of at least 70 percent, and the luminance ratio between the characters and the background should be around 1:8.

Words and sentences should be kept simple, clear, and phonetically balanced, while maintaining the technical language of the field. An active voice should be used, for example "set radar" rather than "radar is set".

Tasks on the checklist should not be too general. They should be specific enough that the objective of the task is clear. Nevertheless, items should not be overly specific.

Advantages and disadvantages of the "human performance approach" and the "engineering approach" must be taken into account when deciding what items to include. Different approaches may be more appropriate in different situations depending on the circumstances. Duplication of items should be used sparingly but can be considered in some cases.

The most important items should be highest on the list, this makes it less likely that they are skipped. Items should be organized by geographical location when possible, so that a natural "flow pattern" can be established. External circumstances must be accounted for, if an item has to be deferred for later this may greatly increase the risk of it being omitted. Similar items can also be organized in logical "chunks", making them easier to remember by taking advantage of the properties of the human working memory.

If a validation that the checklist has been completed is required, the response format should be chosen depending on what is suitable considering the properties of the checklist. In some cases checkboxes may be sufficient; while in others additional options may be required.

It must be clear when the checklist is to be initiated, and who is supposed to initiate it. Preferably the checklist should be initiated when there is a natural "pause point". Basing the initiation on external cues may be disadvantageous if the cue is missed or is absent.

The last item on the checklist should be a call that the checklist is completed. Although this may in some cases seem superfluous, it allows other team members to mentally move on to other tasks.

If the checklist is to be used in several workplaces, it may be necessary to locally modify the checklist to fit it into the existing workflow. While it may not be recommended to change the design or remove important items, some changes can be made without adversely affecting the efficacy of the checklist.

Testing and evaluation is an essential part of the checklist design process. It can help reveal problems with for example the included or excluded items, clarity, order of items. It also allows users of the checklist to give their feedback. Additionally, evaluation the checklist's performance will show whether or not it is providing an actual increase in safety.

Finally, it is important to have a process for revision of the checklist. There are many reasons for a checklist to have to be revised or updated. It should always be clear who the author is and who is responsible for making these updates.

5 Discussion

We chose to do a systematic literature review because this would allow us to analyze most of the existing research on the subject. We believed that this would be the most effective approach for us to answer our research questions.

We wanted to be as systematic and structured as possible when to make sure that we found all or most of the relevant research written on the subject. It was also important to us that the process would be transparent and easily repeatable. Therefore we chose to use the PRISMA method. This method gave us a step by step guide which made sure that we were consistent throughout the process.

We did not follow the PRISMA method down to the letter. Since this method was originally designed for postgraduate students, we had to simplify it slightly to fit our purposes and time frame. We do not believe that this had any effect on the result, but possibly a slight effect on the transparency of the process.

The advantage of a systematic literature review is that it gives a clear picture of the research, gathering what is currently known in one place. This method also makes it possible to cover a very broad range of research.

The disadvantage of using this method was that we did not have much contact with existing shipping companies or the shipping industry. In order to properly answer our first research question it might have been better to use alternative methods such as interviews. This would have given us a better insight into their actual procedures for developing checklists. However, we delimited ourselves from this due to the limited time. Instead we focused on looking at the published research on checklists.

The references used in the thesis were exclusively scientific texts and can therefore be assumed to be trustworthy. The reliability of the information was confirmed both by verifying the credibility of the authors and by cross-referencing whenever possible.

Because of our limited time we only searched four databases. It is possible that there is additional relevant information in other databases in which we have not searched. The possibility that this may have changed the result of the thesis cannot be excluded completely.

One of the most striking results of this review was that we did not find any research relating to checklist design or checklist use in shipping in any of the databases we searched. Although perhaps not surprising, it is still odd that a field that puts so much emphasis on safety has made no research of its own in this area. Shipping is a field that is in many ways bound in old traditions, and it is possible that this is a contributing factor in the reluctance to do research and implement new procedures. Although the absence of published and peer-reviewed research suggests that their approach to checklist development is not scientific, it does not conclusively prove that it is not systematic.

It is possible that individual shipping companies have done their own research when developing their checklists. It is likely that this is something that we would not have found in scientific databases. However, whether or not such research can be called scientific if it is not peer reviewed is questionable.

We found that large developers of checklists in shipping are relying mostly on experience and have an empirical way of creating their checklists. While experience is essential for understanding the procedures for which the checklists are made, experience most likely does not know which the most legible font is, whether upper case or lower case is the most optimal, or what color the paper should be. The cumulative effect of poor design choices could make a big difference for the end result. A combination of procedures designed by experienced individuals presented using the design that human factors research recommends could most likely improve effectiveness and receptivity for the procedure.

In aviation and medicine we found quite a lot of research concerning checklists. Although many studies were quite old, we believe that not much has changed in their relevance. It does not seem too farfetched to assume that not much has changed in the way we perceive text or in what kind of help the human memory needs from a checklist. Most of the research has been made with aviation in mind, however, much of the research in medicine seems to stem from and build on the research made previously in aviation. The fact that almost every conceivable factor in design has been researched in some way clearly indicates that there is a very systematic approach in the design of checklists in these fields.

We also found no evidence to suggest that there is any feedback from the actual users of the checklists in shipping after they are created. It is likely that in many cases nobody outside the team creating the checklist has any chance to provide feedback on the product. From the studies we have read this is not considered optimal. Feedback from the stakeholders should be an integral part of the development process, as well as provide input for future revisions and updates.

In aviation and medicine, in addition to a very well-structured development process, checklists are always tested and evaluated before being put into actual use. The checklists can be redesigned several times before being considered complete. As an example, the WHO checklist was tested in simulations, used on trial in hospitals, and formally evaluated, before finally being released as a completed product.

We believe that proper checklist design could improve compliance with the checklist procedures on ships. It is not hard to imagine how a poorly designed checklist might discourage crewmembers from using it. It is a widespread opinion that checklists inhibit the use of seamanship, or common sense. But we ask ourselves, is it truly common sense to believe that every procedure could be memorized and recalled during times of high stress and workload? The aim of the checklist should not be to remove common sense from the equation, but to free the mind from the strain of remembering procedures, to allow for more common sense and afterthought. A poorly designed checklist may be cumbersome, and may add to any ill will that already exists. On the other hand, a well-designed checklist may prove to be a useful tool for improving safety.

6 Conclusions

This review has examined the development and design of checklists. The findings indicated that checklists are not created using a scientific approach in the maritime industry. Nevertheless, this does prove that the process is not systematic. Communications with the industry suggests that an empiric approach is used in many cases. In other business as aviation and medicine clear indications of a systematic approach, combining scientific and empiric research, were found. Additionally, design recommendations as suggested by human factors research in checklists were presented.

6.1 Further research

This review examined approach to checklist design in shipping, as well as the design criteria recommended by current research. Further research could develop, test and evaluate a checklist in maritime operations using these recommendations as design guidelines. Evaluation of the checklist could be used to show whether or not the checklist is contributing to safety. It could also be interesting to examine such a checklist from a usability perspective.

References

Boeing. (2014). History: B-17 Flying Fortress. Retrieved November 7th, 2014, from Boeing, http://www.boeing.com/boeing/history/boeing/b17.page

Boorman, D. (2001). Today's electronic checklists reduce likelihood of crew errors and help prevent mishaps. *ICAO Journal*, *56*(1), 17-21.

Burian, B. K. (2006). Design guidance for emergency and abnormal checklists in aviation. Paper presented at the Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting, San Francisco. Thousand Oaks: SAGE Publications. doi: 10.1177/154193120605000123

Degani, A., & Wiener, E. L. (1990). Human Factors of Flight-Deck Checklists: The Normal Checklist (NASA Contractor Report, nr 177549). California: NASA Ames Research Station.

Degani, A. (1992). On the Typography of Flight Deck Documentation (NASA Contractor Report, nr 177605). California: NASA Ames Research Station.

Degani, A., & Wiener, E. L. (1993). Cockpit checklists: Concepts, design, and use. *Human Factors*, *35*(2), 345-359.

Denscombe, M. (2009). Forskningshandboken: För småskaliga forskningsprojekt inom samhällsvetenskaperna (2nd ed.). Lund: Studentlitteratur AB.

Federal Aviation Administration. (2007). Flight Standards Information Management System (FSIMS): Volume 3 General Technical Administration: Chapter 32 Manuals, Procedures, and Checklists for 14 CFR Parts 91K, 121, 125 and 135: Section 12 Aircraft Checklists for 14 CFR Parts 121/135. Retrieved October 12th, 2014, from Federal Aviation Administration, http://fsims.faa.gov/wdocs/8900.1/v03%20tech%20admin/chapter%2032/03_032_012.htm

Gawande, A. (2009). The checklist manifesto: How to get things right. New York: Metropolitan Books.

Great Britain Department of Transport. (1987). MV Herald of Free Enterprise: Formal investigation (Report of Court, nr 8074). London: Her Majesty's Stationery Office.

Gross, R. L., & Editorial Staff. (May, 1995). Studies Suggest Methods for Optimizing Checklist Design and Crew Performance: Improved readability, color coding, listing steps in logical sequence, thoughtful indexing, convenient placement within the cockpit, attention to human factors and many other principles will help ensure that checklists are used as intended. *Flight Safety Digest*, *14*(5), 1-11. Retrieved from http://flightsafety.org/fsd/fsd_may95.pdf

Hales, B., Terblanche, M., Fowler, R., & Sibbald, W. (2008). Development of medical checklists for improved quality of patient care. *International Journal for Quality in Health Care*, *20*(1), 22-30. doi: 10.1093/intqhc/mzm062

Hales, B. M., & Pronovost, P. J. (2006). The checklist: A tool for error management and performance improvement. *Journal of Critical Care*, *21*(3), 231-235. doi: 10.1016/j.jcrc.2006.06.002

International Chamber of Shipping. (2013). International Chamber of Shipping: Representing the Global Shipping Industry. Retrieved October 4th, 2014, from International Chamber of Shipping, http://www.ics-shipping.org/

International Chamber of Shipping, Oil Companies International Marine Forum & International Association of Ports and Harbors. (2006). ISGOTT: International Safety Guide for Oil Tankers and Terminals (5th ed.). London: Witherby & Co Ltd.

International Maritime Organization. (2010). ISM Code and Guidelines on Implementation of the ISM Code 2010. Retrieved November 15th, 2014, from International Maritime Organization, http://www.imo.org/OurWork/HumanElement/SafetyManagement/Pages/ISMCode.aspx

Knudsen, F. (2009). Paperwork at the service of safety? Workers' reluctance against written procedures exemplified by the concept of 'seamanship'. *Safety Science*, *47*(2), 295-303.

Marine Accident Investigation Branch. (2009). Report on the investigation of the grounding of mv Maersk Kendal on Monggok Sebarok reef in the Singapore Strait on 16 September 2009 (Investigation report, nr 2). United Kingdom: Marine Accident Investigation Branch.

Mauro, R., Degani, A., Loukopoulos, L., & Barshi, I. (2012). The operational context of procedures and checklists in commercial aviation. Paper presented at the Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting, Boston. Thousand Oaks: SAGE Publications. doi: 10.1177/1071181312561158

Meilinger, P. S. (2004, October 4th). When the Fortress Went Down. *Air Force Magazine*. Retrieved from http://www.airforcemag.com

Microsoft. (n.d.). Relationship between Inches, Picas, Points, Pitch, and Twips. Retrieved November 11th, 2014, from Microsoft, http://support.microsoft.com/kb/76388

Miller, G. A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review*, *63*(2), 81-97. doi: 10.1037/h0043158

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Altman, D., Antes, G., . . . Tugwell, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, *151*(4), 264-269. doi: 10.1136/bmj.b2700

Phonetic balance. (n.d.). In Farlex Partner Medical Dictionary. Retrieved October 14th, 2014, from http://medical-dictionary.thefreedictionary.com/phonetic+balance

Schmutz, J., Eppich, W. J., Hoffmann, F., Heimberg, E., & Manser, T. (2014). Five Steps to Develop Checklists for Evaluating Clinical Performance: An Integrative Approach. *Academic Medicine*, *89*(7), 996-1005. doi: 10.1097/ACM.0000000000000289

Turner, J. W., & Huntley Jr, M. S. (1991). The Use and Design of Flightcrew Checklists and Manuals (U.S. Department of transportation, nr AM-91/7). Washington: Office of Aviation Medicine.

Typography. (n.d.). In Merriam-Webster.com. Retrieved October 14th, 2014, from http://www.merriam-webster.com/dictionary/typography

Verdaasdonk, E. G. G., Stassen, L. P. S., Widhiasmara, P. P., & Dankelman, J. (2009). Requirements for the design and implementation of checklists for surgical processes. *Surgical Endoscopy*, *23*(4), 715-726. doi: 10.1007/s00464-008-0044-4

Weiser, T. G., Haynes, A. B., Lashoher, A., Dziekan, G., Boorman, D. J., Berry, W. R., & Gawande, A. A. (2010). Perspectives in quality: Designing the WHO Surgical Safety Checklist. *International Journal for Quality in Health Care*, *22*(5), 365-370. doi: 10.1093/intqhc/mzq039

Wilson, C. (2013). Checklists. In M. Dunkerley & H. Scherer (Ed.), Credible Checklists and Quality Questionnaires: A User-Centered Design Method (1st ed., pp. 1-27). Boston: Morgan Kaufmann.

Winters, B. D., Gurses, A. P., Lehmann, H., Sexton, J. B., Rampersad, C. J., & Pronovost, P. J. (2009). Clinical review: Checklists - translating evidence into practice. *Critical Care*, *13*(6), 210. doi: 10.1186/cc7792

Ziewacz, J. E., Berven, S. H., Mummaneni, V. P., Tu, Y. H., Akinbo, O. C., Lyon, R., & Mummaneni, P. V. (2012). The design, development, and implementation of a checklist for intraoperative neuromonitoring changes. *Neurosurgical Focus*, *33*(5), E11. doi: 10.3171/2012.9.FOCUS12263

Appendix 1 – Passage Plan Appraisal Checklist

CHALMERS

Passage Plan Appraisal

Valid from 15.07.2008
Version 3

CL047

1.		Have navigation charts been selected from chart catalogue, including	
	1	Large scale charts for coastal waters	
	2	Small scale charts for ocean waters	
	3	Planning charts	
	4	Routing, climatic, pilot and load line zone charts	
2.		Have publications been selected, including	
	1	Sailing directions and pilots book	
	2	IMO Ship's Routeing	
	3	Light lists	
	4	Radio signals	
	5	Guides to port entry	
	6	Tide tables and tidal stream atlas	
3.		Have all navigation charts and publications been corrected up to date, including	
	1	The ordering of new charts/publications, If necessary	
	2	Notices to mariners	
	3	Local area warnings	
	4	NAVAREA navigational warnings	
4.		Have the following been considered	
	1	Ship's departure and arrival draughts, transit draughts, load line zone	
	2	Ship's cargo and any special cargo stowage/carriage restriction	
	3	If there are any special ship operational requirements for the passage	
5.		Have the following been checked?	
	1	Planning charts and publications for advice and recommendations on route to be taken	
	2	Climatologically information for weather characteristic of the area	
	3	Navigation charts and publications for landfall features	
	4	Navigational charts and publication for Ship's Routing Schemes, Ship Reporting Systems and Vessel Traffic Services (VTS)	
	5	Taking into account the marine environmental protection measures that apply, and to avoid, as far as possible, actions and activities which could cause damage to the environment.	
6.		Has weather routing been considered for passage?	

7.		Have the following preparations been made for port arrival?	
	1	Navigation charts and publications studied for pilotage requirements	
	2	Pilot card updated	
	3	Port guides for port information including arrival/berthing restrictions	

Reference:

ISM 7