Master Thesis

The contribution of Human Machine Interfaces on the Situation Awareness of operators.

Evert van der Weit

April 5, 2017

Department of Artificial Intelligence
University of Groningen

Supervisors:

Internal: Fokie Cnossen
External: Jan Siegersma
Redmer van Tijum

Human Machine Communication
***University* of Groningen**

Contents

List of Figures

List of Tables

List of Abbreviations

APC	Advancded Process Control
ATC	Air Traffic Control
CAMS	Cabin Air Management System
Cpk	Process capability index
CTQ	Critical To Quality
DMC	Dot Matrix Code
ECP	Electrical Chemical Process
EID	Ecological Interface Design
GUI	Graphical User Interface
HMI	Human Machine Interface
LOA	Level of Automation
LSL	Lower Specification Limit
NPI	New Product Introduction
QSP	Quality System Production
SA	Situation Awareness
SAGAT	Situation Awareness Global Assessment Technique
SART	Situation Awareness Rating Technique
SASS	Situation Awareness Support System
SOP	Standard Operator Procedure
SPC	Statistical Process Control
SUS	System Usability Scale
USL	Upper Specification Limit

Abstract

Automation is the main focus in the current industry. In this world the role of humans changes from executing to monitoring. One of the ironies of increasing automation is that the requirements for operators increase, as operators have to take over when automation fails.

The ability of operators to take over is linked to the mental model the operator has of that situation, also called Situation Awareness (SA). However, increasing automation is linked to lower SA.

SA is a construct that measures understanding of the current situation and what this implies for future events. Automation has shown to have a negative effect on SA, taking over tasks previously done by the human.

In order to support SA of operators, a Human Machine Interface (HMI) can be used. A HMI can display information efficiently, maximizing the SA of operators. The current study looks at the effect of a HMI on SA with different types of HMI usage. A comparison is made between continuous monitoring versus occasional monitoring.

Results show no difference in SA measures between continuous monitoring and occasional monitoring. This implies that with the right HMI design occasional monitoring is sufficient, which will be necessary with increasing automation.

1 Introduction

The company Philips started out in 1891, with the manufacturing of light bulbs. In 1914 a research division was added and the production expanded to X-rays machines, radios and televisions. The production of shavers started with the introduction of the 'Philishave' in 1939. This production of shavers increased and was moved to a factory in Drachten in 1950. This factory started out with 30 employees, but quickly grew to a peak of 2500 employees in the year 1975. In 2008, Philips Drachten also became the research and development location for the consumer lifestyle department of Philips, with 1500 employees, of which 600 are engineers.

At Philips Drachten the whole production process for shavers is located, from research and innovation to the final assembly of the product. This is split up in different divisions, such as innovation, production and new product introduction (NPI), the division of the current research project. NPI is responsible for the implementation of new products into production, setting up and designing the production lines and assuring these production lines run smoothly and produce at a sufficient level of quality. Within NPI there are different departments responsible for the different aspects of the production line, such as cold forming, finishing and assembly. The current master thesis is done at the department of cold forming.

Philips has produced shavers since 1939 and the basic mechanism has not changed since that time. A shaving cap and cutter provide the basis of this mechanism, which are made with precision out of hard steel. A picture of the different components in a shaving head can be seen in Figure 1. A cutter rotates within a cap together with a so-called 'spider'. The spider lifts up the hairs so the cutter can cut them off at the base of the hair.

It is this production of steel caps and cutters that this master thesis focuses on, specifically looking at the role of operators in that process and at the information displays that operators have and what these should display. For a better understanding of the assignment first the production process of making these caps and cutters will be explained.

Figure 1: Different components of a shaving head.

1.1 Cold forming

Cold forming is defined as the mechanical operation in which a metal shape is permanently deformed into another shape without heating the metal. A punch and a die deform the material under pressure into the desired shape. Using this technique complicated steel shapes can be produced, with benefits in production speed, product quality, strength and material savings. Complicated products such as a cap or a cutter are made with a series of punches and dies, to go from flat steel to the cap or cutter in subsequent steps, sometimes requiring up to 10 steps or even more. To illustrate this process Figure 2 shows how a screw can be made using this technique.

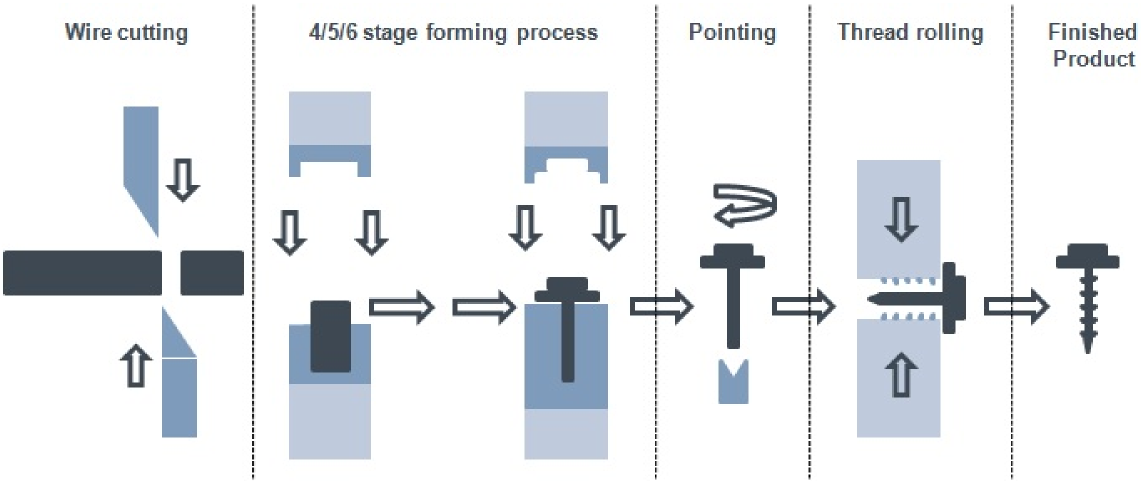


Figure 2: Illustration of the cold forming process.

`http://www.nedschroefmachinery.com/cold-and-warm-forming-technology/138/151/152/`

1.1.1 Production line

The cold forming of caps and cutters is not a standalone process but usually part of a production line consisting of a chain of different machines all completing a different part of the process. At the start of a cold forming line there are coils of steel that decoil and pass through the whole line. The first step is the press which stamps the material in the preferred form. After that the strip of steel passes through a washing machine cleaning off any irregularities before it can undergo the finishing process. In the finishing process the product is sharpened or hardened, depending on the product, and finally polished.

There are different types of production lines within Philips, depending on the type of material used and on the precision required for the end product. Some products are stamped under two different machines while others only undergo one press. Products can be stamped out and then refined using an electrical chemical process (ECP) or they can be made entirely through the ECP method. Some lines are so called 'flow lines' in which products are produced and assembled in one go and other production lines only produce the raw stamped parts in batches, which then serve as input material in other assembly lines. The advantage here is that flow lines reduce the amount of individual parts that have to be kept at a certain stock and the transportation costs that. With this method different parts are made and assembled

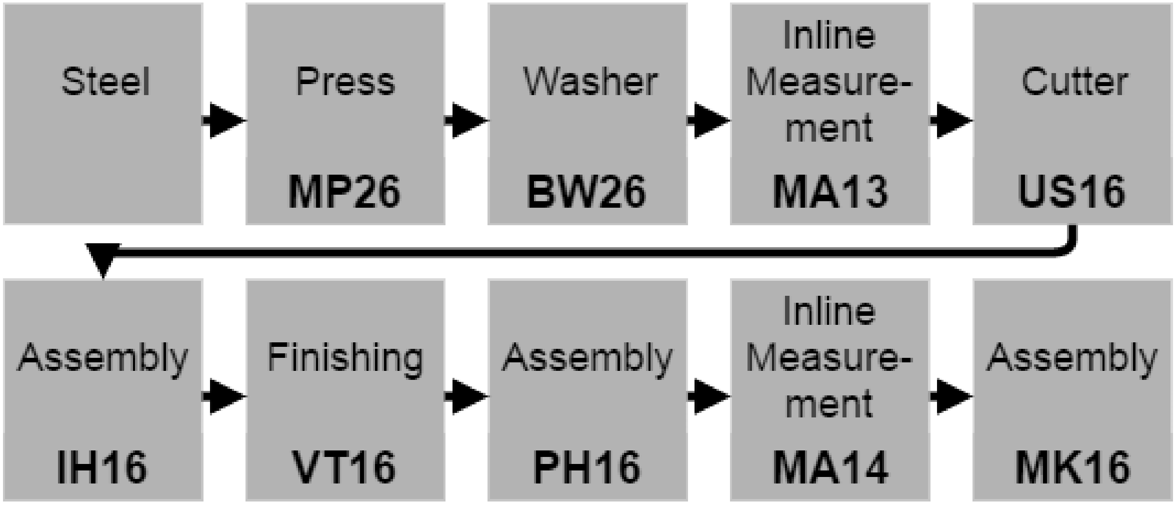


Figure 3: Layout of the cutter production flow line

at one place, compared to the old situation in which each part is made at a certain location and then assembled elsewhere. A flow line is the type of line that will be focused on during this thesis.

These 'flow lines' are relatively new in the domain of cold forming, and they are designed to reduce the stock of different parts of products according to the Lean Six Sigma principles (Brook, 2014). If one machine in the flow line stops the others will stop also when certain buffers reach their maximum capacity. These new lines also include new and sophisticated measurement methods that allow for the measurement of different product specifications within the production line. An overview of such a production flow line can be seen in Figure 3. Each square represents an element in that flow line, with its function at the top and the name of the machine at the bottom. The 'cutter' in the line cuts the stamped products out of the steel strip for further assembly.

1.1.2 Quality control

To assure product quality in production lines products have to be measured and referenced to some sort of quality criteria, which is set by the manufacturer. The ideal situation would be to measure every product that leaves the factory to be sure that it has no flaws and functions as intended. While this may be feasible to do for cars, when making cutters and shavers this would mean measuring up to four products every second. Instead of doing this, so-called statistical process control (SPC) is used to determine the process capability (Cpk) , which is part of the so called Six Sigma methodology (Raisinghani, Ette, Pierce, Cannon, & Daripaly, 2005).

Cpk works by specifying upper and lower specification limits (USL and LSL) that represent the range in which products are acceptable. When measuring individual products the mean and standard deviation of the measured products can be compared to these limits to calculate the Cpk value. The Cpk value is calculated by:

$$Cpk = min(Cp^{lower}, Cp^{upper})$$

where Cp^{lower} and Cp^{upper} values are calculated by the formulas:

$$Cp^{lower} = \frac{(\mu - LSL)}{3 * \sigma}$$

$$Cp^{upper} = \frac{(USL - \mu)}{3 * \sigma}$$

in which μ is the mean of the distribution and σ is the standard deviation of the distribution.

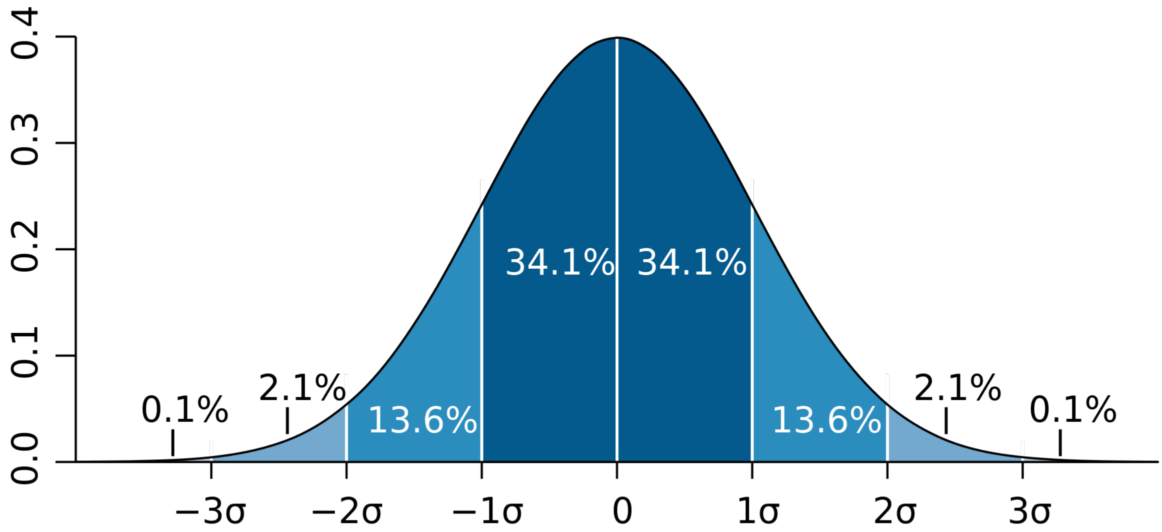


Figure 4: Normal distribution and six sigma

The Cpk value tells us something about the percentage of defects which can be expected when producing for instance a million parts this way. It is a number which reflects how well the 6 sigmas within a normal distribution (see Figure 4) fit between the USL and LSL. The normal distribution states that 99,7 percent of distribution falls within 6 sigma, and when this boundary is a certain distance from the specification limits predictions about product fallout can be made. Philips strives for a short term Cpk value of at least 1.33, representing 2,5 sigma on the long term which translates to a 99,40 percentage of good products which is 6200 defects every 1 million products. This value of 1.33 represents a distance of 4 sigmas between the process mean and the USL or LSL on the short term. When the baseline mean of a production line is determined, a single measurement can be compared to the mean and the USL and LSL of the process to determine a Cpk value. As long as this remains above 1.33 the process is considered capable and every product not measured is assumed to fall within the specification limits. In this way only one out of a few thousand products has to be measured to effectively control the quality.

1.1.3 Different stakeholders

In a production line, different parties work together to ensure the quality of the product that is produced. Before the product comes into production, a team of process engineers designs the process and looks how well different parameters of the product fall within the set boundaries. These boundaries are a result of function research, e.g. how well it shaves. Once the process is stable, the product can go into production and it is then the operators job to keep the quality of the products within the boundaries. Another party that is involved with production quality is maintenance. The tools used on the press and during finishing are subject to wear, and have to undergo maintenance once in a while. Maintenance ensures that the tools are brought back to their desired state, for instance by sharpening worn edges. The specifications of the tool have direct influence on quality, so precision is required.

The different stakeholders have one goal in common, namely optimizing the process, but different priorities within a production line:

- **Operators**
  The operators want to produce products within the specification limits, and want to produce the products smoothly.
- **Process engineers**
  Process engineers want to make the process stable by removing sources of variation within the process. When there is a quality issue they are the ones who have to find the source of the problem.
- **Maintenance**
  Maintenance wants to keep the tools in an optimal state, prevent wear of tools and fix tool problems before they occur.

These different stakeholders need to be taken into account in the current project.

1.2 Available data

The data generated by the quality control measurements is stored in a data system called QSP , which also contains data about production lines such as machine status, operator logs and materials used. However, with the new in-line measurement methods suddenly there is much more data available. In these new lines every product receives a so-called 'dot matrix code' (DMC) which is a unique identifier to which specific data can be linked. Using this DMC, one can take a single cutter or shaver that has been through this production line, read out its DMC and see all the product specifications related to that single product.

This results in a large amount of data that is available within a production line, of which only a small portion is used for quality control, because the Cpk values that are used don't require that many measurements.

1.3 Advanced Process Control

The large amount of data that is available is not yet used to its fullest potential. The current in-line measurement devices are used as a replacement for the old measurement systems into which a product had to be inserted manually. The only difference is that the measurement being conducted in-line is much more frequent (every product or 1 out of 30-40, depending on the in-line measurement tool) compared to the old method, which only measured one product every 4000 products. This will result in earlier detection of possible deviations in the product, but this process is still reactive to the situation, whereas ideally, the process is proactive.

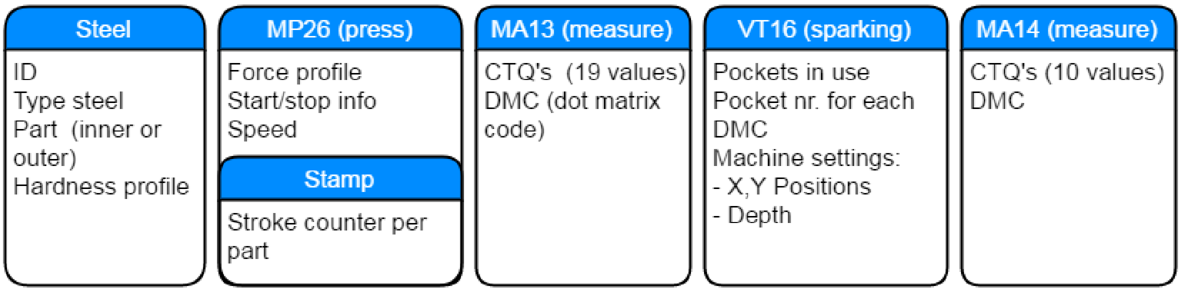


Figure 5: Overview of which data is available where in the production line.

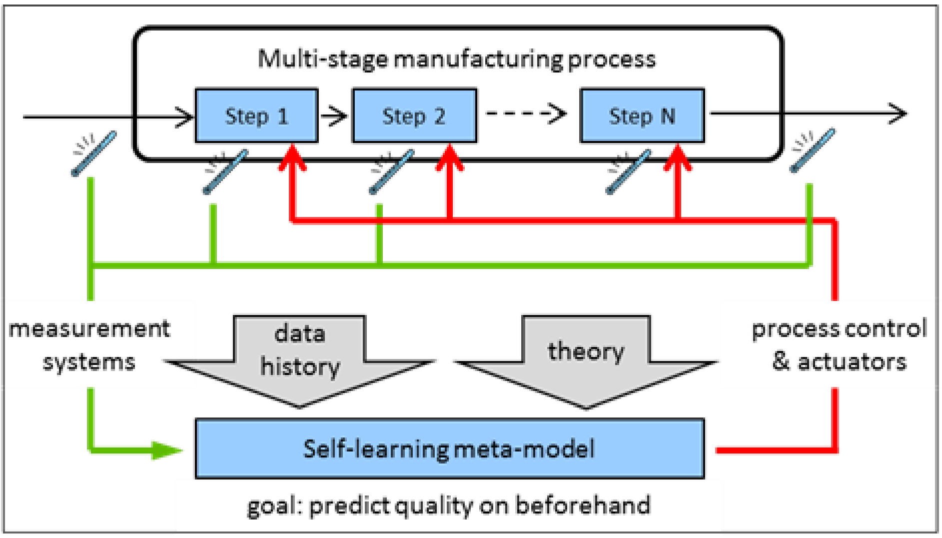


Figure 6: The idea of APC

Proactive process control means preventing the deviations from occurring by making adjustments when a specific critical to quality (CTQ) parameter drifts towards the specification limits or is expected to do so. Of course this ideal is not feasible in every situation, some product deviations are very sudden, such as tool failure, and do not have a gradual onset. But other product deviations may have a gradual onset, for instance resulting from tool wear or slight material property changes, which may very well be predictable.

The project of Advanced Process Control (APC) attempts such proactive control of the process by looking at the process data combined with measurements on the input material. Combining these two into one data set, relations between different parameters can be investigated allowing for predictions of CTQ parameters. Once such relations are known, one can predict CTQ parameters based on the measurements of the input material and adjust the machines accordingly when these predictions are outside the specification limits. For an overview of this process, see Figure 6. When these adjustments are then automated the production line will become highly autonomous, resulting in low workload required from the operator. Furthermore APC would make it easier to control the process more precisely, resulting in better overall quality and less rejection of faulty products.

1.3.1 Project setup

APC will look into proactive control of the cold forming process. As a start, it will try to make a test setup in which adjustments can be made in real-time, without stopping the process. This is done by implementing electrical motors in the tooling replacing manual actuation. Another part of APC in this first stage will be the implementation of new measurements in the production line before the material enters the process. Compared to the already installed measurements, sensing the input of the process is new. In the future it will enable the feed-forward loop, also visible in Figure 6. In practice, these input measurements will be thickness of the strip material, ambient temperature and Eddy Current properties.

Because these new measurements are not implemented yet they are not in scope of this thesis. For the current scope only a measurement of thickness will be simulated if no real data is available at the time of the experiment.

1.4 Assignment

It is expected that in future more and more data for every product will become available, including existing relations within these data. For this reason the human machine interface (HMI) that will present all measurements to the operators for that specific situation becomes more and more important.

The design of this HMI is the scope of this thesis. A HMI will be developed that displays the available data in an efficient way and matches the needs of the stakeholders using the HMI. The assignment will focus on the HMI in context of a specific production line, namely the cutter line. A HMI design and realization of this design will be delivered at the end of this thesis, and tested through an experiment. The focus will be on the presentation of the data rather than on which data to present.

1.4.1 Operator of the future

This design and implementation will be experimentally tested on the operators, that will be the main users of the HMI. The task of the operator has already seen a great shift in the last fifty years. Manual labour and assembly have been replaced by robots and flow lines that have changed the task of the operator. The main task of the operator nowadays is to monitor the performance of an automated process and to ensure the process keeps running. The project of APC also looks at automating the adjusting of machine settings, a task up to now done by the operator. APC might also predict the adjustments to be made, making the operator merely the executor. New technologies keep improving and reducing the workload of the operator and increasing the monitoring task of the operator. Eventually the operator might become obsolete, but current technology still has its flaws and does not always do what is intended. Operators still play a vital part in this process, but in five or ten years time a single operator might monitor multiple production lines from inside a control room. What are the implications for the operator in such a situation and how can the operator be supported optimally in such a situation?

1.4.2 Research question

As the task of the operator changes so do the requirements for that operator. This is also depicted as one of the ironies of automation (Bainbridge, 1983; Baxter, Rooksby, Wang, &

Khajeh-Hosseini, 2012), as automation increases in complexity, the expertise required by the operator to step in when automation fails also increases. Adding to this is the unsuitability of humans to act as monitors, with humans not being able to focus effectively on a display for over 30 minutes (Mackworth, 1950).

How well a human is able to monitor a process can be translated to the so called Situation Awareness (SA) that they have of that process (Endsley, 1995b). The human as monitor has a passive role and this can lead to a decrease in the SA of the operator (Endsley, 1996), which in turn leads to the out-of-the-loop performance problem (Endsley & Kiris, 1995), meaning that operators are less able to take over in case of automation failure. It is therefore interesting to look for possibilities of maintaining good SA whilst doing a monitoring task. This loss of SA can be related to change of feedback (Endsley & Kiris, 1995), meaning that different or less information is available to the operator. A HMI might be used to fill this gap by effectively presenting the information necessary to the operator.

The research question of this thesis is thus:

In what way does a Human Machine Interface add to the Situation Awareness of operators?

This research question requires a suitable HMI for experimental use and sub-questions can be formulated about the needs for this HMI and how these needs will be met. Another issue here is that in practice, operators also will not be able to monitor a certain HMI 100% of the time, because of distractions, coffee breaks or the HMI not being the central information source in a given process. To account for this there are a few sub questions which need to be answered:

- What are the needs of the operator for such a HMI?
- How can these needs be met and the data visualized efficiently?
- Is there a difference between continuous and occasional monitoring of a HMI?

1.4.3 Thesis outline

In section 2 the theoretical background for the research question will be given, followed by the requirements for the HMI in section 3. Afterwards the conducted experiment and the HMI designs used with the consequent results will be described in the sections 4, 5 and 6 respectively. Finally section 7 will discuss the results and conclusions will be drawn from it in section 8.

2 Theoretical Background

Chapter 1 describes the company of Philips in which production lines are already highly automated. At one end a sheet metal strip is pushed in to get a finished cap or cutter out on the other side, with minimum human effort in between. The humans job is to finetune the machines and supply the production line with materials and export the finished products to their next location. However as technology improves, manufacturers seek to automate even more, improving efficiency and cost price. With every automated part added to a production line the whole becomes more complex, requiring more knowledge from the person operating it to fix problems when they occur. This is what Bainbridge (1983) refers to as the irony of automation (Bainbridge, 1983); the more complex and advanced an automated system is, the more crucial may be the job of the human operator. An operator is asked to monitor a complex process to see if it functions the way it should do. Thirty years after the paper by Bainbridge Baxter et al. (2012) reviewed this issue and concluded that the ironies are still present today (Baxter et al., 2012). The take away message of that paper summarizes this:

> *"The more we depend on technology and push it to its limits, the more we need highly-skilled, well-trained, well-practised people to make systems resilient, acting as the last line of defence against the failures that will inevitably occur."*

With increasing automation, the role of the operator also changes in a different way; the technology places the human in the role of monitor, a role which does not suit us well. We are unable to maintain attention towards a source of information where not much happens for over 30 minutes (Mackworth, 1950). We don't pay attention when we are not kept busy which of course has implications for when something does happen. Studying the monitoring of processes requires some measurement of how well the operator is involved in the process and knows what is going on.

2.1 Situation Awareness

Situation awareness (SA) is the 'buzzword' used to describe a sort of mental model that operators have of their surroundings and of the device they are operating. As with many buzzwords the term is used a lot and also in different contexts, not always with the same meaning. To avoid confusion, it is therefore useful to first define the term.

One of the pioneers for SA is Endsley (1988) and she defines SA as:

> *Situation awareness is "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future"*(Endsley, 1988)

First of all SA does not directly infer good performance. Someone can have a good SA for example but lack the skill to perform the task well. Likewise good performance is possible without having a good SA. Developing and maintaining SA is seen as a important part of the activities of operators (Adams, Tenney, & Pew, 1995). An operator's SA has to be constantly updated because small changes in the process might have large consequences for the process as a whole. The term SA originally was used within the aviation domain, in which flight automation called for such a concept. Failure of automation was the cause of multiple crashes, which could have been prevented if the pilot had better known what was happening (Endsley, 1996). To illustrate the concept of SA in aviation a specific definition was given by Regal, Rogers, and Boucek (1988):

Situation awareness "means that the pilot has an integrated understanding of factors that will contribute to the safe flying of the aircraft under normal or non-normal conditions. The broader this knowledge is, the greater the degree of situational awareness" (Regal et al., 1988)

2.1.1 Levels of SA

A model of SA distinguishing different levels of SA was defined by Endsley (1995b), as seen in Figure 7. This figure is a representation of a decision making loop, with SA as a component used for decision making. Outside of this loop we see different factors, individual and task/system factors, which influence different elements of the decision making loop. The SA component of this model consists of three levels of SA.

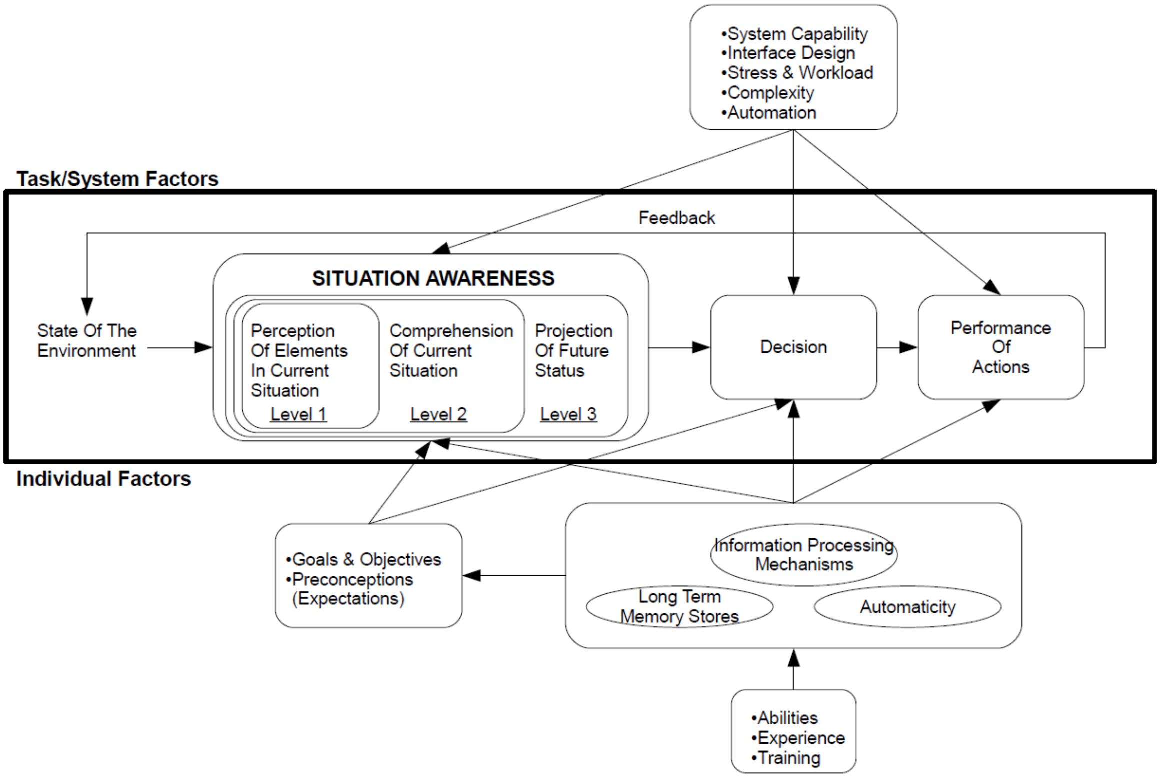


Figure 7: A model of SA in dynamic decision making. Figure from (Endsley, 1995b)

.

Level 1: Perception of elements in current situation Level 1 SA is about the perception of all the relevant elements in the environment of the operator. For a production line this would mean things such as the status of the production line, what product is currently produced, what is the status of all the warning lights and what info about the production line does the line monitor display?

Level 2: Comprehension of current situation Level 2 SA is about the comprehension of the current situation, putting everything together from level 1 SA and translating this relative to the goals that the operator has. Inferring from the information gathered at level 1 this means that the operator knows whether everything is going correctly and knows which action to take if there is an alarm or some machine has stopped working. Novice operators may have the same level 1 SA as experienced operators but may lack the experience to comprehend the whole situation as well as a more experienced operator.

Level 3: Projection of future status Level 3 SA is about the projection of future events that will occur, at least in the near future. This can be achieved through combining the level 1 and 2 SA. This type of SA is important to be proactive rather than reactive when making decisions. In the production line example there may be some process positions that often causes errors at a certain point for which anticipating actions may be taken to solve it quickly when they occur. Research in the aviation domain has shown that experienced pilots also spend quite some time anticipating for possible events that may occur in the future (Amalberti & Deblon, 1992).

2.1.2 Errors in levels of SA

With SA being part of a decision making framework, as seen in Figure 7, can also result in errors in decision making that can be attributed to a lack of SA on a certain level. Possible reasons for errors on different levels of SA were identified by Endsley (1995b), of which a few will be listed here.

Level 1 An operator may simply fail to perceive information that is necessary for good SA, this may happen for multiple reasons. The information might be hard to detect or hard to discriminate due to the physical nature of the information. Information that has to be seen visually can be obscured or hard to notice due to non-optimal system design. Certain information might only come to light when an error occurs, for instance a person driving a car might not realize it's slippery until he or she loses control over the car (Rasmussen, 1986). The failure to perceive important information is also related to the ability to divide attention across tasks (Martin & Jones, 1984). Human limitations often lead to level 1 SA errors.

Another factor which affects that is stress, which leads to attentional narrowing (Janelle, Singer, & Williams, 1999). Attentional narrowing means that attention is less divided and more devoted to one thing, for instance the problem at hand. If thus a problem occurs an operator might focus on that and fail to notice other important information.

The last reason for errors in level 1 SA identified by Endsley is misperception of the information. Perceiving for instance a 3 as an 8 is a simple example of such an error.

Level 2 Errors in level 2 SA are a result of the inability to translate observations and perceptions of level 1 SA to a meaning relative to the goals of the operator. This may be because the operator lacks the knowledge to translate these observations and perceptions into meaning, or focuses on the wrong cues because of that. Another cause for errors may be the method in which level 2 SA is usually acquired. A common strategy is to select a certain model based on previous experiences and match that to the current situation. This way, operators make assumptions about what observations mean, linking for instance a specific warning light to a certain scenario. In aviation flight crews were also found to obtain a substantial amount of information after they had made a decision (Mosier & Chidester, 1991). Selecting the wrong model for the situation can lead to errors in level 2 SA, causing operators to wrongly interpret observations, also possibly due to the confirmation bias (Nickerson, 1998). When no model is available for the current situation level 2 SA has to be developed in working memory. This can lead to errors in level 2 SA due to limitation of working memory.

Level 3 Errors in the projection of the future status are also related to the model that an operator has of the situation. A situation might be understood clearly but if the implications of that situation on the future are not know or incorrect, level 3 SA can be lacking or incorrect.

2.2 Automation

Before the relation between SA and automation can be made this section will describe the research field of automation and the issues there. Automation is defined as "a function that was previously carried out by a human" (Parasuraman & Riley, 1997). As technology improves, more functions that were previously carried out by humans are taken over by automation, mainly because automation is thought to prevent human error and improve efficiency (Lee, 2006). Automation also redefines the role of humans in complex systems and imperfect automation has many pitfalls (Parasuraman & Riley, 1997). This section will review some of these pitfalls and the different roles automation can have within a complex system.

2.2.1 Automation use, disuse, misuse and abuse

Terms often used within the automation design include the use, disuse, misuse and abuse of automation (Lee, 2008; Parasuraman & Riley, 1997). This is a useful distinction used to guide automation-related research and these different uses of automation will be explained next.

Use of automation refers to the operator using the automation for tasks they would otherwise perform manually. With the choice of the operator to use some kind of automation many different factors play a role, such as workload, trust in automation and risk (Lee, 2008). Automation is usually expected to decrease the workload of the operator and make the task easier, but this is not always the case. Easy tasks are usually easy to automate and hard tasks are left with the operator, making the easy task easier and the hard tasks harder (Wiener, 1989).

Misuse of automation refers to situations in which operators use an automation system that performs poorly. This is also called over-trust or over-reliance on automation, and results in the operator trusting the automation and failing to notice conflicting signals or use automation in situations that are not appropriate. For instance Riley (1994) performed a study on automation use with pilots and students and found that even though the task had nothing to do with aviation, half of the pilots kept on using the automation when it failed, compared to almost all the students who turned it off. Parasuraman and Riley (1997) gives another real-life example of misuse of automation in which the GPS signal of a large ship failed, and the crew failed to notice this until 24 hours later the ship ran aground.

Disuse of automation refers to the situations in which operators do not use automation while it could enhance performance. This is basically the counterpart of misuse because disuse mostly comes from under-trusting automation (Lee, 2008). Reasons for not trusting an automation system come from false alarms. If an automated system gives a lot of false alarms the operator will at some point negate this alarm and also fail to notice if the alarm is correct. Responses to such alarms will become slower if the alarm has a high false alarm rate compared to the hit rate, therefore decreasing the effectiveness of the alarm (Parasuraman & Riley, 1997).

Abuse of automation refers to situation in which automation is designed and implemented without looking at its effect on operators and the overall process. This often occurs when designers believe that adding automation will reduce errors and increase efficiency, while it only creates new problems. Parasuraman and Riley (1997) gives an example of this where a weight-on-wheels sensor has been added to planes, to make sure the plane is on the ground before pilots can reverse the trust of the engines. If this sensor fails however, pilots can't reverse the trust of the engines, disabling their ability to break properly. This abuse of automation is one of the main reasons of the ironies of automation depicted by Bainbridge (1983). The irony here is that when more automation is added to a system, the system becomes more complex and the skills needed to monitor the system and ensure its correct functioning increase with it. These ironies have been reviewed and concluded to still be present this day (Baxter et al., 2012).

2.2.2 Levels of automation

The previous section shows that there are many pitfalls with humans and automation and that automating more is not necessarily better. Also in assembly work, the human is found to be an important factor not replaceable by robots or automation (Pfeiffer, 2016). What remains is a team performance of automation and humans, assigning functions to each party to get the optimal performance.

What this assignment of functions between humans and automation is, is usually referred to as the level of automation. A global distinction between levels of automation was offered by Sheridan (1992) with three levels: manual control (everything is done by humans), supervisory control (human is supervisor and can take over control at any time) and fully automatic control (everything is done by automation). This distinction soon proved to be too simple and more levels were distinguished by Sheridan (2002), as can be seen in Table 1. Similarly in the aviation domain Billings (1997) proposed different levels of automation, as seen in Table 2.

When comparing Table 1 and 2 we see that they are very similar, defining similar degrees of human and automation roles. Table 1 makes more distinctions at a higher level of automation with a difference between the automation of the informing of humans.

Table 1: Levels of automation from Sheridan (2002)

Level	**Description**
1	the computer offers no assistance; the human must do it all
2	the computer suggests alternative ways to do the task
3	the computer selects one way to do the task and (see Level 4)
4	executes that suggestion if the human approves, or (see Level 5)
5	allows humans a restricted time to veto before automatic execution, or (see Level 6)
6	executes automatically, then necessarily informs the human, or (see Level 7)
7	executes automatically, then informs the human only if asked
8	the computer selects the method, executes the task, and ignores the human

Table 2: Levels of automation from Billings (1997)

Level	**Description**	**Role of Automation**	**Role of Humans**
1	Direct manual control	No automation is used.	Human manually controls all functions and tasks.
2	Assisted manual control	Primarily manual control with some automation support.	Human manually controls with assistance from partial automation.
3	Shared control	Automatic control of some functions task	Humans control some functions/tasks
4	Operation by delegation	Automatic control when directed by human to do so.	Human provides supervisory commands that automation follows.
5	Operation by consent	Full automatic control under close monitoring and supervision.	Human monitors closely, approves actions, and may intervene.
6	Operation by exception	Essentially autonomous operation unless specific situation or circumstances are encountered.	Human must approve of critical decisions and may intervene.
7	Autonomous operations	Fully autonomous operation. Human not usually informed. System may or may not be capable of being disabled.	Human generally has no role in operation, and monitoring is limited.

Table 3: Levels of automation adapted from Endsley (1999) by OHara and Higgins (2010)

Level	**Description**	**Agent responsible***			
		Mon	**Gen**	**Sel**	**Imp**
1 - Manual	The human performs all tasks.	H	H	H	H
2 - Action Support	Automation assists the operator in performing the selected action, although some human control actions are required.	H/A	H	H	H/A
3 - Batch Processing	Although humans generate and select the options to be performed, they are completed automatically.	H/A	H	H	A
4 - Shared Control	Both the human and the automation generate decision options. The human still retains full control in selecting which option to implement; however, carrying out the actions is shared.	H/A	H/A	H	H/A
5 - Decision Support	The automation generates a list of decision options from, which humans select, or they may generate their own options. Once the human has selected an option, it is implemented automatically.	H/A	H/A	H	A
6 - Blended Decision Making	The automation generates a list of decision options, selects one, and carries it out with human consent. The human may approve of the automations option or select one from among those generated by the automation or the operator. The automation l then completes the human-selected action.	H/A	H/A	H/A	A
7 - Rigid System	Automation presents a limited set of actions from which the human selects one; humans cannot generate other options. Automation implements the selected actions.	H/A	A	H	A
8 - Automated Decision Making	The system selects and implements the best option from a list of alternatives it generated (augmented by alternatives suggested by the human).	H/A	H/A	A	A
9 - Supervisory Control	The system generates options, selects one to implement and carries out that action. The human monitors the system and intervenes if necessary. Intervention requires that the human select a different option from those generated by automation or by the human).	H/A	A	A	A
10 - Full Automation	The system carries out all actions. The human is out of the control loop and cannot intervene.	A	A	A	A

* Abbreviations: Human (H), Automation (A), Monitoring (Mon), Generating (Gen), Selecting (Sel), Implementing (Imp).

2.2.3 Functions of automation

Tables 1 and 2 focus on process control which can be split between humans and automation but there are more functions in which automation could play a role. Four general functions can be identified within the context of automation (Parasuraman, Sheridan, & Wickens, 2000):

- *Information acquisition* refers to the sensing of data in the environment. Automation may do this by reading out certain sensors at a given interval.
- *Information analysis* refers to the processing of the data by making inferences, averaging or predicting.
- *Decision and action selection* refers to the selecting of a decision from among different alternatives.
- *Action implementation* refers to the execution of the decision made, replacing the hand or voice of the human.

Endsley (1999) combined the levels of automation and the different functions that automation can fulfill to ten levels of automation, as seen in Table 3. This framework distinguishes ten levels of automation looking at the role of humans and automation on the different functions that automation can have.

2.2.4 LOA at Philips

When applying the framework of Table 3 to the situation of the production line at Philips subject of this project, we see that there is still quite a low level of automation there. It can be best described as a combination of level 2 and 3 automation, because almost all of the information analysis and action selection is done by the operator. As soon as the production stops the automation stops and waits for the operator, also no suggestions about what to do are given at that time. If the automation will also start to do adjustments to the process, as planned within the project of APC, the level of automation will change. This change may have implications for performance and SA of operators because the role of the operator shifts more towards a monitoring role. Consequences of this role on performance and SA will be discussed in Sections 2.3 and 2.4.

2.3 Automation and monitoring

The effect of automation on human performance has been studied intensively in the aviation domain. An extensive list of issues with aviation automation has been made by Funk et al. (1999), identifying and ranking all the different issues found with automation. OHara and Higgins (2010) organised the top issues into categories, with 'Impaired Monitoring and Situation Awareness' being one of those categories. Fifty percent of the top issues were part of this category, showing that impaired monitoring and SA is a big issue within automation.

First, the factors responsible for impaired automation monitoring will be discussed of which three can be identified (OHara & Higgins, 2010):

- Automation's reliability - Operators are less likely to monitor automation they consider reliable.

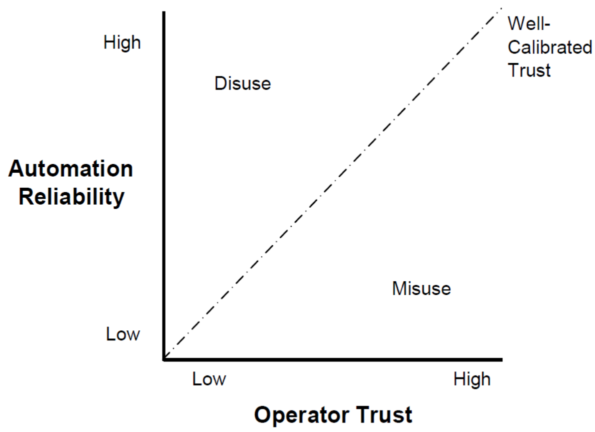


Figure 8: Relationship between operator trust and automation reliability

- Operator's workload - Operators are less likely to monitor automation when they are busy with other tasks.
- HMI design - Operators are less likely to monitor automation when the HMI does not offer an easy means to do so.

2.3.1 Reliability

The automation's reliability indicates how often the automation is correct, and this is closely related to the operators trust in automation. However, the amount of trust an operator has in automation does not always match the reliability of the automation. If the operator trust in automation exceeds the automation reliability it results in misuse of automation (called overreliance or overtrust), and disuse of automation occurs when automation reliability exceeds the operator trust (called underreliance or undertrust) (Lee & See, 2004). This relationship is illustrated in Figure 8.

One study looking at the effect of automation reliability on operator trust and reliance (usage of automation) was done by Ross, Szalma, Hancock, Barnett, and Taylor (2008). Participants performed an identification task in which they were operating a simulated unmanned ground vehicle to identify locations of terrorists, civilians and explosive devices. They were shown videos after which they had to decide if they had seen one of these things. An automation aid was available that could give advice on the decision they had to make but this aid had different levels of reliability, from 75% to 99%. They were not told about the aid's reliability and were also not forced to use or follow the aid. Results showed that as the aid's reliability increased the operators reliance and trust in that aid also increased, meaning that participants could infer the aids reliability from experience with the system.

If an automation is reliable the operator is more likely to trust it and, consequently, pay less attention to monitoring and checking the automation. This would be no problem if the automation was 100% reliable, but in practice no automation ever is. There are always components that can fail or situations that occur for which the automation was not designed and is not able to handle.

2.3.2 Workload

The effect of workload on performance is illustrated in Figure 9 although this figure and the underlying model is simplified (OHara & Higgins, 2010). If workload is too low this can result in loss of vigilance, simply boredom, and workload is to high there is overload which causes deterioration of performance. For optimal performance the operator has to have a workload that is somewhere in between these extremes.

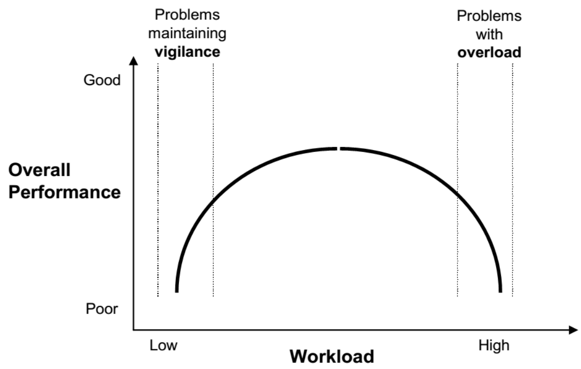


Figure 9: Relationship between workload and performance

The effect of automation on workload is mixed and complex (Parasuraman et al., 2000). Automation can reduce the overall workload of a task, by taking over parts of the task. In aviation automation was found to have this effect (Wiener, 1989). As is also described in 2.2.1, automation makes easier tasks easier but harder tasks harder. Automation in aviation was also found to make the task of landing an airplane more complex under abnormal circumstances (Wiener, 1989). Pilots usually revert to more manual modes of flight under abnormal circumstances and first having to disable or override automation in that case poses additional workload.

Another study however found no reduction of workload with automation compared to no automation (Endsley & Kiris, 1995). A navigation task using different levels of automated assistance were compared, but workload did not differ between any of the different levels, ranging from manual to full automation. A possible explanation here was the type of workload shifting from execution to monitoring where the load remained similar (Billings, 1991). Monitoring automation is usually not the only task an operator has, and if other problems occur that require attention and result in workload the monitoring part will become less important.

An example of this can be seen in the example of the crash of a Boeing 757 in Cali, Columbia in 1995. The pilots used a flight management system (FMS), which is an automated navigation- and flight-control system. It was so very reliable and accurate that it could take care of the whole flight except for landing and takeoff. When approaching the airport the air traffic control gave the pilots an option of a direct approach instead of flying around. To do this the pilots had to reprogram the FMS or continue flying manually. They chose to reprogram the FMS but entered the wrong beacon by entering 'R' and commanding the FMS to fly there. However there was another beacon starting with 'R' in the wrong direction and the plane was automatically steered in that direction. However when manually overriding the FMS destination, the FMS entered a mode in which it did not automatically avoid terrain. The pilots, unaware of this mode of the FMS and unaware that the plane was not going the way they intended, were too late to prevent the plane from crashing into the mountains when they noticed the mistake.

2.3.3 HMI design

The previous example also shows the importance of HMI design on monitoring a system. In the example it was unclear to the pilots what the mode of the automation was and to which beacon they were headed. Monitoring is improved when the behaviour of automation is clear in the HMI and the attentional demands for this information are minimized (Parasuraman & Riley, 1997). A HMI should be designed to provide necessary information about what the automation is doing (OHara & Higgins, 2010). A HMI can make the monitoring task easier and result in better SA, all of which will be discussed further in Section 2.7.

2.4 SA and automation

Impaired monitoring as a part of the category responsible for fifty percent of the top issues in automation has been discussed in the previous sections and this section will focus on the part of impaired SA as a reason for automation issues.

Supported by many examples of failures that occurred with automated systems Endsley (1996) gives a summary of the problem:

> *"Situation awareness, a person's mental model of the world around them, is central to effective decision making and control in dynamic systems. This construct can be severely impacted by the implementation of automation."*

A review of commercial aviation accidents concluded that 88 percent of the accidents involving human error could be attributed to problems with SA (Endsley, 1995a).

2.4.1 Out-of-the-loop

Loss of SA awareness has been linked to the out-of-the-loop performance problem that is a negative consequence of automation (Endsley & Kiris, 1995). This problem entails that operators of automated systems are handicapped in their ability to take over manual operations in event of automation failure. Effects of automation on performance and SA as a result of being out-of-the-loop have been found (Carmody & Gluckman, 1993). The study looked at automating aviation-relevant tasks and requiring the human to take over when the automation failed. Endsley and Kiris (1995) explains this relationship between impaired SA and the performance decrement:

> *"Operators who have lost SA may be slower to detect problems and require extra time to reorient themselves to relevant system parameters in order to proceed with problem diagnosis and assumption of manual performance."*

Endsley categorized the evidence for this claim into three categories: (1) loss of vigilance and increase in complacency associated with the assumption of a monitoring role; (2) change from active to passive processing of information and (3) change in the type of feedback provided to the operators.

The first two have been discussed in previous sections and include many of the uses of automation such as use and disuse. The change in feedback refers to the state of the system, of which an operator no longer receives direct feedback, because the operator is the one doing the action, but rather has to be informed about what the automation did. The design of displays may occlude information from the operator, show processed information instead of raw information about the process or fail to give information about whether the automation succeeded in completing an action.

2.4.2 LOA effect

The out-of-the-loop performance problem is also affected by the level of automation (LOA). Lorenz, Di Nocera, Röttger, and Parasuraman (2001) looked at the effect of low, medium and high level of automation with a system that had a reliability of 90%. The task consisted of operating a Cabin Air Management System (CAMS) with assistance of an automation aid. The performance of the participants during the 10% of the trials that the automation failed was evaluated and results showed impaired performance in the medium and high levels of automation compared to the low LOA condition. Taking over when automation fails is thus easier with lower levels of automation.

The relation between LOA and SA was studied by Willems and Heiney (2002), who looked at the effect of automated decision support on the behavior of air controllers. They found that SA was usually higher with a low level of automation compared to no automation or full automation. However when the workload of the task was high, SA was worse under full automation conditions compared to the other conditions. The authors also found that operators spend a lot of time transitioning between the two displays the task used, which might be one of the problems. They suggested to integrate these displays better to resolve this problem, something to keep in mind when using multiple displays.

Not only the LOA but also the function of automation has an influence on SA (Kaber, Perry, Segall, McClernon, & Prinzel, 2006). As described in Section 2.2.3 there are three functions that automation can fulfill: information acquisition, information analysis and action implementation. An experiment with an Air Traffic Control (ATC) task was conducted by Kaber et al. (2006). Participants performed this task with an automated aid. The function of this aid switched between the three functions mentioned above, and the impact on SA was measured. Of these three, automation of information analysis was found to have a big effect on SA (Kaber et al., 2006).

2.5 SA and HMI design

The previous sections have shown the importance of a good HMI within automation, linking the automation and the human together. A HMI is the main tool to keep the operator informed about the actions of the automation. This importance is stated by Liu, Nakata, and Furuta (2004):

> *"In fact, incidents and accidents have actually been caused because operators either did not understand the goals or working states of the automation or had a misperception of their own tasks since they could not construct an up-to-date mental model of the current system state in time. Therefore, it is very difficult for operators to adapt to manual control if the automatic systems fail. This out-of-the-loop problem has often been attributed to inadequate design of the human-machine interface.*
> *Designing an effective user interface for an automatic system is very challenging*

due to complexity of the controller algorithm, the amount of information that is potentially relevant and the complex interaction of the underlying process. However, only by visualizing the working of automatic systems in an interface can an operator track what an automatic system is doing, why and how it is doing it, and what it will do next. That is, operators can build an appropriate mental model of the system."

HMI design is directly linked to the operator's mental model of the system, which is basically what SA entails. HMI design is thus an important factor influencing SA in automated systems, and effects of HMI design are indeed reflected in SA. For instance, for power plant operators a functional display was compared to a traditional schematic display. The functional design resulted in higher SA with the same task (Tharanathan, Laberge, Bullemer, Reising, & McLain, 2010).

Also HMI designs that follow the Ecological Interface Design (EID, which will be explained in Section 2.7) have been shown to improve SA compared to traditional displays (Burns et al., 2008; Kim, Suh, Jang, Hong, & Park, 2012).

In the research on this topic the influence of HMI design on SA is seen when comparing an old design with a new design, which are totally different from each other. The effects of details in HMI design, such as different visualizations for the same information, on SA have not been studied widely, nor the effect of HMI design on human-automation interaction in general (OHara & Higgins, 2010). A study by Bowden and Rusnock (2015) has shown no effect of details in HMI design on SA. They compared the type of information presentation (numeric versus graphic) and the information arrangement (functionally grouped versus spatially mapped) and found no significant effects on SA or task performance. More studies doing similar research were hard to find in the literature.

2.6 SA measures

To measure SA, different methods are available, of which two of the best known methods will be used in this thesis. These two methods are known as SART and SAGAT, which will be described next.

2.6.1 SART

The Situation Awareness Rating Technique (SART) was developed by Taylor (1990) as a tool to measure SA of pilots. It is a simple and subjective questionnaire that is administered after the period over which the SA needs to be determined. It uses ten different dimensions across three domains to measure SA, which are listed in Table 4.

The participants are asked to rate each dimension on a seven point scale with 1 equals low and 7 equals high. These ratings are combined to calculate a measure of SA, where the scores of the ten dimensions are summed over three domains, D (attentional demand), S (attentional supply) and U (understanding). SA is then calculated with the following formula:

$$SA = U - (D - S) \tag{1}$$

SART is one of the best known and most thoroughly tested subjective SA measures (Jones, 2000). Its validity and sensitivity have been shown by numerous studies (Jones, 2000), showing sensitivity to task difficulty and operator experience as well as discriminating between different

Table 4: Different SART dimensions from Taylor (1990).

Domain	Dimension	Description
Attentional demand	Instability of the situation	Likeliness of situation to change suddenly
	Complexity of situation	Degree of complication of situation
	Variability of situation	Number of variables that require attention
Attentional supply	Arousal	Degree that one is ready for activity
	Concentration of attention	Degree that one's thoughts are brought to bear on the situation
	Division of attention	Amount of division of attention in the situation
	Spare mental capacity	Amount of mental ability available for new variables
Understanding	Information quantity	Amount of knowledge received and understood
	Information quality	Degree of goodness of value of knowledge communicated
	Familiarity with situation	Degree of acquaintance with situation experience

displays. Other studies found SART not to be sensitive to display manipulation (Endsley, Sollenberger, Nakata, & Stein, 2000; Satuf, Kaszkurewicz, Schirru, & de Campos, 2016). SART is still being used in recent studies as a measure of SA (Satuf et al., 2016; Kim et al., 2012; Salmon et al., 2009; Naderpour, Lu, & Zhang, 2016), but is usually used alongside other methods such as SAGAT. Correlation between SAGAT and SART has not been found in multiple studies (Endsley, Selcon, Hardiman, & Croft, 1998; Naderpour et al., 2016), so they might measure different concepts of SA.

2.6.2 SAGAT

The Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1988), is a tool to assess SA based on the different levels of SA defined in Section 2.1.1. SAGAT is situation specific as questions for each level of SA have to be developed for each new situation. To use SAGAT a simulation is frozen and the displays blanked after which the participant has to answer a few questions. By doing this multiple times during a simulation a SA score can be calculated from these questions by adding up all the correct questions. For this experiment a total of 15 questions is used, 3 per SA level, to determine the SA score.

Similar to the SART technique, SAGAT is one of the most used SA assessment techniques, often used alongside SART or other methods (Jones, 2000). SAGAT is also used in a few recent studies (Satuf et al., 2016; Naderpour et al., 2016; Wulf, Rimini-Döring, Arnon, & Gauterin, 2015; Bowden & Rusnock, 2015) with different domains, other than the aviation domain in which it was developed. This method might be the most widely tested measure of SA, and has shown to have good levels of sensitivity and reliability (Endsley, 2000).

However, SAGAT was found not to be sensitive to display manipulation in one study in which different methods of visualizing data were compared (Bowden & Rusnock, 2015). Satuf et al. (2016) used SAGAT to test the effectiveness of an Ecological Interface Design (EID) and the SAGAT scores were higher for this design compared to the old design. These results however were not significant because of a small participant group (n=10). Another study found that SAGAT level 2 and level 3 scores were significantly higher for a system with an additional Situation Awareness Support System (SASS) than without this additional system (Naderpour et al., 2016).

2.7 Design guidelines

One of the design principles used for HMI development is the Ecological Interface Design (EID) framework. EID is a theoretical framework for designing complex human-machine systems (Vicente & Rasmussen, 1992). This framework is based on the skill, rules and knowledge taxonomy (Rasmussen, 1983). This taxonomy distinguishes three types of behaviour to which EID should provide optimal support:

1. Skill-based behaviour - To support interaction via time-space signals, the operator should be able to act directly on the display, and the structure of the displayed information should be isomorphic to the part-whole structure of movements.

2. Rule-based behaviour - Provide a consistent one-to-one mapping between the work domain constraints and the cues or signs provided by the interface.

3. Knowledge-based behaviour - Represent the work domain in the form of an abstraction hierarchy to serve as an externalized mental model that will support knowledge-based problem solving.

These principles are quite abstract and are usually translated to more concrete design guidelines that can be followed when making user interfaces. The rest of this section will discuss some of these design guidelines.

Design guidelines have been around since the computer made its appearance, with the earliest coming from 1976 (Cheriton, 1976). These design guidelines contain rules to follow when designing user interfaces but are formulated in a broad way. Two of the best known lists that formulate design guidelines are in table 5.

As can be seen these lists have items that are very similar and some of the items might appear to be self-evident. These guidelines leave room for interpretation and can lead to conflicts when trying to adhere to every rule. Many of these rules however have their bases in cognitive psychology. It makes sense to look at design guidelines from a cognitive psychology perspective, and this is what is done by Johnson (2013). The following sections will review these areas of psychology and the implications for interface design briefly.

2.7.1 Perception by expectation

Our visual system is heavily influenced by what we expect to see. This expectation can be based on our previous experience, the context, or the goals we have at that time. An example of experience can be seen in for example a picture of a collection of dots. Once you see a meaningful image in this collection of dots you can no longer look at that picture and see only a collection of dots. Context alters our perception by changing the meaning of words

Table 5: Best known lists of user interface design guidelines, from (Johnson, 2013)

Shneiderman (2010)	**Nielsen and Molich (1990)**
- Strive for consistency	- Consistency and standards
- Cater to universal usability	- Visibility of system status
- Offer informative feedback	- Match between system and real world
- Design task flows to yield closure	- User control and freedom
- Prevent errors	- Error prevention
- Permit easy reversal of actions	- Recognition rather than recall
- Make users feel they are in control	- Flexibility and efficiency of use
- Minimize short-term memory load	- Aesthetic and minimalist design
	- Help users recognize, diagnose, and recover from errors
	- Provide online documentation and help

depending on what words are around it but also what we see can be affected by what we hear, smell or see. Our senses do not work independently of each other. Goals also affect our perception by focusing only on items that are related to that goal. If a website is scanned for information, items not relevant to the goal are not only ignored but often not even noticed, a phenomenon referred to as 'perceptual filtering'.

Design implications This leads to the consistency guideline; all users should interpret the user interface correctly and in the same way. Ambiguity should be avoided and standards should be kept. For instance, a lot of icons have an associated meaning, coming from the use of computers and smart-phones. Although these are mere conventions, not following them can lead to users misunderstanding the interface. Another implication is to take into account the different goals users may have with the interface, and that the information needed for those goals should be easily found.

2.7.2 Perception of structure

We are wired to perceive visual information as shapes, figures and objects instead of as disconnected lines, edges and areas. A group of German psychologists in the early twentieth century investigated how our visual perception works and came up with what they called 'Gestalt' principles. These principles are a description of factors that influence our perception, which have to be taken into account with interface design.

Proximity The relative distance between objects affects our perception of whether or not objects appear to be grouped. Objects that are near to each other relative to other objects appear grouped. The way objects are placed in a grid determines whether we see columns or rows and buttons on an interface that are together appear as a group. Grouped buttons in turn suggest some sort of similarity in function.

Similarity The same holds for similarity; things that are similar to each other appear to be grouped. If there are different kinds of buttons, the ones that are similar to each other will appear grouped.

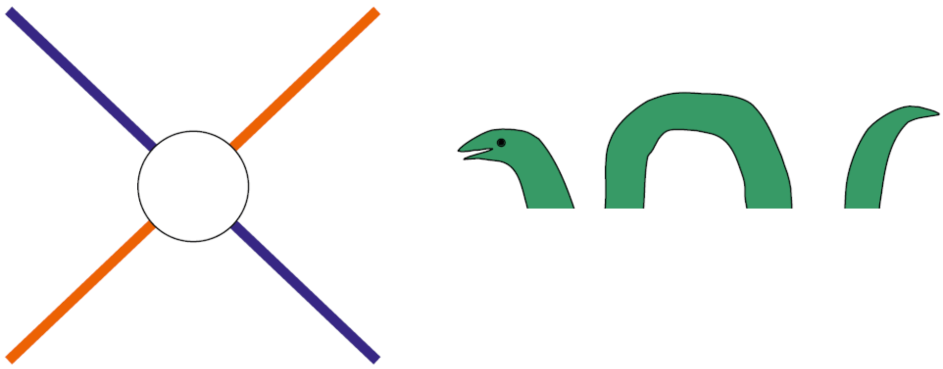


Figure 10: The gestalt principle of continuity, we fill in missing data. Figure from Johnson (2013).

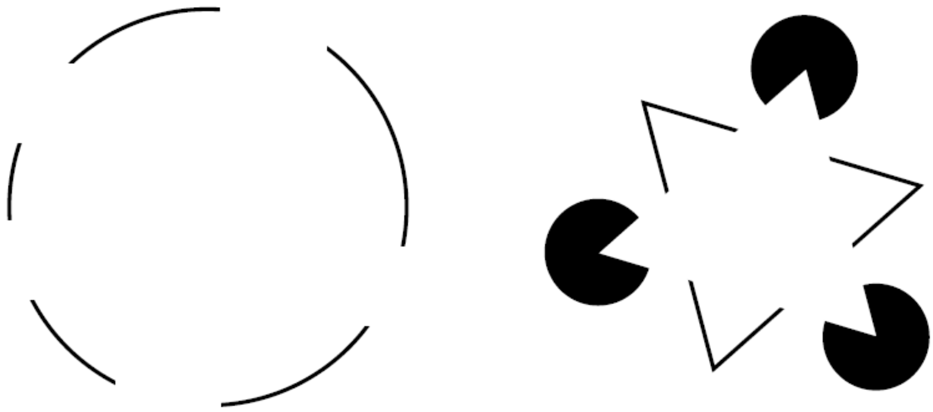


Figure 11: The gestalt principle of closure, partial objects are perceived as whole. Figure from Johnson (2013).

Continuity Our visual system tends to fill in gaps to perceive a continuous form rather than disconnected segments. An example of this can be seen in Figure 10. Within a user interface sliders with a handle in the middle work according to this principle.

Closure Related to the continuity principle is closure, the principle that our visual system automatically closes open figures so that we perceive whole objects. In Figure 11 a circle and two triangles between three black circles can be seen even though this is not what is visible. This principle is used to represent a collection of objects such as a stack of documents in which only the edges of the objects behind can be seen.

Symmetry The principle of symmetry is also related to the previous two principles in seeing objects rather than individual shapes. For instance, we interpret two overlapping diamond shapes as two diamond shapes and not as two tetris figures touching each other or an eight shape with a square in the middle, see Figure 12. A combination of these three principles allows us to see three-dimensional scenes in two-dimensional images.

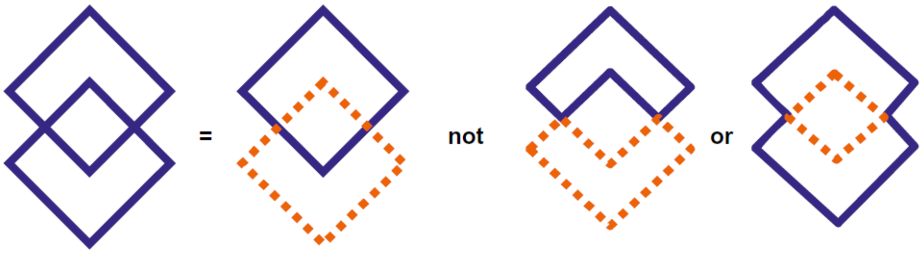


Figure 12: The gestalt principle of symmetry. Figure from Johnson (2013).

Figure/ground The following principle is about the structuring of the visual field by our visual system into a figure and a ground. When objects are overlapping the smaller one will be seen as the foreground (figure) and the larger one as background (ground). This principle is used to display something (e.g. a photo) over the current page by dimming that page and using it as background. This can be beneficial for keeping people orientated within the user interface.

Common fate This principle concerns moving objects, compared to the other ones discussed so far. The principle states that objects or items that move together appear grouped. Common motion can be used to show relationships between items, for instance when representing something on a grid with different dots. When selecting a property which is true for some of the dots one can let these dots move together to indicate they are the ones belonging to that property.

Design implications The major implication for interface design following from these gestalt principles is to prevent unwanted relationships. When the interface is finished the developer should look at it with each of the principles in mind and see whether there are any relationships visible that were not intended.

2.7.3 Reading

Reading is not a natural ability, compared to speaking language (Sousa, 2014). This means that our brain is not wired to learn reading even though we do it daily. Reading is a skill that is learned just as playing a musical instrument is. This also means that not everyone is equally good at it, with skilled readers using different mechanisms to read compared to novice readers. Brain research has shown that there are two modes of reading; automatic, context-free reading and conscious, context-based reading. The first works by recognition and requires experience but does not require a lot of effort, leaving room to analyze what is read. The second mode requires more effort as we consciously read a text, burdening working memory and reducing comprehension. This mode costs more energy and makes us tired, so if possible this should be avoided.

Given a skilled reader which uses automatic reading there are multiple things that can revert the reader back to the conscious mode. When encountering unknown words we have to analyze the meaning based on the context of the word. A font that uses different shapes for letters or uses all caps makes it hard to read automatically, as well as fonts that are too tiny. Text that is displayed on a noisy background also interrupts automatic reading, as well as text that is centered instead of left-outlined. Our eyes are used to the move back to the position we started reading a line, only then one line down; if the start of the line we need to read next is on a different location, we need to adjust for it and this interrupts automatic reading.

Design implications This implies a few design implications, all focused on supporting automatic reading, not interrupting it:

- Use a standard font, with normal size
- Use simple words, avoid 'geek' speak
- Use a simple background, maximizing contrast
- Don't use centered text

Besides that it is often possible to reduce the amount of reading needed within an application by removing obsolete text or using icons to convey information.

2.7.4 Color vision

The way our color vision works also has implications for interface design. The workings of our vision system will not be explained, but rather the limitations that it has will be discussed. First of all our vision is optimized to detect edges, not to detect brightness or exact color. Multiple factors influence our perception of color, such as the background on which the color is presented or if there is shading involved, also see Figure 13.

This also makes us bad at discriminating colors from each other, especially when they are pale colors, small patches or if they are separated by distance as can be seen in Figure 14.

On top of that roughly 1 out of 20 people have some form of color blindness, making it hard for them to distinguish between certain colors. Furthermore not every display shows colors in the same way, the display on which the interface is designed might show colors different than the display on which the interface will be used. Colors that are easy to distinguish when designing might not be distinguishable upon use of the interface. A good way to check if the colors used are easy to distinguish is to look at it in grayscale. If there are no problems in a grayscale image color-blind people will also be able to distinguish the colors used.

Design implications When designing an interface one has to keep in mind the limitations of the visual system. This means using colors that are easy to distinguish, also for people that have color blindness. There are tools available to select which colors to use in for instance a graph, such as `http://colorbrewer2.org`. Also, if possible, do not rely on color alone and accompany color with either text or a symbol. Legends in graphs should be accompanied by sufficiently large color maps, making it easier to match the legend to the graph.

2.7.5 Peripheral vision

The center of our visual field, also called the fovea, has a much better resolution than the rest of our visual field. About 50% of the visual cortex in the brain is devoted to process the input of this area which is about 1% of the size of our retina (Lindsay & Norman, 2013).

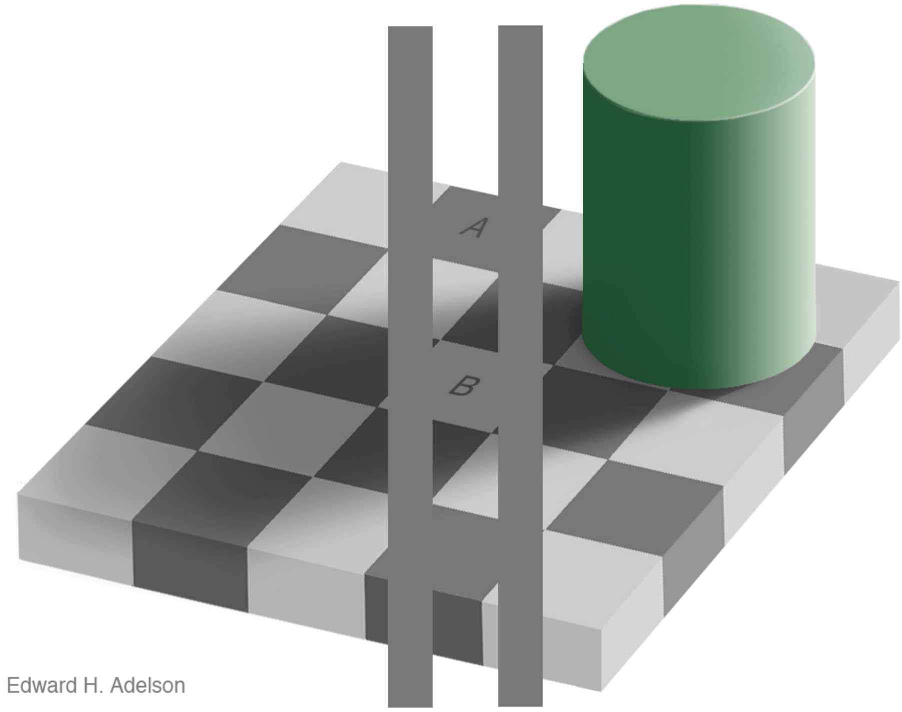


Figure 13: Area A and B have the same shade of grey but appear different because of shading. Created by Edward H. Adelson.

This means that we are not able to see many details in the our peripheral field of vision, as can be seen in Figure 15.

We are thus able to see some details in our peripheral vision but small changes we will not notice. However we are good at detecting motion in our peripheral field, drawing attention to that location to inspect what we see there.

Design implications This too has implications for interface design. Warning messages for instance should be near the spot where the person is currently looking. Upon pressing a button in the bottom right corner the warning message should also be displayed near that button otherwise it may be missed. If it is critical that a message is seen by the user movement such as a wiggle can also help for attracting attention.

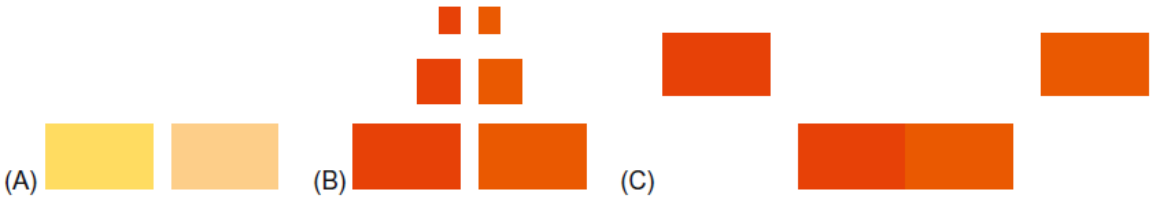


Figure 14: Factors that influence the ability to distinguish colors: (A) paleness, (B) size, (C) distance in between. Figure from (Johnson, 2013).

Figure 15: Illustration of the resolution of peripheral vision. The size of letters required to be able to read them when focused at the center.

2.7.6 Memory and attention

Memory can be divided into short-term memory and long-term memory. The first is a storage for things we are currently working on and the latter for everything else. The capacity of short-term memory, often called working memory, was initially determined at 7 plus or minus 2 items, such as words, (Miller, 1956), but later this was adjusted to a limit of 3 to 5 (Cowan, 2001). Recent research however leaves the assumption of short-term memory as a storage but links this to attention (Jonides et al., 2008). We can divide our attention between a limited number of items and get distracted easily as other things draw our attention away from what we are currently doing. Likewise our long-term memory is imperfect, we need repetition to prevent a memory from decaying and many external factors influence how well we can remember something.

The bottom line here is that our short-term memory is limited and our long-term memory is imperfect which has implications for interface design. Another aspect of attention is that it is driven by the goals that we have at that time, depending on the goal certain keywords or pictures will draw our attention. This is related to 'Perception by expectation' which was discussed earlier.

Design implications The main implication of this is that an interface design should not place unnecessary burdens on memory. Different modes for instance can be very useful in an interface but users will often forget what the current mode is they are operating in. When using modes the current mode must be made clear to the user at all times. When using search functions the term users entered to search needs to remain visible and when giving instructions on how to handle a certain problem these instructions must remain visible during the process.

An interface should be viewed in light of the different goals that that user could have with it and then look if the right items draw the users attention in that case. Users should be guided towards their goal.

2.7.7 Recognition and recall

Recognition is easy and recall is hard, because our brain has a preference for recognition. When recalling we have to search in our memory for the answer that we seek. What we basically do is reactivating a neural pattern of that particular memory, which is difficult and can easily fail. With recognition we do not have to search our memory, we either do or do not recognize a picture, face or a certain song. A face will produce a certain neural pattern which we either have or haven't seen before. Of course when a new face looks exactly like a face you have seen before you might recognize it because the neural pattern is similar to something encountered before.

Design implications Our ease of recognition can be exploited by using pictures to convey function in interfaces. A lot of pictures already have a recognized function because of all the previous programs using certain pictures to indicate a function. A gear wheel icon for instance is used to indicate the settings function. Thumbnails, small screenshots of a page can be used for navigation, as is often done in web navigation.

2.7.8 Learning

There are multiple factors that influence the speed at which a user will learn to work with an application.

Operations are task-focused, simple and consistent Operations are the actions users have to perform in the application to get to the goal they have. These operations should not be unnecessarily complex, but get the user to the goal quickly. Applications should also be as simple as possible, more functionality means more complexity and if users will not use the functionality it should not be added. Being consistent means that a certain function in a program will have the same effect on different objects.

Vocabulary is task-focused, familiar and consistent Similarly the vocabulary used in an application should be focused on the task and not on the technology. Users don't care about how something is implemented but on what they want to do with the application. The terms in an application should also be familiar, meaning that computer jargon should be avoided. The vocabulary should use terms that are already in use by the users of the application, for instance task specific vocabulary such as machine names and abbreviations for control limits. Furthermore only one term should be used for one concept, multiple terms for one concept or multiple concepts mapping to one term only cause confusion.

The risk in using the application is low Users of the application should not be afraid to explore and try-out the application. However when the application is designed in such a way that it is dangerous to just go and try-out some buttons users will not explore the options. Errors should be hard to make and if they are made they should be easy to recover from.

Design implications To promote learning of an application the application should provide the functionality needed but not more, the vocabulary used in the application should make sense to the user and be consistent throughout the application and exploration of the application should be promoted by for instance deactivating commands and buttons that are not applicable in that situation or should not be used at that time.

2.7.9 Time requirements

Cognitive processes take time and this has to be taken into account in applications. More important even is the fact that we have some standard of perceived responsiveness, if something takes longer than a certain amount of time we consider it to be slow. For instance to make the brain understand that some action caused a certain effect these two need to be within 140 milliseconds from each other. Another example is the maximum gap we expect in a conversation, which is about 2 seconds. When interacting with an application this works similarly, an application should do what we asked within 1-2 seconds or else give some sort of progress indicator. No response from the system within this time will make users impatient. For a complete list of the time requirements that humans have, see Johnson (2013).

Design implications The time requirements are important to acknowledge to make users perceive the user interface as responsive. The following guidelines should be followed:

- Acknowledge users actions instantly, even if it's just a message saying it will take time
- Use busy indicators to indicate if the application is doing something
- Animate movement smoothly and clearly
- Allow users to abort an action if it takes to long
- Use progress indicators to show how long something will take

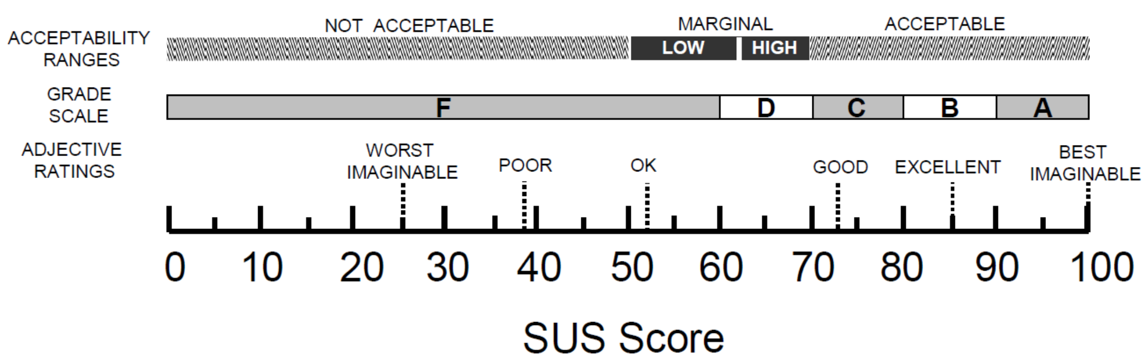


Figure 16: Comparison of different ratings compared to the SUS score. Figure from (Bangor et al., 2009).

2.8 System usability

A HMI design can also be evaluated on usability, to get a general measure of how useful an interface is. The System Usability Scale (SUS) is a quick and general tool to get such a measure of the user's subjective view of the usability of the system. It can be used with different technology such as hardware, software, websites or operating systems (Brooke, 2013). It has proven to be an effective and valid measure of usability even with small samples of 8-12 users (Tullis & Stetson, 2004).

The SUS consists of 10 statements on which the user can rank his agreement on a five-point scale, from strongly agree, to strongly disagree. Of these 10 items, five are phrased positively and five negatively. The scale positions correspond to a value from 1 to 5 and for the positively phrased questions the statement score is scale position minus 1. For the negatively phrased values the statement score is 5 minus the scale position. The sum of the statement scores are then multiplied by 2.5 to get to a score between 0 and 100. SUS scores relate to subjects ratings in terms of adjectives such as 'good', 'poor' or 'excellent' (Bangor et al., 2009). This is illustrated by Figure 16.

3 Requirements analysis

To investigate the effect of a HMI on SA an experiment with different HMI designs is required. These HMIs have to be applicable in the situation of the operator working at Philips and also be useful to other people who want to gain insight into that specific production line, so for that reason a requirements analysis will be conducted. To learn about the needs of the different parties, first the task of different parties has to be clear. This will be done by doing a task analysis; observe operators and see first hand what it entails to be an operator and what an operator has to do during a shift. Similarly engineers, when solving a problem with a production line, have certain methods they use when tackling that problem. To make the HMI useful for all parties they will be interviewed about their methods and how a HMI could assist them. Besides that a function analysis has to be conducted, to get a clear view of what functions the HMI has to fulfill. These functions combined with the task analysis and constraints that the different parties have will serve as a guideline for the design process of the HMI.

Scope Whilst different production lines are visited this thesis will focus on one of those specifically. This production line will be more closely examined and the HMI will be created with the needs of operators of that production line in mind. This is done because production lines differ a lot from each other and it is out of the scope of the project to make a general HMI that would be applicable in every production line. The one chosen is the cutter flow-line, which is a new line that includes the in-line measurement systems talked about in sections 1.1.1 and 1.2. Within this line the scope is limited to the cold-forming process because the project of APC is focused on that part.

3.1 Task analysis

This section will analyze the task of the different parties, because a clear view of the tasks of the user is necessary to get an answer to the questions: "What does it have to do?" and "How well does it have to do it?".

3.1.1 Operators

Operators at Philips work in 8-hour shifts, and in teams of two or three persons, depending on the production line. To get a global view of the task of the operator different teams at different production lines were observed and interviewed. This created the opportunity to see first hand what operators have to do and to ask them about their work. Also the project of APC was explained and the operators questioned about what they would want to know in a situation in which additional measurements are available and automated adjustments are made to machines within the production line.

Line monitors To give a task description of the operator, knowledge about the line monitors is necessary. The work of operators depends heavily on line monitors present at the production lines. An image of this line monitor can be seen in figure 17.

The line monitors give information about the production line during the current shift (of 8 hours) and the machine status during that time. Product fall-out during the shift and per machine can also be seen, with a refresh rate of 30 minutes. The line monitors were designed

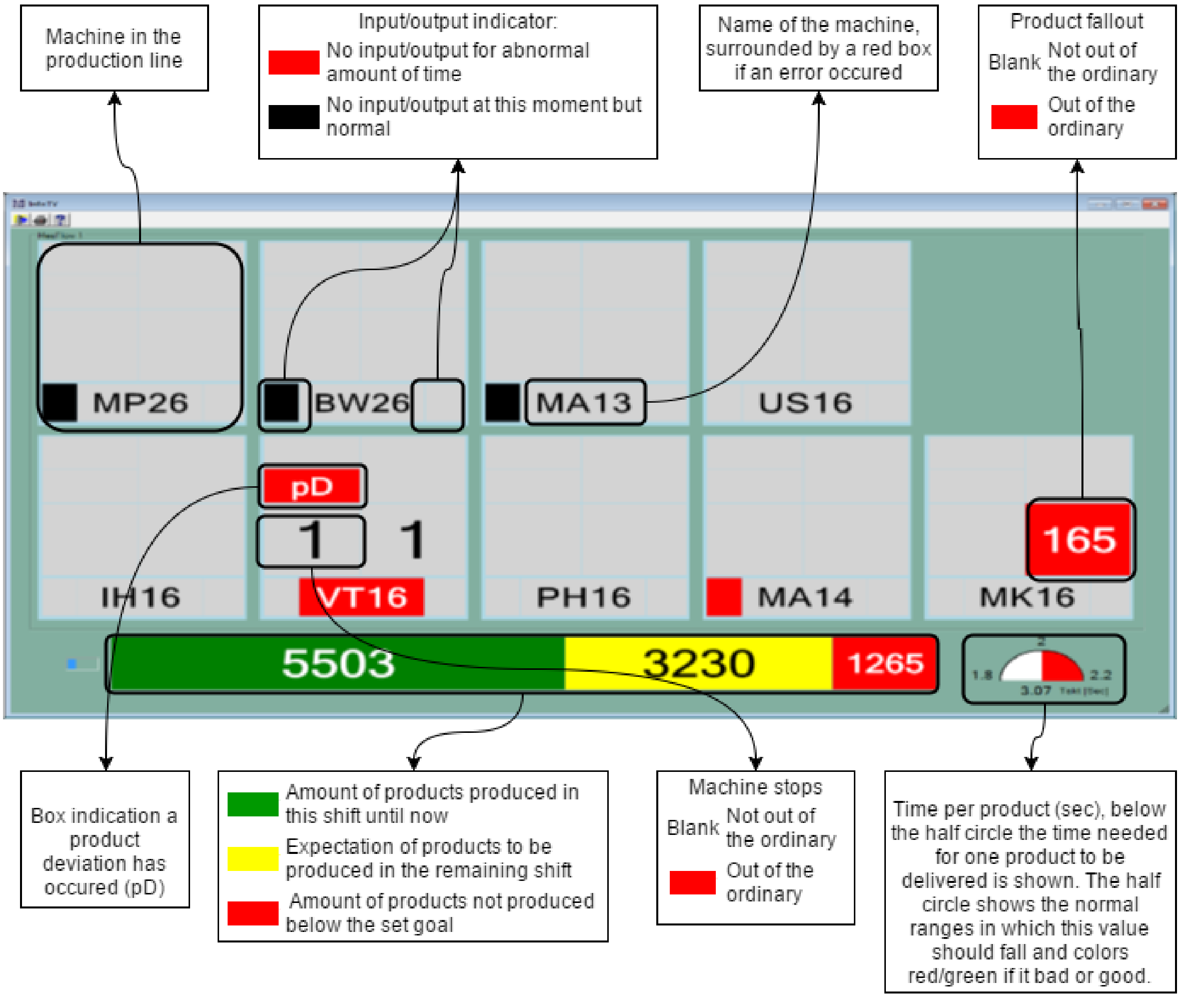


Figure 17: Line monitor used in production line

to give an overview of the production line to the operators and to gain rapid insight into the status of that line.

Task description The main task of the operator is to ensure that the production line keeps running. The main tool the operator uses for this is the line monitor described in the previous section. In addition to the line monitor every machine has indicator lights, which can be either red, orange or blue. The operator constantly has to check the screen and/or the machine indicator lights to check whether the process is still running. Experienced operators also start to use the sound of production lines as a means to detect errors as a correctly functioning production line has a distinct repetitive sound from which deviations are easy to distinguish (at least for the trained ear).

Once the operator has noticed the line has stopped the operator has to determine the reason for the stop. The production line will stop if:

- One of the machine produces an error (red box under the name)
- A product specification is exceeded (red pD)

- A machine runs out of input parts or if output is full

If a product specification is exceeded the operator has to find out the reason behind the deviation. This can be a faulty part somewhere in one of the machines, an error in the measurement system or a machine setting which needs adjustment. The reasons for product deviations are usually very similar and can mostly be avoided by preventive adjustments to machine settings for instance.
An operator can see the live measurement results from the in-line measurement systems and detect if a product CTQ is close to the USL or LSL. An image of the display of these live measurements can be seen in Figure 18.

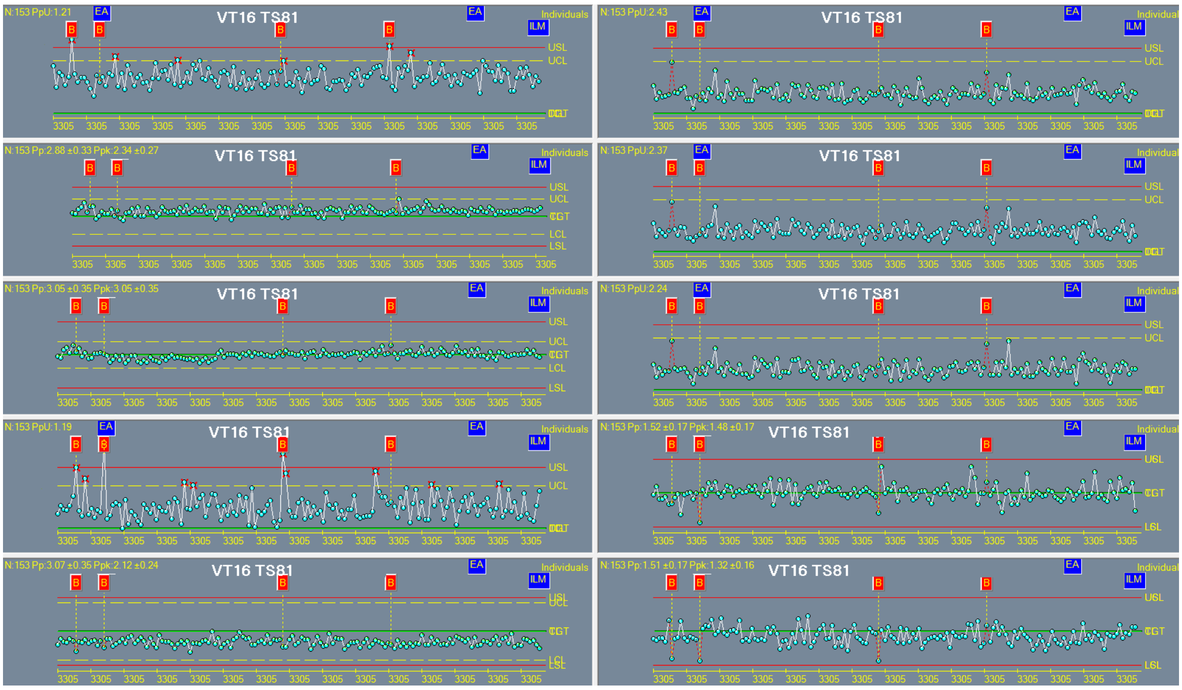


Figure 18: Overview of the display available to the operators.

A zoomed in version of one of the CTQ's is shown in Figure 19. All the blue dots represent measurements that are plotted on a time axis with shift numbers. A shift number consists of the week number (the first two digits) and the shift. The number 3305 represents the fifth shift of 8 hours in week 33, so that would be Monday from 8 am until 4 pm.

If one of the machines produces an error the operator has to go to that specific machine to figure out what the error is. Most machines have its own user interface on which an error message is displayed that gives information about the error. This usually gives the operator enough information to know where he has to look or what he has to do to fix the error. If the error is generic then the operator has to look if he can spot the problem or do some manual runs with the machine to see if he can reproduce the problem. The operator can also manually take a faulty product and scan its code to see where it has been, in what machine it was at what time and which part of the machine it was processed in.

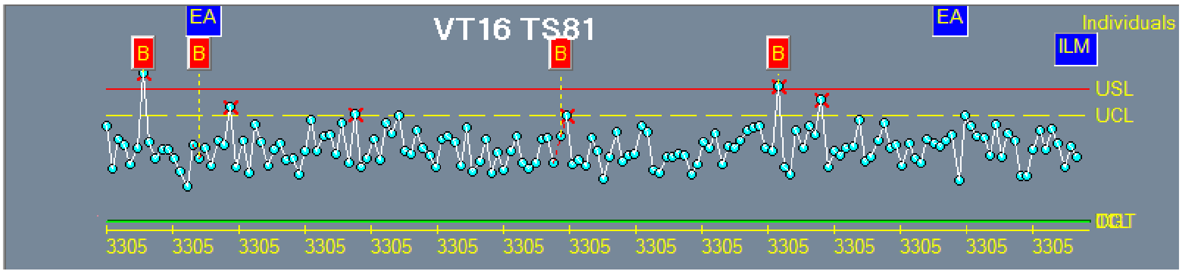


Figure 19: Live measurement data for a single CTQ. The red lines are the USL and LSL, the yellow dotted lines are the control limits; when exceeded, action has to be taken.

3.1.2 Engineering

Besides the operators there are also quality engineers involved in the process of optimizing the production lines. They are responsible for setting up these production lines and organizing them efficiently but also monitor the quality of the products coming from these lines. If a problem occurs that an operator cannot fix the quality engineer has to find the root cause of the problem and implement a fix. Also if the products are within specification limits the quality engineer attempts to find sources of deviation to further stabilize the process.

3.2 Function analysis

With the project of APC the implementation of automated machine adjustment is investigated. Some form of feedback about this automation has to be implemented in the form of a HMI, also with the possibility to override the system manually.

Operator The operator's job is to keep the production line running and he is mainly interested in the actions necessary to get it running again or to prevent a failure. The APC project can produce a lot of information but the operator only wants to know the practical implications of these relations. Questioning the operators resulted in the following list of needs and functions they would like the HMI to have:

- Minimal information, only display necessary information.
- Display actions the operators needs to take.
- Prioritize these actions if there are multiple.
- Indicating when to do reference measurements.
- Show calculated time until next action.
- Give an overview of errors; what are the errors and how often do they occur?

Engineering Questioning the quality engineers about their methods revealed that they use the same QSP interfaces to gain insight in the data. Once they have formed a hypothesis about certain relationships within the data they test this by importing the data into a statistical package to verify that hypothesis. Requirements of the quality engineers for the HMI include:

- Display relations between parameters
- Display deviations in the process, what parameters are responsible for a lot of deviation?
- Display the predicted CTQ's compared to measured CTQ's.
- Check if the input material meets the delivered specifications of the manufacturer.

3.3 Overview

The requirements for operators and quality engineers differ and are difficult to combine in one view. Operators require a minimal view and only the information useful to their goals while quality engineers want insight in the production line, looking at possible relations between parameters that may or may not be interesting. The two needs may be met by using different modes in the HMI, one for operator use and one for the engineers.

4 HMI Design

The HMI designed is closely linked to the project of Advanced Process Control (APC) that looks at the automatic adjustment of CTQ values. If in the future automated adjustment is implemented the HMI should give information about this process in a way that enables the operator to act quickly and correctly in case of any errors with the system.

Because of the relation between APC and the HMI, the experimental design is made with this functionality in mind. The aspect of proactive control and the requirements that come with that aspect are not implemented for the final design. Proactive control requires that predictions can be made about output parameters and these relations are currently unknown. The inclusion of realistic predictions require the development of sophisticated models, which goes beyond the scope of the research.

The requirements of the HMI in Section 3 were gathered based on inclusion of proactive control. Therefore, the requirements have a slight mismatch with the HMI design. Nevertheless, the HMI is prepared to include the future proactive control, when the predictive models are available.

4.1 Design process

Before implementing the design, operators and engineers were asked about their preferences. This resulted in the requirements listed in Section 3. These requirements were however all functional so a design was sketched up from scratch. During the design process, the operators were actively involved. Operators were asked for feedback on provisional implemented designs, so they could give feedback about the layout and items that were visible in the design. This process of designing, implementing, evaluating and adjusting was done twice before the design was finalized.

4.2 Implementation

The HMI is implemented in Python (version 3.4.4) together with the design package Qt (version 5) which is available under a Python binding platform known as PyQt5. These are the main components used in the implementation of the HMI. Specifically, a bundled version of Python known as WinPython is used. Together, this allows for an easy design of a graphical user interface (GUI) , with drag and drop functionality to position different items in the HMI. Another tool used that builds upon this platform is pyqtgraph, a plotting library that enables the fast and responsive real time graphs.

4.3 Final design

A screenshot of the final design can be seen in Figure 20. This section will describe the make-up of this design and the reasons for the layout and the different components.

4.3.1 General layout

As a starting point the screen was split in three panels:

1. A top panel, spanning all across;
2. A side panel on the right;

Figure 20: Final design of the HMI

3. The main panel, serving as a graph area.

The layout of the main panel keeps changing depending on the amount of graphs the user opened. The layout of the other panels remains static although the content can change. This setup was chosen because some information is not always necessary to display. By dividing the screen into different parts the whole layout does not have to change all the time. This way, information in the static panels can always be found in the same place, while keeping the HMI flexible. A dark color theme was chosen to keep the HMI soothing to the eyes, even if used in darker surroundings.

4.3.2 The right panel

This panel consists of two components, the action list seen in Figure 21a and the so called 'status widget' seen in Figure 21b. The action list is made to list the actions that the HMI has for the users, displaying information about an event that needs special attention or some problem that needs to be fixed. The items in this list are prioritized based on a priority rating which is currently not visualized because the simulation only covers one action at a time. Actions can be toggled to indicate that they are finished and those actions are grayed out, but remain visible in the list. The items are prioritized based on the time added when grayed out. This was done to provide some history of events and can be scrolled through to see what the previous actions.

The status widget is an overview of a CTQ value for all of the nine cutter legs. As a background image here a cutter is displayed as it is produced by the press and the legs are numbered corresponding to the internal numbering used at Philips. This is done to make the abstract CTQ values concrete and show what the CTQ is about. On each leg the mean value of the CTQ is displayed as a green circle which is positioned on a slider. The yellow

(a) Action list

(b) Status widget

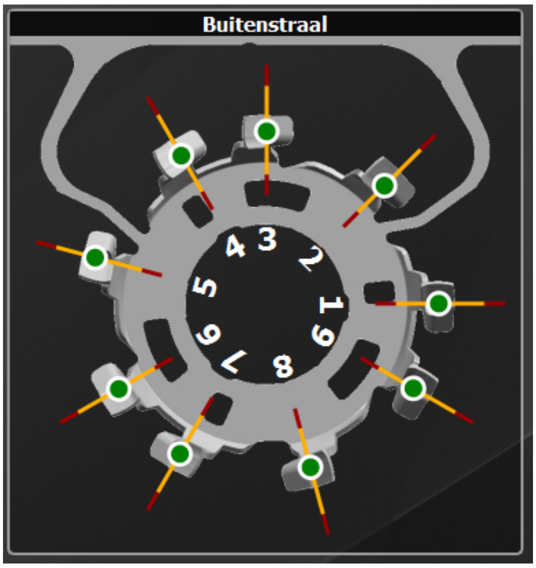


Figure 21: Subparts of the HMI: I

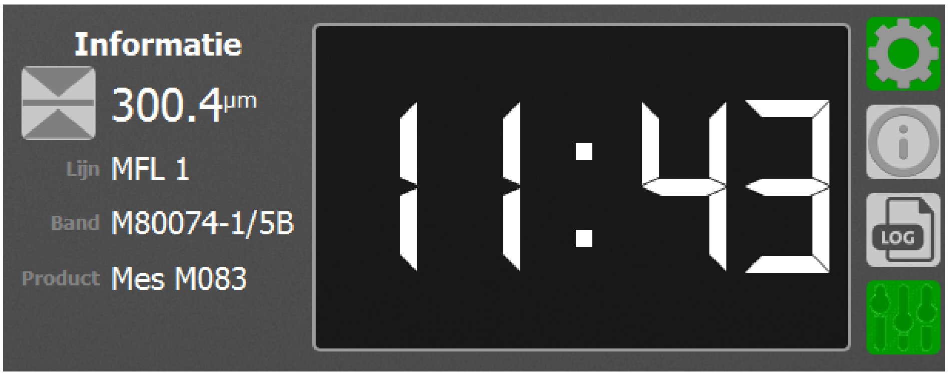
and red areas on these sliders correspond to the control limits and specification limits which are also visible in the graphs, see Figure 23. The mean value is calculated based on the last 30 measurements received. When the CTQ value of a certain leg will drift towards the upper control limit the circle will slide along outwards in this status widget.

4.3.3 The upper panel

The upper panel also consists of two components, the information panel seen in Figure 22a and the key-control shown in Figure 22b. The information panel shows general information about the production line, such as the name of the production line, the material used and the product that is manufactured. Also the thickness of the material and the current time is displayed. On the right side of this panel are a few buttons; one to get to the settings menu, a button to get information about how to use the system, a log file of the automatic adjustments and a switch to enable/disable the automatic adjustment mode. For the buttons an icon was used to convey the message, to keep the amount of text on the HMI limited. An information pop-up is shown when hovered over the button to make their purpose clear if this is not intuitive. Coloring of the buttons is used to convey an additional message. A green color on the settings button means that there currently is a connection with the data server and the automatic adjustment mode is green to indicate that automatic adjustment is currently enabled. These buttons were disabled during the simulation.

The key-control shows a visual representation of the keys that are inside the stamp. Currently operators can set the keys by turning a screw which slides the key back and forth. This is now done when the press is stopped, but with APC adjustments will be automated, enabling adjustments while producing. The working of a key is similar to that of a slide bar and that's why this representation of the keys is chosen. A key is shown for each leg (and CTQ value) that can be adjusted. The position of the slider corresponds to the position of

(a) Information panel



(b) Key-control

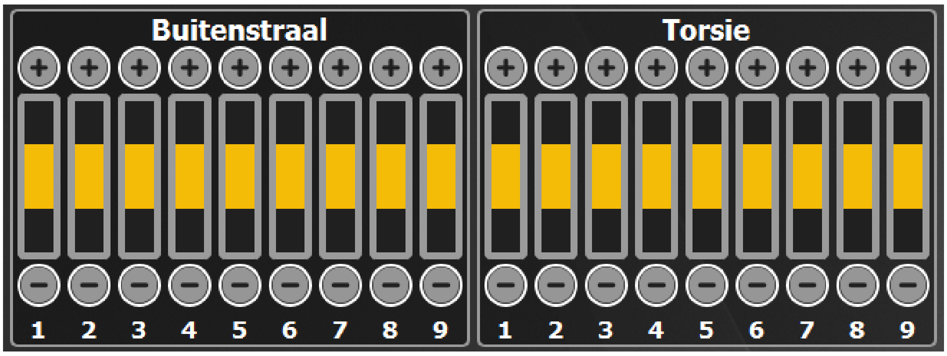


Figure 22: Subparts of the HMI: II

the key and also shows the range limit that the key has. The keys are automatically adjusted based on the data received but buttons for manual control are also included in case that would be necessary. In the simulation these manual controls are disabled, as automatic adjustment is always active. The different sliders act as buttons to open the corresponding graph of that particular leg, which will become visible in the main panel. When a slider is automatically adjusted it will light up green for a few seconds to indicate that an adjustment is made.

4.3.4 The main panel

The main panel of the HMI is where the graphs of the different legs are shown. This panel is flexible and can contain multiple graphs which can be moved and resized as desired. Figure 23 shows a few of these graphs opened. At the top of this panel a legend is visible together with a few buttons to change the timescale of the X-axis or to close all opened graphs. The legend with the buttons is only shown when a graph is opened. Otherwise it is hidden to keep the HMI clean and to prevent showing unnecessary information.

The graphs are based on the graphs shown in Section 3 in Figure 19. Each measurement is plotted as a dot with the time on the X-axis and the measurement value on the Y-axis. Instead of having the time displayed as a week and shift number it is shown as a time stamp, which changes depending on the view range. Control limits and specification limits are shown

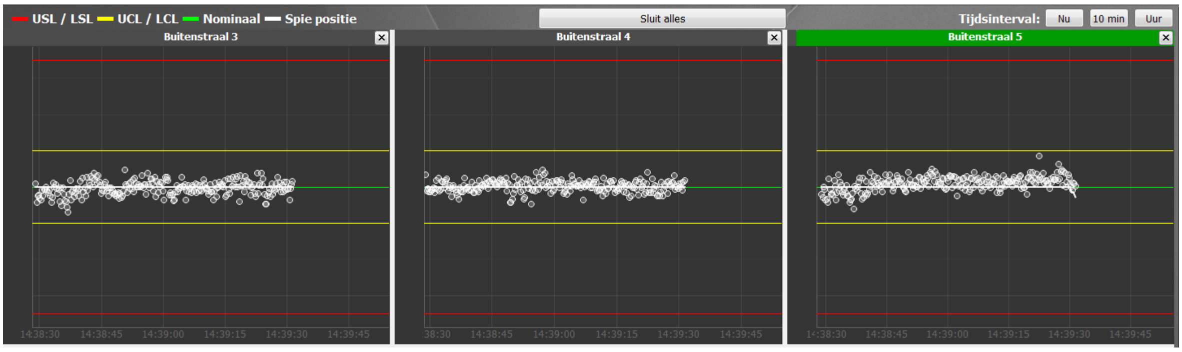


Figure 23: Graphs in the HMI

in the graph as well as the position of the key, which is relative to the Y-range and not the specific value of the Y-axis. The white line representing the key position will thus be in the middle (as seen in the figure) when the key is in the middle of its range. This way automatic adjustments are visualized in the graph, as a line that goes up and down depending on the key position. As an additional indicator that automatic adjustments are happening, the title bar of the graph turns green for a few seconds.

Each graph has a fixed Y-axis range. The range of the X-axis can be adjusted manually according to the preferences of the user, either by scrolling or dragging with the mouse. The amount of measurements visible will change according to the view range to prevent a cluttering of plotted points. When the view range exceeds 5 minutes visible data is downscaled to show only the means of 10 measurements, and further downscaling is done at view ranges of 60 minutes or a single day.

4.4 Relation to design guidelines

This section will relate the HMI design to the different design guidelines stated in Section 2.7.

Perception by expectation The HMI takes into account the expectations of the operators by using graphs which have all the same elements that the graphs they know also have. This makes the HMI consistent with other tools they use in their daily jobs.

Perception of structure There are no unwanted relationships in the HMI based on the gestalt principles, described in Section 2.7.2. Buttons and items belonging together are grouped or surrounded by a border and different panels of the HMI are divided by some white space grouping their functionality together.

Reading The amount of reading necessary for the HMI is kept to a minimum, one font is used (although with different sizes) and the contrast is maximized for the information that is important to the user. Also icons are used where possible to convey a meaning that could also be represented with text.

Color vision The amount of colors used in the HMI is kept to a minimum to keep the colors meaningful, not giving unnecessary colors to items which do not need it. None of the participants in the experiment reported difficulties distinguishing colors, however one participant mentioned that the use of color could be exploited more.

Peripheral vision To account for the limited vision in the peripheral field the HMI uses short flashing of for instance the keys to draw attention to a certain event. Also the flashing of keys is accompanied by the flashing of the corresponding title bar of the graph, so that the user will still notice the event if their attention is focused on the main panel.

Memory and attention The current HMI design leaves some room for improvements on this aspect, making the different CTQ's easier to distinguish. Currently both CTQ's have similar key representations and similar looking graphs, only differing in the text used. Some remarks from users indicated that they had difficulty telling which graph or key belonged to what CTQ value.

Recognition and recall For the settings and other menu buttons, icons were used to adhere to this guideline. The experiment does not include usage of these buttons so their effectiveness is not evaluated.

Learning Task-specific vocabulary from different systems already in use by operators are also used in the HMI. For example the names for the graph labels and the information displayed in the top panel. Also an image from a cutter as it is manufactured is used as a background image to align the HMI with the knowledge an operator already has of the process.

Time requirements The HMI takes into account the time requirements needed for a responsive system, by updating the graphs real-time so the system feels responsive. To keep the live updating smooth, some down-sampling of data is performed at bigger view ranges, otherwise the system can not keep up.

5 Methods

To test the effectiveness of the HMI designed (see Section 4) an experiment is designed with the objective to test this in a realistic setting. The setting in which the HMI will be implemented is that of a production line where the HMI will be located at the start of the process and will only be visible to operators when they actively walk over there to inspect some sort of problem. Another monitor, called the line monitor (see section 3.1.1), is already available to give information about the production line as a whole. This monitor is continuously monitored by the operators and will act as a trigger to obtain more information about a certain problem. One of the sources to gain more information in this case will be the HMI designed.

One of the challenges with this setup is that the operator only briefly interacts with the HMI. Whether or not this brief interaction is enough to gain a sufficient understanding of the process is the main question in the experiment described. In order to answer this question the experiment will test the understanding of the situation comparing continuous and occasional use of the HMI. This understanding is measured as SA with two different SA assessment methods, which will be described later in this section. If a difference is seen for SA between these uses, the situation at Philips is not optimal for maximizing SA of operators.

5.1 Participants

The ideal participants in this experiment would be the operators who will also have to work with the HMI if it is to be implemented. However there are not many operators working at a production line as only a few are needed at one time to monitor the process. To have more participants to do the analysis on the choice was made to include regular participants, consisting of colleagues working at the department of NPI, consisting of engineers responsible for the production lines.

In total a group of 11 operators and 19 engineers participated in the experiment. Of each group one participated in a pilot experiment. The group of operators consisted of experienced operators, with at least one year of experience in the field. Age was not recorded for either group but ranged between 25 and 65. The operators were all male and of the engineers there were 17 males and 2 females. All of the participants signed an informed consent before the start of the experiment.

5.2 Experiment design

A within-subject randomized block design was used for the current experiment. Because of the limited amount of participants, each participant did both conditions of the experiment. The main factor in the experiment is occasional versus continuous use; either the HMI was visible only during events or all of the time. A possible learning factor in this experiment was controlled for by alternating the condition which was done first by the participant, which resulted in two variants of the experiment. Another factor which might be of influence is the type of participants. Operators are more experienced at such a task and this might result in higher SA scores. To account for this, both experiment variants were balanced in the operator group and the regular participants.

5.3 The task

The task in the experiment is to operate a simulation of a production line in which the HMI is active. This simulation includes the automatic adjustment of parameters in the press, ensuring that the CTQ values measured will remain within the set boundaries. The simulation did not include the prediction of the CTQ values based on input measurements. Automatic adjustment could be the next step for Philips Drachten, so a simulation was designed to reflect a plausible setup that could be reality within a few years. A material thickness measurement is also simulated in the task. Automatic adjustments were only reactive and not proactive.

The main focus within the task is to monitor the key-values and look at their influence on the process. In reality a key represents a screw within the stamp which can be turned to influence how much a certain stamp process bends the material. Each key has a certain length and therefore a certain range in which it can adjust the process, which is a limitation which has to be kept in mind.

The experiment consisted of two simulations in which the data that was sent to the HMI was adjusted in such a way to give rise to certain situations on which the operator had to act. Five situations were created during a simulation, based on the possible behavior of the system when it would be implemented. For instance certain keys would auto-adjust to the maximum of their range, or the automatic adjustment of a key would fail. Around these five situations a SA questionnaire was built with three questions for each situation, one for each level of SA. Besides these five situations some other distracting events occurred to which the participant had to perform an action. For example one of the product handling machines had an error, which participants had to fix. These situations were added to keep participants actively participating in the situation.

In reality, a standard operator procedure (SOP) will be followed by the operator to solve a certain problem. Because the simulation had no physical machines on which some action could be performed the SOP's were replaced by short video fragments, completely unrelated to the problem. This was done to simulate the time away from the system that will otherwise occur when a SOP is completed. This also represents the distractions that occur in reality within the experiment. The goal of the situations thus was not to look at how participants would react to the situations, the actions that they had to take were prescribed and unrelated to the situation itself. The SAGAT questions asked during the task looked at the understanding of participants of the situations. Did they notice a key did not respond to automatic adjustment? Did they notice a drift in certain quality parameters?

The goal was to look at the information taken in by the participants about the situation, which was measured with two SA measures. No performance measurements were collected during the experiment as each participant had to complete the same actions, visible in the action list in Figure 21a, which stated what they had to do to solve the current problem.

5.4 Measures of Situation Awareness

Situation Awareness is used as a measure to determine the understanding of participants of the situation they are in. This measure is used as the dependent variable in the experiment to see which factors and manipulations have an effect on this measure. The SA measures used in this experiment are SART and SAGAT, as described in Section 2.6.

5.4.1 SART

In the current study a translated version of SART is used because most of the operators have an insufficient proficiency of the English language and this would probably lead to them misunderstanding the question if phrased in English. A validated version of the SART questionnaire in Dutch does not exist by the current knowledge of the author. The questionnaire is translated carefully and the version used in the experiment can be found in Appendix A.

5.4.2 SAGAT

The used SAGAT questions for each part of the experiment can be found in Appendix B. The questions differ slightly but are based on similar situations in both parts of the experiment. The questions are phrased in Dutch, for better understanding of the questions by the operator.

5.5 System Usability Scale

The original System Usability Scale (SUS) statements are translated to Dutch for use in this experiment. Because no validated version is currently known to the author, a new translation was made. The English statements however are concise and easy to translate so this should provide no problems. The SUS used in the experiment can be found in Appendix C, and is described in further detail in Section 2.8.

5.6 Procedure and Apparatus

The experiment setup can be seen in Figure 24. The experiment lasted approximately 60 minutes, depending on the speed with which the participant answered the questionnaires and did the actions. Each part of the experiment consisted of 5 sub-parts which all lasted approximately 5 minutes. Each sub-part was followed by three SAGAT questions, which were summed over a part of the experiment to get a SA score. After each part the SART questionnaire was administered, resulting in 4 measures of SA. At the end of the experiment the usability of the HMI was measured using the System Usability Scale (SUS) (Brooke, 1996).

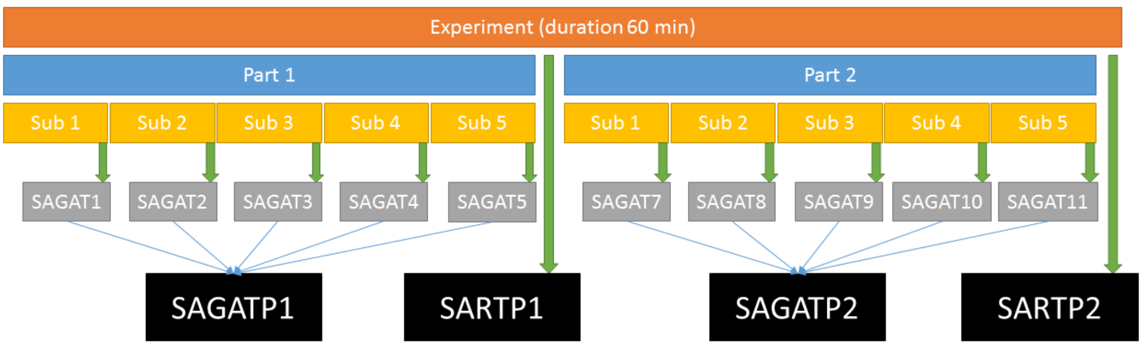


Figure 24: Overview of the experiment setup

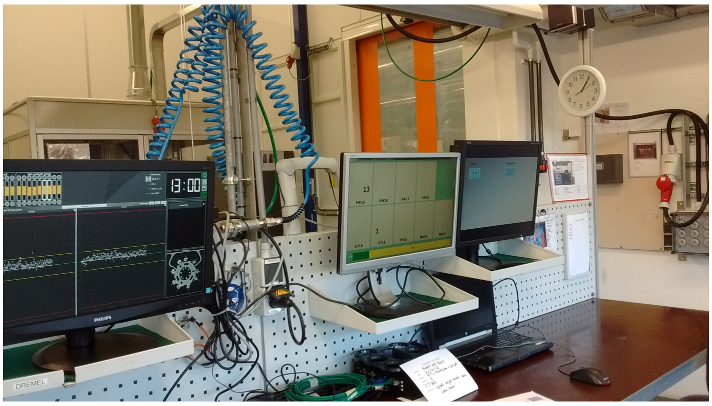


Figure 25: Photo of the experiment simulation setup

The experiment was setup in one of the testing facilities located at the factory, see Figure 25. A simulation of a production line was setup here with a line monitor, the HMI and a third screen to display the questionnaires and the distractions that occurred during the experiment. A participant had to monitor the process and perform the actions that popped up during the experiment. All participants were briefed for 10 minutes before the start of the experiment. They were given instructions about the experiment and also the HMI was introduced, showing the functionality of the HMI so everybody who started with the experiment knew what the different elements of the HMI meant. After this the experiment was started, which is described previously in Section 5.3. The SA questionnaires that were administered during the experiment, as well as the SUS, were recorded with a digital questionnaire tool.

During the experiment the participants were observed by the researcher and the researcher was available for any questions that the participant might have during the experiment.

5.7 Data collection and analysis approach

5.7.1 Data collection

After all participants had completed the experiment the data was collected and prepared for analysis. For analysis all of the SA scores from both the SART and SAGAT questionnaires are converted to a percentage for better comparison. The max of the summed SAGAT scores is 15 so here the percentage is calculated by: $(SAGAT/15) * 100$.

For SART this is somewhat more complicated, because SART is calculated from the different domains, see Equation1 in Section 2.6.1. Here U and D are calculated out of three dimensions and S from four dimensions. The maximum SART score is achieved by maximizing U and S and minimizing D. Each dimension is ranked on a scale from 1-7, resulting in the following calculation:

$$Max = 3 * 7 - (3 * 1 - 4 * 7) = 21 - (3 - 28) = 46 \quad (2)$$

The percentage score of SART is thus calculated as followed: $(SART/46) * 100$. The maximum score for SART will probably never be reached because people have a so called central tendency bias, avoiding extremes when filling out scales. The SART scores have to be evaluated in this light.

5.7.2 Data analysis approach

First the SA scores are analyzed for unwanted effects such as the training effect and subject type. Based on this outcome the data is either collapsed or analyzed separately.

SA scores were compared for different conditions of the experiment. Depending on the results of normality tests (Anderson-Darling and Shapiro-Wilk tests) either a paired t-test was used or a Wilcoxon signed rank test. Similar for non-paired data a t-test or a Wilcoxon rank sum test was used if necessary. For the analysis the means of different conditions are compared. This was done with hypothesis testing, where the null-hypothesis was always the assumption of no difference between the compared means.

SA scores were analyzed looking at the components that build up the score. For the SART this meant looking at the scores on different domains and for SAGAT of the different sub-parts and SA levels (1,2 and 3) of each sub-part.

Similarly SUS scores were compared between groups and collapsed over these groups when appropriate. Log-linear models were used to look at the relation between the SUS score and SA measures.

6 Results

6.1 Preliminary analysis

Before continuing with the primary analysis the data was checked for artifacts and unwanted effects from the experiment design. These results and any adjustments made are described next.

6.1.1 SART justification

With the translation of the dimension 'Concentration of attention' of the SART questionnaire there was an issue with it being ambiguous, meaning that not every participant interpreted this question the correct way. This dimension was opposite to the 'Division of attention' dimension within the same domain, and participants are thus expected to rate both dimension with opposite scores, scoring the dimensions with 1 and 7 or vice versa. However upon analyzing the answers given by the participants almost half of the participants did not have the expected opposite scores, while the other half clearly did have.

This lead to the conclusion that half of the participants misinterpreted the statement and the scores for those participants should be inverted for correct analysis of the results. So for further analysis the scores for those participants are inverted for that specific dimension, changing a score of 7 to 1, 6 to 2 and so on.

6.1.2 Missing values SAGAT

Because of some errors with the online questionnaires, some of the SAGAT sub-parts were not recorded for some of the participants. This led to missing values on sub-parts of the SAGAT for some participants, meaning that the maximum score for those participants could no longer be 15. Out of the 27 participants, 7 had some missing values, so instead of removing all of their SAGAT scores, their SAGAT percentage score is calculated by: $(SAGAT/12) * 100$. Only participants that were missing one of the sub-parts of SAGAT were treated this way, if multiple parts were missing the resulting SAGAT value was excluded from the analysis. Only one SAGAT score was excluded from the analysis for one part of the experiment.

6.1.3 Training Effect

Each participant did both conditions of the experiment in random order. Because both conditions involved similar situations a training effect might be present in the data, with participants performing better on part 2 compared to part 2. SA scores for part 1 and part 2 of the experiment can be seen in Figure 26a. A Welch two sample t-test showed no difference between the first and second part of the experiment for the SART, $t(26) = 0.06$, $p = 0.95$. For the SAGAT the difference between the first and second part was also insignificant, $t(25) = 0.85$, $p = 0.40$. No training effect is present in the data, so the data can be collapsed over the experiment order for further analysis.

6.1.4 Operators vs. engineers

The experiment was conducted with two subject groups, a group of operators and a group of engineers. Because of experience with monitoring a production line operators might have higher SA compared to engineers. Figure 26b shows the SA scores for both experiment

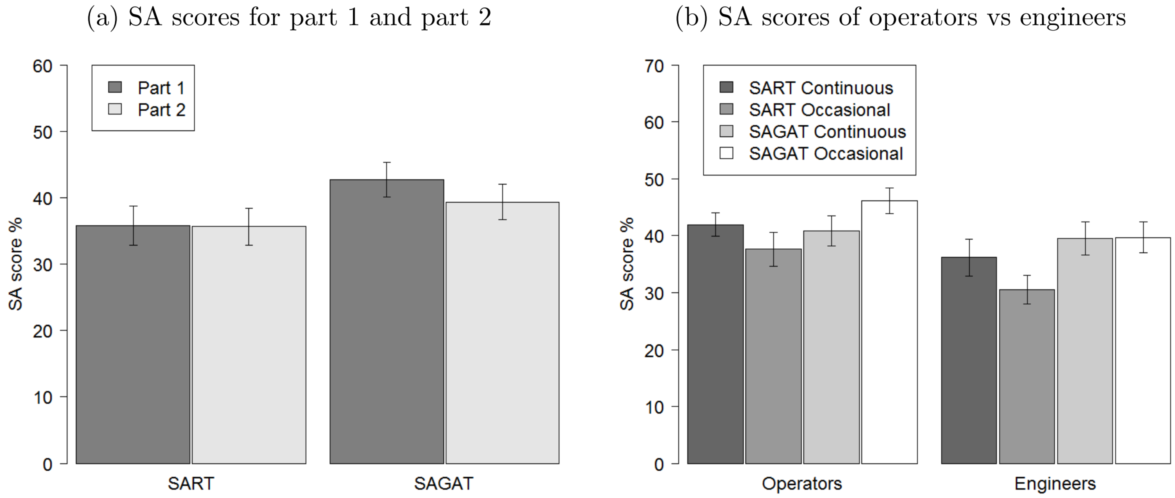


Figure 26: SA scores evaluated for training effects and subject type. Error bars indicate the standard error of the mean.

conditions split by subject type. Welch two sample t-tests show no significant difference between means of operators and engineers on either of the four categories; SART continuous ($t(24.7) = 1.08$, $p = 0.29$), SAGAT continuous ($t(20.57) = 0.23$, $p = 0.82$) and SAGAT Hidden ($t(21.54) = 1.28$, $p = 0.21$).

Normality tests for the SART occasional data indicated deviations from normality. A Wilcoxon rank sum test with the same data is also non-significant, W=116 (W refers to the test statistic), $p = 0.12$. No difference between operators and engineers is thus found, so the data of SA measures is collapsed over subject type for further analysis.

6.2 Primary analysis

The previous section showed no training effect and difference between subject type, so the data is collapsed and compared between continuous and occasional use.

6.2.1 Continuous vs. occasional use

To see if occasional use results in similar SA compared to continuous use we compare the SA scores for both conditions. Figure 27 shows the mean SA score of both the SART and SAGAT for each of the experiment conditions. Exact mean values are available for look-up in Appendix D in Table 6. A paired t-test shows no difference between continuous and occasional HMI use for the SART measure, $t(26\ (df)) = 1.99$, $p = 0.057$. Similarly for the SAGAT, a paired t-test shows no difference between continuous and occasional HMI use, $t(25) = -0.44$, $p = 0.66$. No difference between conditions is found for the total SA scores.

6.2.2 Measures of Situation Awareness

The SA scores from the SA measures used are composed of different components. The SART consists of demand (D), supply (S) and understanding (U), representing the different domains of SA. SAGAT is calculated from the scores on different levels of SA, level 1 (L1), level 2 (L2)

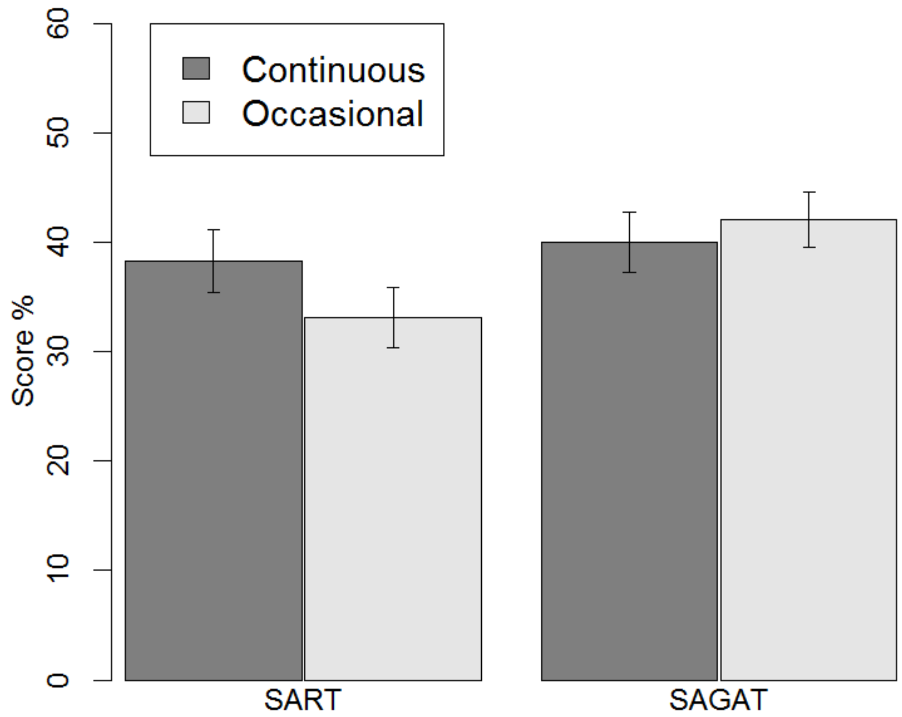


Figure 27: Mean SA values for each experiment condition. Error bars show the standard error of the mean.

and level 3 (L3). The scores of participants on these different components are seen in Figure

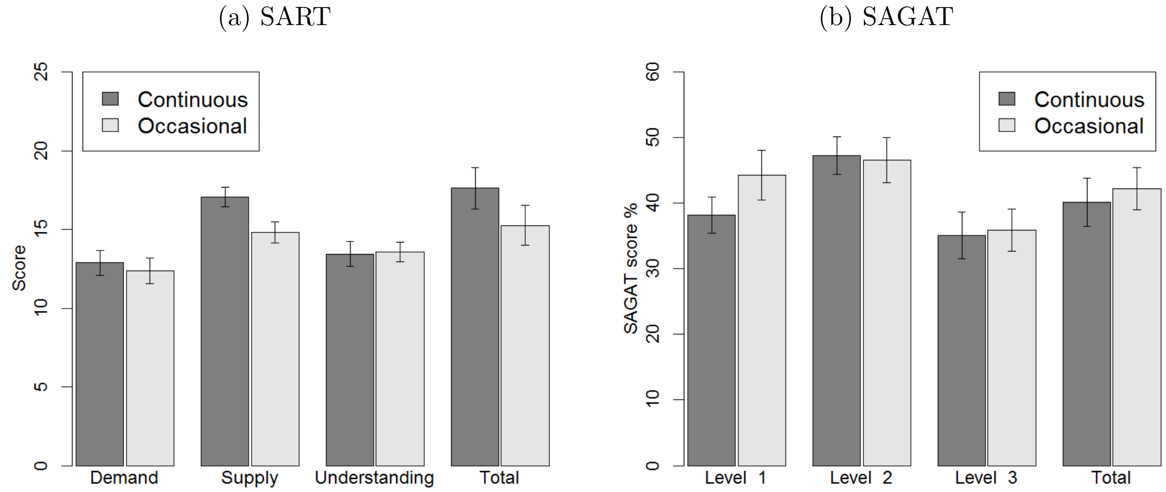


Figure 28: SART and SAGAT mean scores split up in domains and levels of SA. Error bars indicate the standard error of the mean. For the SART the shown values are absolute, as a formula is used to calculate the total score from the domain scores.

28a and 28b. Effects of continuous versus occasional use might be visible when looking at these components of each SA measure.

Figure 28a shows the different domains of the SART. Only the Supply component of the SART questionnaire shows an effect of experiment condition, with a higher mean for the continuous use of the HMI. Because of non-normality and paired data a Wilcoxon signed rank test is used. The test is insignificant with W = 231.5 and p = 0.02, thus rejecting the null-hypothesis of the means being equal. The other domains of SART and the levels of SAGAT show no significant difference between means between both conditions.

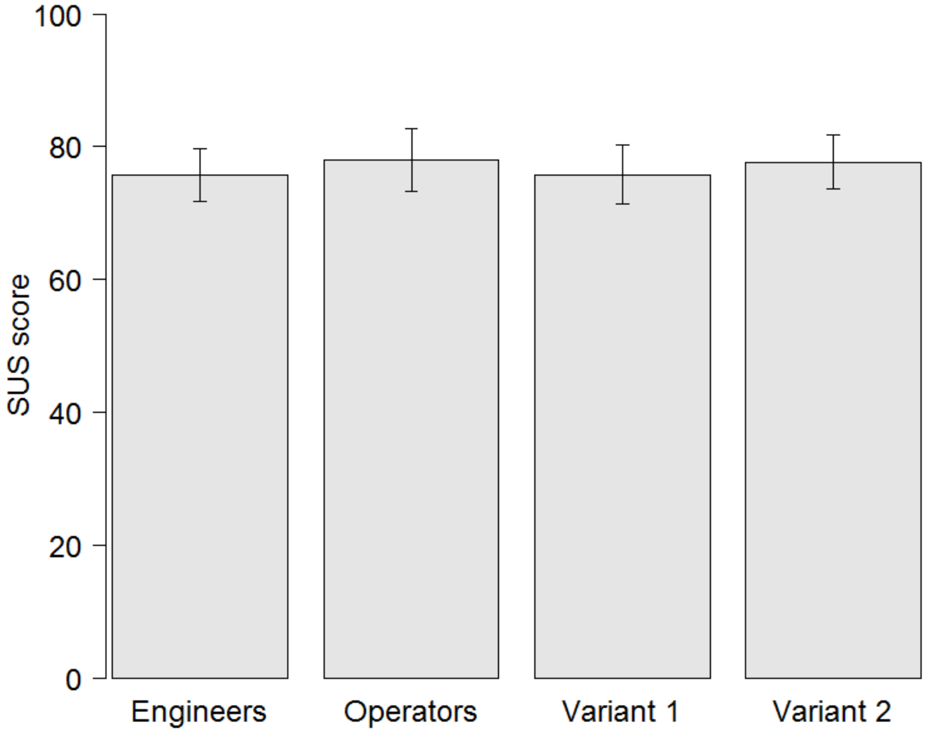


Figure 29: Mean SUS scores per subject type and experiment variant. Error bars show the standard error of the mean.

6.2.3 System Usability Score

The mean System Usability Score (SUS) of all participants is 77, corresponding to an adjective rating of 'good'. Subject type and experiment variant (which condition was done first) might influence this score. Figure 29 shows the mean SUS scores for each of these groups. A Welch two sample t-test shows no difference between operators and engineers, $t(18.21) = -0.34$, $p = 0.74$.

Normality tests for the variant 1 data show deviations from normality. A Wilcoxon rank sum test shows no difference between variants, $W = 70$, $p = 0.95$. SUS scores are thus not influenced by the first encounter participants had with the HMI.

6.2.4 Situation Awareness and System Usability

For further analysis the relation between SA measures and SUS was analyzed. Hypothesis here is that with a higher system usability information is easier to access for the user. With most systems the HMI is the primary source of information to construct a mental model of the situation. Expectation is thus that a higher SUS score will result in higher SA, as the user can extract information from the HMI more easily.

However this relation is not expected to be linear, as a system that has a usability score lower than 50 is considered not acceptable. With SUS scores below this threshold a user has problems using the HMI and thus extracting information from it. Expected is that SA will increase after a certain SUS has been reached, resulting in a non-linear relationship. For the model to be fitted to the data a log-linear model is chosen, as this fits the relation described. This model uses the log values of the SA measures.

Log-linear models for both SART and SAGAT were constructed for each experiment conditions, four models in total. No violations of the assumptions for linear models were found. Three subjects with a missing SUS score were excluded from the correlation analysis.

SART vs SUS A linear model of the SART score as a function of SUS was created for each condition, with the experiment variant and subject type as blocking factors. For continuous use of the HMI this model was significant (F(3,20) = 6.37, p = 0.003, R^2 = 0.49) with SUS being the only significant coefficient, p = 0.001. For occasional HMI use this model was not significant (F(3,20) = 2.05, p = 0.14, R^2 = 0.24). The model fit for the continuous use of the HMI can be seen in Figure 30a.

(a) SART score vs SUS score

(b) SAGAT score vs SUS score

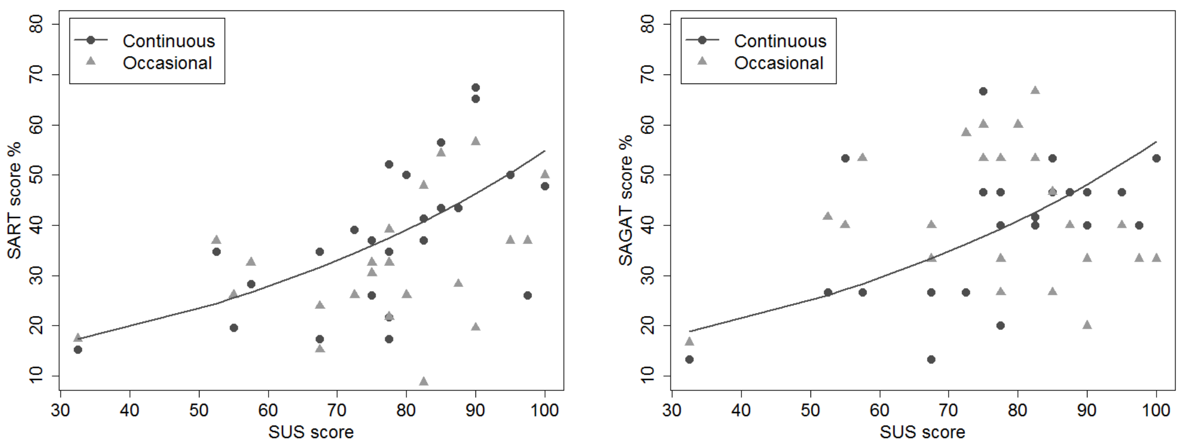


Figure 30: SA score plotted against SUS score, with one point for each subject on both conditions.

SAGAT vs SUS A linear model of the SAGAT score as a function of SUS was created for each condition, with the experiment variant and subject type as blocking factors. For continuous use of the HMI this model was significant (F(3,19) = 4.45, p = 0.016, R^2 = 0.41) with SUS being the only significant coefficient, p = 0.003. For occasional HMI use this model was not significant ($F(3,20) = 1.22$, $p = 0.33$, $R^2 = 0.15$). The model fit for the continuous use of the HMI can be seen in Figure 30b.

7 Discussion

The current study tried to find the influence of HMI on SA, specific to the use case at Philips. The needs of this HMI and how to visualize those needs efficiently are prerequisites to answer the question of continuous versus occasional use of that HMI. Future monitoring of production lines might include having one operator for multiple production lines, requiring the operator to maintain SA of both. With this scenario in mind the operator has to resort to occasional monitoring of each production line. The experiment conducted thus looks at this occasional monitoring, can occasional monitoring result in similar SA compared to continuous monitoring?

The main results of the experiment show no difference in SA between occasional and continuous use of the same HMI. This indicates that continuous monitoring is not necessary to keep good SA of a production line. The current study does not allow for inferences on the specifications of the HMI that are needed for this level of SA. The mean SUS score of 77 does indicate that the HMI is judged as a good interface. The needs of operators in the situation of Philips are thus visualized efficiently, by keeping the information presented to a minimum and using representations that resemble actual functionality.

Other situations with a different interface might lead to different results, and not every situation requires occasional use of an interface. Most cases of interface research focus on for instance plant control (Tharanathan et al., 2010; Bowden & Rusnock, 2015), in which a HMI is the only source of information. The current situation at Philips does not allow for such an approach, here a HMI is only one of the multiple sources of information an operator has at his disposal. The results can thus not directly be applied to other situations not resembling the setup at Philips.

The following subsections will further discuss the results of the previous sections and other limitations of this study.

7.1 SA measures

The scores for both SART and SAGAT have similar averages, around 40%. These values are hard to compare however, as they use different methods of acquiring these values. A SAGAT score depends on the difficulty of the questions, which have to be constructed for each situation. SART uses a generic questionnaire but a perfect SA of 100% will probably never be reached as this would require extreme responses on the answer scales, which tend to be avoided. Naderpour et al. (2016) reported an absolute SART score between 20-30, compared to which the current scores are quite low. This might be an indication of the difficulty of the task. Both SART and SAGAT results were analyzed for abnormalities

7.1.1 SAGAT

When looking at the SAGAT scores these don't show the expected pattern that scores from level one to level three are descending. This can be further explored by looking at the different sub-parts of the SAGAT questionnaire, seen in Figure 31. This figure shows that SAGAT 5 and 8 show a reversed pattern, with more correct answers on level 3 questions. As level 3 SA is supposed to be achieved by combining level 1 and 2 SA this result is unusual. A possible explanation for this is that the questions chosen for each of the SA levels were not appropriate. This is one of the drawbacks of using the SAGAT method, as the used questions differ for each situation and ideally require a full experimental study to fine-tune.

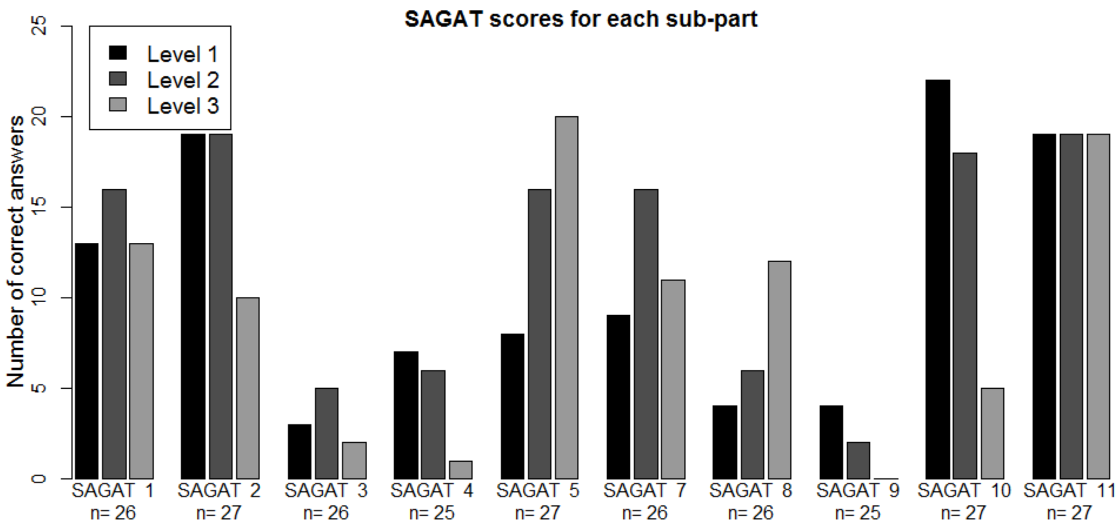


Figure 31: Scores on each SAGAT sub-part, for all participants.

Another abnormality in Figure 31 is the low overall score on SAGAT parts 3, 4 and 9. The corresponding SAGAT questions covered events that did not occur directly preceding the questions. For example part 3 and 9 referred to a situation before a change of input material in the experiment. This change of input material caused some automatic adjustments but the situation of interest was a quality parameter going out of bounds before the material change. The questions were explicit about which situation they referred to but apparently some participants did not read the questions thoroughly. Another explanation would be that they simply did not retain the information long enough to answer the questions correctly.

7.1.2 SART

When looking at the SART scores we see that the difference in the total score can be attributed to the difference in supply between both conditions. The difference between both conditions is also found significant. Being able to monitor the HMI longer thus leads to a higher supply of information, which is not surprising. This difference in supply however is not reflected in the resulting understanding of the situation. A more in-depth view of the SART score build up can be seen in Figures 32a and 32b.

This figure shows the SART scores for each condition of the experiment for all participants, ordered from the lowest SA score to the highest. It shows that higher SA scores are achieved by a lower score on the 'Demand' domain, which is subtracted from 'Supply'. The difference in 'Understanding' for participants with a high SA and a middle SA score is very small. These patterns can be seen for both conditions of the experiment. A higher 'Supply' does not seem to have a direct effect on the total SA score, which Figure 28a does suggest.

7.2 Visibility HMI

The results show no effects of HMI visibility on the SA of participants, but for the SART measure it approaches significance. The sample size (n=27) however exhausted the main

(a) HMI visible (b) HMI hidden

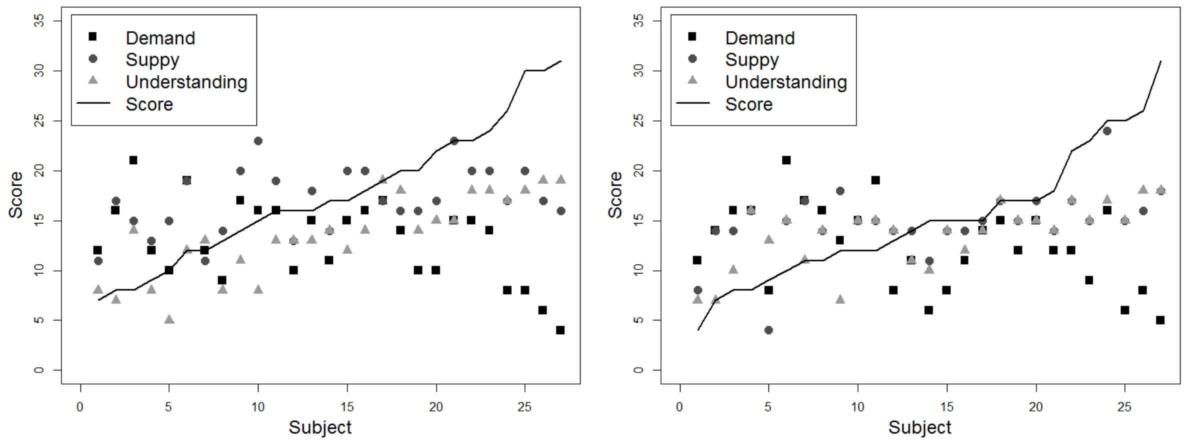


Figure 32: Scores of each subject on both SART questionnaires.

population of interest, so this cannot be attributed to a small sample size. The increase in the SART total score is mainly contributed to an increase in the supply-domain of the SART. Continuous or occasional monitoring had no effect on the understanding domain of the SART. The current study does not look at task performance, so this can't be analyzed to see if the difference in supply had any effect here. Previous research also only relates SA as a total measure to performance, see Section 2.4.1.

The SAGAT measure showed no significant effect of the experimental manipulation, showing no effect of continuous versus occasional HMI use. SAGAT has been found to be sensitive to display manipulation before (Tharanathan et al., 2010), specifically to level 1 and level 2 SA. The current study also shows no difference between individual levels of the SAGAT, see Figure 28b. The results in this study concerning SAGAT have to be interpreted with care because of the unexpected results of that questionnaire, discussed earlier.

Overall the results have implications for further automation of production lines. It shows that SA does not suffer from occasional monitoring of a HMI, compared to continuous monitoring. If this result can be validated this could generalize to other domains and might enable operators to monitor multiple processes, while maintaining enough SA to act upon failure. An interesting study would be to see if one operator monitoring two processes simultaneously can achieve similar SA to one operator monitoring only one of these processes.

7.3 System usability

To measure the appreciation of the HMI by participants the System Usability Scale (SUS) was used. The mean SUS score is 77 which reflects to an adjective rating of 'Good', see Figure 16. Thus meaning the used HMI is perceived as a good and usable system. The SUS score does also not vary between subject type and the order in which the experiment was conducted.

Additional analysis was done to relate the SUS score with SA, as a relation might be present there. A system with a good usability should also have provided the operator with the necessary information. A lower rating of SUS would relate to the ease of use and inconsistency, which could also affect the amount of information an operator obtains from the HMI. To test if such a relation between SUS and SA consisted the SA and SUS scores were plotted for each participant. This could not be related to the visibility of the HMI as only one interface was used for both conditions and only one SUS was administered.

When analyzing these plots visually, an indication of a relation can be spotted, warranting the use of linear models. For the SART measures a model was found significant, indicating a relation between SUS and SART. For the SAGAT measures this relation was not found significant, probably because of the previously discussed problems with the SAGAT measure. Another could be that the SUS measure is only obtained once, at the end of the experiment. This SUS is thus based on the use of both occasional and continuous use of the HMI. Although the HMI used for both types of use is the same, some influence of the amount of use might effect SUS, which is currently unknown.

From the plots it is not clear if a non-linear (log) or a linear approach is better suited for this relation. A linear approach is most straightforward, but the log models for the SART and SAGAT in the HMI visible conditions are a better fit to the data. A linear model would suggest that SUS scales with SA linearly. It can also be argued that up to a certain point, bad interfaces (SUS < 50), are not easy to use and thus provide not enough information to the user. Only after an interface achieves a minimum level of SUS it will then result in increasing SA levels.

7.4 Further research

One of the main topics of further research could be the validation of occasional monitoring versus continuous monitoring. This should be validated within multiple domains and with different SA measures. If similar SA can be achieved with both methods, this could be implemented into multiple process monitoring by one operator. As mentioned earlier, other effects of monitoring two processes instead of one should also be researched.

The relation between SUS and SA also needs further research, validating it's existence across different domains and with different SA measures. Potential uses of this relation could be predicting SA from SUS, as SUS is easier to administer compared to using a SART or SAGAT measure. Optimizing the SUS could then be used to also optimize the resulting SA from the HMI design.

8 Conclusion

This thesis looked at the effect of a HMI on SA of operators, with the following research question:

In what way does a Human Machine Interface add to the Situation Awareness of operators?

Specific to the situation at Philips three sub-questions are derived from this question, which were answered in this thesis. Those questions were:

- What are the needs of the operator for such a HMI?
- How can these needs be met and the data visualized efficiently?
- Is there a difference between continuous and occasional monitoring of a HMI?

The first two questions are specific to the HMI to be designed for the case at Philips, and are answered in Sections 3 and 4 respectively. This was done by interviewing the operators and reviewing the design guidelines from the literature.

The third question required an experiment that is described in Section 5. Results of this experiment justify the use of occasional monitoring at production lines. The System Usability Scale (SUS) has been found to correlate with SA, so in order to maximize the SA of operators the interfaces they use should be evaluated for usability using SUS. Effective Human Machine Interface (HMI) design is necessary to enable sufficient Situation Awareness (SA) for operators.

Aiming for system usability may be one of the most important aspects of using a HMI within production lines. This aspect is currently not emphasized, as interfaces that are available to the operator are usually designed from a functional aspect, not with the user in mind.

8.1 Use for Philips

The results from the experiment align with the current practice at Philips where a lot of information is available to the operator, but only when actively seeking this information. This type of monitoring can result in similar SA compared to continuous monitoring, justifying the current practice.

The current direction of production industry will eventually result in highly automated processes were an operator is required to monitor multiple processes from some sort of control room. This setup can have consequences for operator SA and has to be investigated carefully. Effective interface design that is user-centered, optimized for SA, will become more and more important.

References

Adams, M. J., Tenney, Y. J., & Pew, R. W. (1995). Situation awareness and the cognitive management of complex systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 85–104.

Amalberti, R., & Deblon, F. (1992). Cognitive modelling of fighter aircraft process control: a step towards an intelligent on-board assistance system. *International Journal of Man-Machine Studies*, *36*(5), 639–671.

Bainbridge, L. (1983). Ironies of automation. *Automatica*, *19*(6), 775–779.

Bangor, A., Kortum, P., & Miller, J. (2009). Determining what individual sus scores mean: Adding an adjective rating scale. *Journal of usability studies*, *4*(3), 114–123.

Baxter, G., Rooksby, J., Wang, Y., & Khajeh-Hosseini, A. (2012). The ironies of automation: still going strong at 30? In *Proceedings of the 30th european conference on cognitive ergonomics* (pp. 65–71).

Billings, C. E. (1991). Human-centered aircraft automation: A concept and guidelines. *NASA Technical Memorandum 103885*.

Billings, C. E. (1997). *Aviation automation: The search for a human-centered approach.*

Bowden, J. R., & Rusnock, C. F. (2015). Evaluation of human machine interface design factors on situation awareness and task performance. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 59, pp. 1361–1365).

Brook, Q. (2014). *Lean six sigma and minitab.* Opex Resources.

Brooke, J. (1996). Sus-a quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4–7.

Brooke, J. (2013). Sus: a retrospective. *Journal of usability studies*, *8*(2), 29–40.

Burns, C. M., Skraaning, G., Jamieson, G. A., Lau, N., Kwok, J., Welch, R., & Andresen, G. (2008). Evaluation of ecological interface design for nuclear process control: situation awareness effects. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(4), 663–679.

Carmody, M. A., & Gluckman, J. P. (1993). Task specific effects of automation and automation failure on performance, workload and situational awareness. In *Proceedings of the seventh international symposium on aviation psychology* (Vol. 1, pp. 167–171).

Cheriton, D. R. (1976). Man-machine interface design for timesharing systems. In *Proceedings of the 1976 annual conference* (pp. 362–366).

Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and brain sciences*, *24*(01), 154–176.

Endsley, M. R. (1988). Situation awareness global assessment technique (sagat). In *Aerospace and electronics conference, 1988. naecon 1988., proceedings of the ieee 1988 national* (pp. 789–795).

Endsley, M. R. (1995a). A taxonomy of situation awareness errors. *Human factors in aviation operations*, *3*(2), 287–292.

Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 32–64.

Endsley, M. R. (1996). Automation and situation awareness. *Automation and human performance: Theory and applications*, 163–181.

Endsley, M. R. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, *42*(3), 462–492.

Endsley, M. R. (2000). Direct measurement of situation awareness: Validity and use of sagat. *Situation awareness analysis and measurement*, *10*.

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(2), 381–394.

Endsley, M. R., Selcon, S. J., Hardiman, T. D., & Croft, D. G. (1998). A comparative analysis of sagat and sart for evaluations of situation awareness. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 42, pp. 82–86).

Endsley, M. R., Sollenberger, R., Nakata, A., & Stein, E. S. (2000). *Situation awareness in air traffic control: Enhanced displays for advanced operations* (Tech. Rep.). Washington, D.C: DTIC Document.

Funk, K., Lyall, B., Wilson, J., Vint, R., Niemczyk, M., Suroteguh, C., & Owen, G. (1999). Flight deck automation issues. *The International Journal of Aviation Psychology*, *9*(2), 109–123.

Janelle, C. M., Singer, R. N., & Williams, A. M. (1999). External distraction and attentional narrowing: Visual search evidence. *Journal of Sport & Exercise Psychology*.

Johnson, J. (2013). *Designing with the mind in mind: simple guide to understanding user interface design guidelines.* Elsevier.

Jones, D. G. (2000). Subjective measures of situation awareness. *Situation awareness analysis and measurement*, 113–128.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual review of psychology*, *59*, 193.

Kaber, D. B., Perry, C. M., Segall, N., McClernon, C. K., & Prinzel, L. J. (2006). Situation awareness implications of adaptive automation for information processing in an air traffic control-related task. *International Journal of Industrial Ergonomics*, *36*(5), 447–462.

Kim, S. K., Suh, S. M., Jang, G. S., Hong, S. K., & Park, J. C. (2012). Empirical research on an ecological interface design for improving situation awareness of operators in an advanced control room. *Nuclear Engineering and Design*, *253*, 226–237.

Lee, J. D. (2006). Human factors and ergonomics in automation design. *Handbook of Human Factors and Ergonomics, Third Edition*, 1570–1596.

Lee, J. D. (2008). Review of a pivotal human factors article:humans and automation: use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(3), 404–410.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *46*(1), 50–80.

Lindsay, P. H., & Norman, D. A. (2013). *Human information processing: An introduction to psychology.* Academic Press.

Liu, Q., Nakata, K., & Furuta, K. (2004). Making control systems visible. *Cognition, Technology & Work*, *6*(2), 87–106.

Lorenz, B., Di Nocera, F., Röttger, S., & Parasuraman, R. (2001). The effects of level of automation on the out-of-the-loop unfamiliarity in a complex dynamic fault-management task during simulated spaceflight operations. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 45, pp. 44–48).

Mackworth, N. H. (1950). Researches on the measurement of human performance. *Med. Res Council Spec Rep Ser*(268).

Martin, M., & Jones, G. (1984). Cognitive failures in everyday life. *Everyday memory, actions and absent-mindedness*, 173–190.

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.

Mosier, K., & Chidester, T. (1991). Situation assessment and situation awareness in a team setting. *Designing for everyone*, 798–800.

Naderpour, M., Lu, J., & Zhang, G. (2016). A safety-critical decision support system evaluation using situation awareness and workload measures. *Reliability Engineering & System Safety*, *150*, 147–159.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, *2*(2), 175.

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 249–256).

OHara, J., & Higgins, J. (2010). Human-system interfaces to automatic systems: Review guidance and technical basis. *Human Factors of advanced reactors (NRC JCN Y-6529) BNL Tech Report No BNL91017-2010*, 1–23.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(2), 230–253.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, *30*(3), 286–297.

Pfeiffer, S. (2016). Robots, industry 4.0 and humans, or why assembly work is more than routine work. *Societies*, *6*(2), 16.

Raisinghani, M. S., Ette, H., Pierce, R., Cannon, G., & Daripaly, P. (2005). Six sigma: concepts, tools, and applications. *Industrial management & Data systems*, *105*(4), 491–505.

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. *IEEE transactions on systems, man, and cybernetics*(3), 257–266.

Rasmussen, J. (1986). *Information processing and human-machine interaction. an approach to cognitive engineering.* New York: North-Holland.

Regal, D. M., Rogers, W. H., & Boucek, G. P. (1988). Situational awareness in the commercial flight deck: Definition, measurement, and enhancement. *SAE Technical Paper*.

Riley, V. (1994). A theory of operator reliance on automation. *Human performance in automated systems: Current research and trends*, 8–14.

Ross, J. M., Szalma, J. L., Hancock, P. A., Barnett, J. S., & Taylor, G. (2008). The effect of automation reliability on user automation trust and reliance in a search-and-rescue scenario. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 52, pp. 1340–1344).

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., & Young, M. (2009). Measuring situation awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics*, *39*(3), 490–500.

Satuf, E. N., Kaszkurewicz, E., Schirru, R., & de Campos, M. C. M. M. (2016). Situation awareness measurement of an ecological interface designed to operator support during alarm floods. *International Journal of Industrial Ergonomics*, *53*, 179–192.

Sheridan, T. B. (1992). *Telerobotics, automation, and human supervisory control.* MIT press.

Sheridan, T. B. (2002). *Humans and automation: System design and research issues.* John Wiley & Sons, Inc.

Shneiderman, B. (2010). *Designing the user interface: strategies for effective human-computer interaction.* Pearson Education India.

Sousa, D. A. (2014). *How the brain learns to read.* Corwin Press.

Taylor, R. (1990). Situational awareness rating technique(sart): The development of a tool for aircrew systems design. *AGARD, Situational Awareness in Aerospace Operations 17 p(SEE N 90-28972 23-53).*

Tharanathan, A., Laberge, J., Bullemer, P., Reising, D. V., & McLain, R. (2010). Functional versus schematic overview displays: Impact on operator situation awareness in process monitoring. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 54, pp. 319–323).

Tullis, T. S., & Stetson, J. N. (2004). A comparison of questionnaires for assessing website usability. In *Usability professional association conference* (pp. 1–12).

Vicente, K. J., & Rasmussen, J. (1992). Ecological interface design: Theoretical foundations. *Systems, Man and Cybernetics, IEEE Transactions on*, *22*(4), 589–606.

Wiener, E. L. (1989). *Human factors of advanced technology (glass cockpit) transport aircraft.* Moffett Field, Calif: NASA Ames Research Center.

Willems, B., & Heiney, M. (2002). *Decision support automation research in the en route air traffic control environment.* Department of Transportation, Federal Aviation Administration.

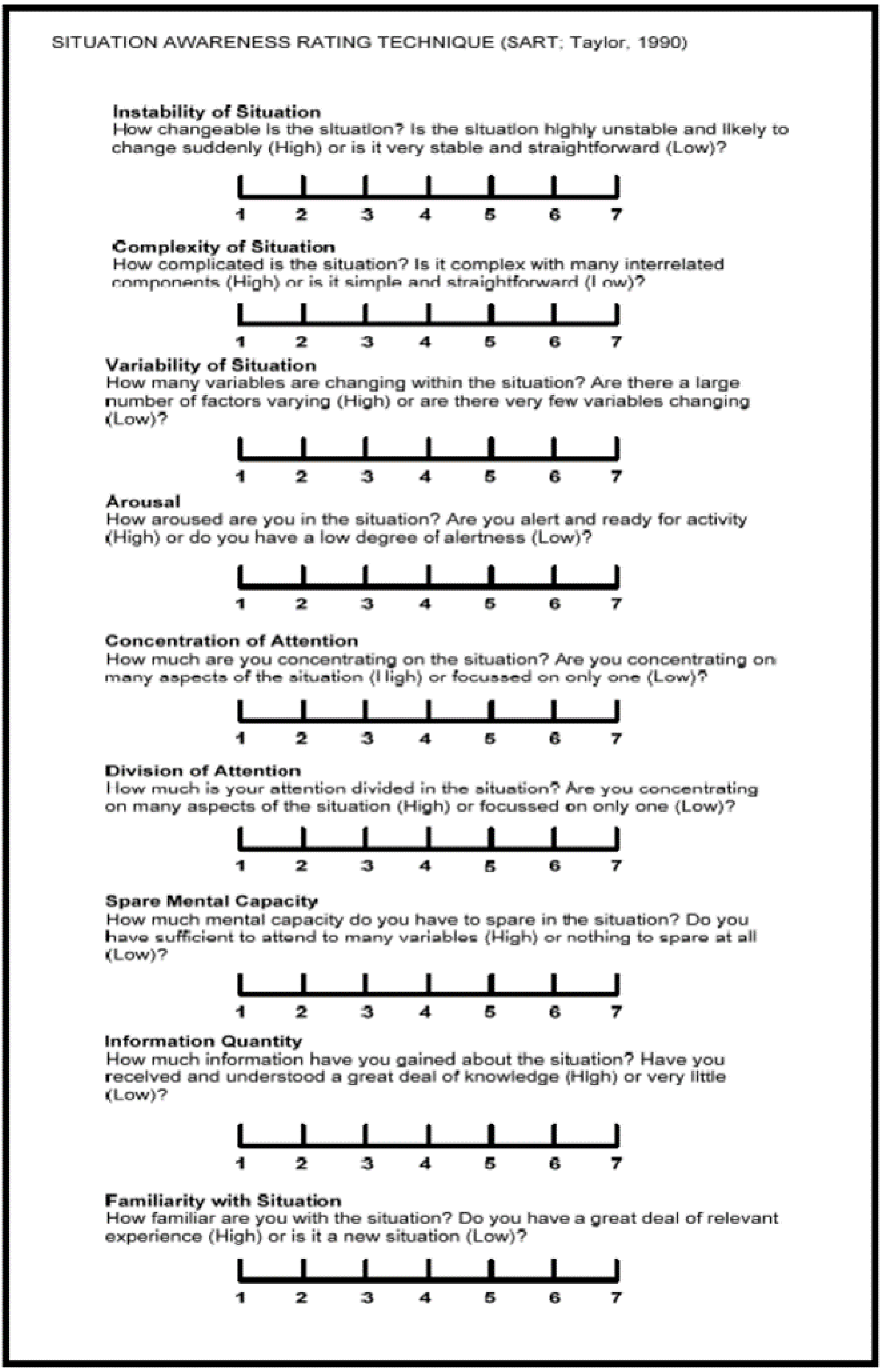
Wulf, F., Rimini-Döring, M., Arnon, M., & Gauterin, F. (2015). Recommendations supporting situation awareness in partially automated driver assistance systems. *IEEE Transactions on Intelligent Transportation Systems*, *16*(4), 2290–2296.

Appendices

A SART questionnaire

This is the questionnaire used to get a SA value using the SART technique. The questionnaire is translated from English to Dutch staying as close as possible to the original phrasing. The original questionnaire can be seen in Figure 33. Figure 34 shows the questionnaire as implemented in a digital form, used in the experiment.

- **Instabiliteit van de situatie**
  Is de situatie instabiel en is het waarschijnlijk dat het zomaar verandert (hoog) of is het het tegenovergestelde (laag)?
- **Complexiteit van de situatie**
  Hoe complex is de situatie? Is het complex met veel van elkaar afhankelijke variabelen (hoog) of is het simpel en direct duidelijk (laag)?
- **Variabiliteit van de situatie**
  Hoeveel variabelen veranderen in de situatie? Zijn er veel factoren die veranderen (hoog) of zijn dit er maar weinig (laag)?
- **Alertheid**
  Hoe alert ben je in de situatie? Ben je alert en klaar om in actie te komen (hoog) of heb je moeite om alert te blijven (laag)?
- **Concentratie van aandacht**
  In welke mate ben je geconcentreerd op de situatie? Concentreer je op vele aspecten van de situatie (laag) of op maar een enkel aspect (hoog)?
- **Verdeeldheid van aandacht**
  In welke mate is je aandacht verdeeld in de situatie? Concentreer je op veel aspecten (hoog) of maar op een enkel aspect (laag)?
- **Mentale belasting**
  Wat is je mentale belasting in de situatie? Is deze erg hoog en kon je er niks meer naast doen (hoog) of is deze laag en kun je nog meer dingen in de gaten houden (laag)?
- **Kennis van de situatie**
  Hoeveel informatie heb je verkregen over de situatie? Heb je veel informatie opgedaan en begrepen (hoog) of maar weinig (laag)?
- **Kwaliteit van de informatie**
  Hoe goed is de informatie die je hebt verkregen over de situatie? Is de kennis die verkregen is erg nuttig (hoog) of is het een nieuwe situatie (laag)?
- **Bekendheid met de situatie**
  Hoe bekend ben je met de situatie? Heb je veel relevante ervaring (hoog) of is de situatie helemaal nieuw (laag)?



SITUATION AWARENESS RATING TECHNIQUE (SART; Taylor, 1990)

Instability of Situation
How changeable is the situation? Is the situation highly unstable and likely to change suddenly (High) or is it very stable and straightforward (Low)?

1 2 3 4 5 6 7

Complexity of Situation
How complicated is the situation? Is it complex with many interrelated components (High) or is it simple and straightforward (Low)?

1 2 3 4 5 6 7

Variability of Situation
How many variables are changing within the situation? Are there a large number of factors varying (High) or are there very few variables changing (Low)?

1 2 3 4 5 6 7

Arousal
How aroused are you in the situation? Are you alert and ready for activity (High) or do you have a low degree of alertness (Low)?

1 2 3 4 5 6 7

Concentration of Attention
How much are you concentrating on the situation? Are you concentrating on many aspects of the situation (High) or focussed on only one (Low)?

1 2 3 4 5 6 7

Division of Attention
How much is your attention divided in the situation? Are you concentrating on many aspects of the situation (High) or focussed on only one (Low)?

1 2 3 4 5 6 7

Spare Mental Capacity
How much mental capacity do you have to spare in the situation? Do you have sufficient to attend to many variables (High) or nothing to spare at all (Low)?

1 2 3 4 5 6 7

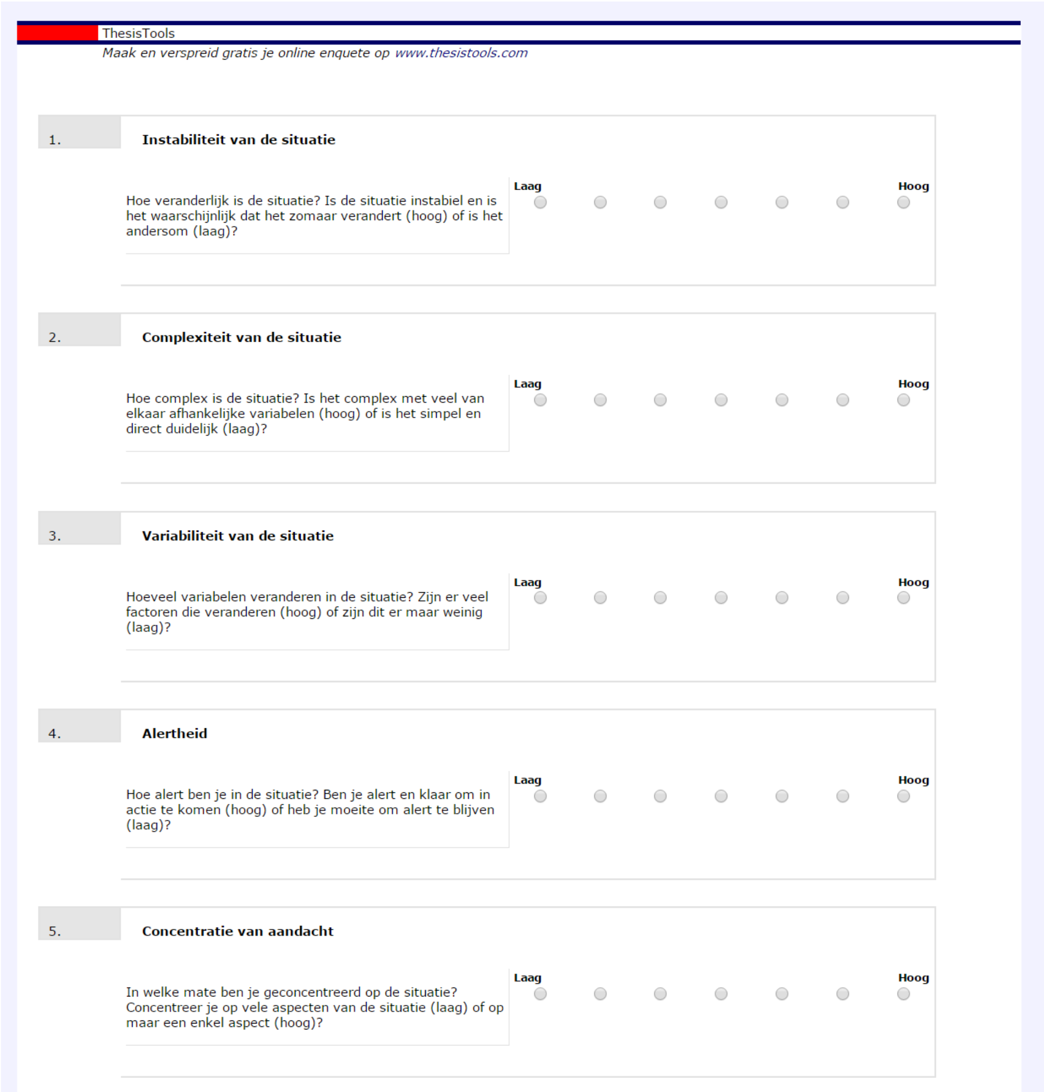
Information Quantity
How much information have you gained about the situation? Have you received and understood a great deal of knowledge (High) or very little (Low)?

1 2 3 4 5 6 7

Familiarity with Situation
How familiar are you with the situation? Do you have a great deal of relevant experience (High) or is it a new situation (Low)?

1 2 3 4 5 6 7

Figure 33: SART questionnaire as introduced by Taylor (1990).



ThesisTools

Maak en verspreid gratis je online enquete op www.thesistools.com

1. **Instabiliteit van de situatie**

Hoe veranderlijk is de situatie? Is de situatie instabiel en is het waarschijnlijk dat het zomaar verandert (hoog) of is het andersom (laag)?

Laag ○ ○ ○ ○ ○ ○ ○ Hoog

2. **Complexiteit van de situatie**

Hoe complex is de situatie? Is het complex met veel van elkaar afhankelijke variabelen (hoog) of is het simpel en direct duidelijk (laag)?

Laag ○ ○ ○ ○ ○ ○ ○ Hoog

3. **Variabiliteit van de situatie**

Hoeveel variabelen veranderen in de situatie? Zijn er veel factoren die veranderen (hoog) of zijn dit er maar weinig (laag)?

Laag ○ ○ ○ ○ ○ ○ ○ Hoog

4. **Alertheid**

Hoe alert ben je in de situatie? Ben je alert en klaar om in actie te komen (hoog) of heb je moeite om alert te blijven (laag)?

Laag ○ ○ ○ ○ ○ ○ ○ Hoog

5. **Concentratie van aandacht**

In welke mate ben je geconcentreerd op de situatie? Concentreer je op vele aspecten van de situatie (laag) of op maar een enkel aspect (hoog)?

Laag ○ ○ ○ ○ ○ ○ ○ Hoog

Figure 34: The SART questionnaire as it was administered with a digital form.

B SAGAT questionnaires

The SAGAT questionnaire consists of five times three questions that were conducted at regular intervals during the experiment. The amount of correct answers resulted in the SA score.

B.1 SAGAT questionnaire 1

1. **Nadat het stempel vervangen is, hoeveel CTQ's zijn er veel bijgestuurd?**
    - Twee of minder
    - Meer dan twee
    - Nul, het proces was stabiel
    - Weet ik niet
2. **Welke CTQ had meer bijsturing nodig?**
    - De buitenstraal
    - De torsie
    - Beide evenveel
    - Er was geen bijsturing nodig
    - Weet ik niet
3. **Wat was de situatie geweest als er niet was bijgestuurd?**
    - Buitenstraal en torsie waren buiten controle limiet geweest.
    - Alleen de buitenstraal was buiten controle limiet geweest.
    - Alleen de torsie was buiten controle limiet geweest.
    - Alles was binnen controle limieten gebleven.
    - Weet ik niet
4. **De torsie werd actief bijgestuurd, voor hoeveel van de 9 pootjes was dit het geval?**
    - Een pootje
    - Twee tot drie pootjes
    - Vier of meer pootjes
    - Weet ik niet
5. **Naar welke kant werden de spie's van de torsie gestuurd?**
    - Allemaal naar boven
    - Allemaal naar beneden
    - Naar beide kanten
    - Weet ik niet
6. **Op basis van de huidige trend, bij welke CTQ krijg je een probleem?**

- Bij de torsie
- Bij de buitenstraal
- Bij beide CTQ's
- Er komt geen probleem
- Weet ik niet

7. **In het gedeelte voor de materiaal wissel, hoeveel CTQ's waren er buiten de controle limieten?**
    - 0
    - 1
    - 2
    - 3
    - Weet ik niet
8. **Voor de materiaalwissel, wat was de situatie die plaatsvond?**
    - Er was een CTQ die over de controle limiet heen ging
    - Er was een spie die aan het einde van zijn bereik zat
    - Beide bovenstaande uitspraken zijn waar
    - Er was niks aan de hand
    - Weet ik niet
9. **Wat is een mogelijke oorzaak voor deze situatie?**
    - De spie reageerde niet op aansturing
    - De CTQ week door het process erg snel af
    - Beide bovenstaande oorzaken zijn mogelijk
    - Geen van de genoemde oorzaken is mogelijk
    - Weet ik niet
10. **Van hoeveel CTQ's zat de spie op het maximale bereik?**
    - 0
    - 1
    - 2
    - 3
    - Weet ik niet
11. **In de afgelopen serie, wat was de situatie?**
    - Er was een CTQ die over de controle limiet heen ging
    - Er was een spie die aan het einde van zijn bereik zat
    - Beide bovenstaande uitspraken zijn waar

- Er was niks aan de hand
- Weet ik niet

12. **Wat is een mogelijke oorzaak voor deze situatie?**

    - Het materiaal werd plotseling 10 mu dikker
    - De CTQ week door het process erg snel af
    - Beide bovenstaande oorzaken zijn mogelijk
    - Geen van de genoemde oorzaken is mogelijk
    - Weet ik niet

13. **Hoeveel producten zijn er geproduceerd buiten de controle limiet op de MP26 in het afgelopen deel van het experiment?**

    - Minder dan 100
    - Tussen de 100 en 200
    - Tussen de 200 en 300
    - Meer dan 300
    - Weet ik niet

14. **De materiaaldikte fluctueerde in het laatste deel. Hoe groot was deze fluctuatie?**

    - Plus minus 1 mu
    - Plus minus 2 mu
    - Plus minus 3 mu
    - Weet ik niet

15. **Had de dikte invloed op de CTQ's en zo ja, op welke?**

    - Ja, op de buitenstraal en de torsie
    - Ja, alleen op de buitenstraal
    - Ja, alleen op de torsie
    - Nee
    - Weet ik niet

B.2 SAGAT questionnaire 2

1. **Nadat het stempel vervangen is, hoeveel CTQ's zijn er veel bijgestuurd?**
   - Twee of minder
   - Meer dan twee
   - Nul, het proces was stabiel
   - Weet ik niet
2. **Welke CTQ had meer bijsturing nodig?**
   - De buitenstraal
   - De torsie
   - Beide evenveel
   - Er was geen bijsturing nodig
   - Weet ik niet
3. **Wat was de situatie geweest als er niet was bijgestuurd?**
   - Buitenstraal en torsie waren buiten controle limiet geweest.
   - Alleen de buitenstraal was buiten controle limiet geweest.
   - Alleen de torsie was buiten controle limiet geweest.
   - Alles was binnen controle limieten gebleven.
   - Weet ik niet
4. **De buitenstraal werd actief bijgestuurd, voor hoeveel van de 9 pootjes was dit het geval?**
   - Een pootje
   - Twee tot drie pootjes
   - Vier of meer pootjes
   - Weet ik niet
5. **Naar welke kant werden de spie's van de buitenstraal gestuurd?**
   - Allemaal naar boven
   - Allemaal naar beneden
   - Naar beide kanten
   - Weet ik niet
6. **Op basis van de huidige trend, bij welke CTQ krijg je een probleem?**
   - Bij de torsie
   - Bij de buitenstraal
   - Bij beide CTQ's
   - Er komt geen probleem

- Weet ik niet

7. **In het gedeelte voor de materiaal wissel, hoeveel CTQ's waren er buiten de controle limieten?**
   - 0
   - 1
   - 2
   - 3
   - Weet ik niet
8. **Voor de materiaalwissel, wat was de situatie die plaatsvond?**
   - Er was een CTQ die over de controle limiet heen ging
   - Er was een spie die aan het einde van zijn bereik zat
   - Beide bovenstaande uitspraken zijn waar
   - Er was niks aan de hand
   - Weet ik niet
9. **Wat is een mogelijke oorzaak voor deze situatie?**
   - De spie reageerde niet op aansturing
   - De CTQ week door het process erg snel af
   - Beide bovenstaande oorzaken zijn mogelijk
   - Geen van de genoemde oorzaken is mogelijk
   - Weet ik niet
10. **Van hoeveel CTQ's zat de spie op het maximale bereik?**
    - 0
    - 1
    - 2
    - 3
    - Weet ik niet
11. **In de afgelopen serie, wat was de situatie?**
    - Er was een CTQ die over de controle limiet heen ging
    - Er was een spie die aan het einde van zijn bereik zat
    - Beide bovenstaande uitspraken zijn waar
    - Er was niks aan de hand
    - Weet ik niet
12. **Wat is een mogelijke oorzaak voor deze situatie?**

- Het materiaal werd plotseling 10 mu dunner
- De CTQ week door het process erg snel af
- Beide bovenstaande oorzaken zijn mogelijk
- Geen van de genoemde oorzaken is mogelijk
- Weet ik niet

13. **Hoeveel producten zijn er geproduceerd buiten de controle limiet op de MP26 in het afgelopen deel van het experiment?**

    - Minder dan 100
    - Tussen de 100 en 200
    - Tussen de 200 en 300
    - Meer dan 300
    - Weet ik niet

14. **De materiaaldikte fluctueerde in het laatste deel. Hoe groot was deze fluctuatie?**

    - Plus minus 1 mu
    - Plus minus 2 mu
    - Plus minus 3 mu
    - Weet ik niet

15. **Had de dikte invloed op de CTQ's en zo ja, op welke?**

    - Ja, op de buitenstraal en de torsie
    - Ja, alleen op de buitenstraal
    - Ja, alleen op de torsie
    - Nee
    - Weet ik niet

C SUS

Each statement is answered on a five-point scale, from strongly disagree to strongly agree.

Vul bij elke stelling in hoeverre je het eens of oneens bent met de stelling.

1. Ik denk dat ik deze HMI vaak zou gebruiken als die er was.
2. Ik vond de HMI onnodig complex.
3. Ik vond de HMI makkelijk te gebruiken.
4. Ik denk dat ik ondersteuning van een technisch persoon nodig heb om deze HMI te gebruiken.
5. Ik vond de verschillende functies in de HMI goed gentegreerd.
6. Ik vond dat er teveel inconsistentie was in de HMI.
7. Ik stel me voor dat de meeste mensen makkelijk met deze HMI leren omgaan.
8. Ik vond het ongemakkelijk om deze HMI te gebruiken.
9. Ik voelde me zeker in het gebruik van de HMI.
10. Ik moest veel leren voordat ik de HMI kon gebruiken.

D Results table

Table 6: Data for both SA measures on different conditions. SAGAT values are shown as %. Total SART scores are calculated with a formula, thus it's sub-parts are shown as absolute values.

	SAGAT				**SART**				
	L1	**L2**	**L3**	**Total**	**D**	**S**	**U**	**Total**	**%**
HMI Visible	38,17	47,33	35,11	40,20	12,89	17,07	13,44	17,63	38,33
HMI Hidden	44,27	46,56	35,88	42,24	12,37	14,81	13,55	15,26	33,17
Total	41,22	46,94	35,49	41,22	12,63	15,94	13,49	16,44	35,74