Infusion pumps and patient safety

A comparison of Infusion pump interfaces by means of psychophysiological measures

By Jan Sommer

07 April 2014

Master Thesis

Jan Sommer

Student number: 1024833

07 April, 2014

Faculty of behavioral sciences
Department of Cognitive Psychology & Ergonomics (CPE)
University of Twente, Enschede, the Netherland

EXAMINATION COMITTEE

1. **Dr. Matthijs Noordzij (University of Twente, Department CPE)**
2. **Prof. Dr. Jan Maarten Schraagen (TNO Behavioural and Societal Sciences)**

UNIVERSITY OF TWENTE.

TNO innovation for life

Preface

This Master's thesis concludes my studies of Human Factors and Media Psychology at the University of Twente. I will hereby grasp the chance and thank a few people, who have accompanied me during this time.

First, I want to thank my supervisors Dr. Matthijs Noordzij and Prof. Dr. Jan Maarten Schraagen for their advice, good support and feedback throughout the course of this project. Moreover, I would like to thank Raphaela Schnittker and Frauke van Beek for their cooperation during the planning and conducting of the experiment.

Of course a lot of thanks also go to the whole project team and all participants, without whom the whole project would have been impossible to conduct.

Furthermore, I want to thank my parents, brother and my friends for their support during my years of study at the University of Twente. It has been a great time, to which I will surely look back with gratitude.

Jan Sommer

Enschede, 07.04.2014

Abstract

English:

Infusion and thereby infusion pumps are an irreplaceable asset in modern health care. However, there are 56.000 adverse events related to them reported yearly. These are mainly due to use errors. Therefore, a new prototype interface was developed and its usability was tested against a commonly used one. Nurses from both the intensive care and the nursing department were asked to complete three sets of comparable tasks for both interfaces, during which electrodermal activity of all participants was recorded. Results indicate no significant differences in habituation between both interfaces as well as no differences on psychophysiological parameters of amplitude, NS.SCR frequency and SCL. Further, these psychophysiological measures did not correlate well with subjective measures of workload as well. Tasks, which exhibited a high error rate often showed higher EDA parameters, even though this relationship was not significant. It is concluded that the prototype interface is more usable than the old one, although EDA could not differentiate between them. This, however, is a positive outcome in itself considering the short design-cycle of the new interface. Considering the host of innovations which have been introduced with the prototype interface, even significantly higher EDA values would have been perfectly natural.

Dutch:

Infusies en zodoende infuuspompen zijn een onvervangbaar deel van moderne ziekenzorg. Desondanks zijn er jaarlijks 56.000 ongelukken gerelateerd aan gebruiksfouten met infuuspompen. Daarom werd een prototype interface ontwikkeld en werd deze getest in vergelijking met een vaak gebruikte interface. Verpleegkundigen van de intensive care en de verpleegafdeling werden tijdens het experiment gevraagd om drie maal een set van 8 verglijkbare taken met elke interface uit te voeren. Daarbij werd van elke proefpersoon de elektrodermische activiteit opgenomen en geanalyseerd. Er werd geen verschil tussen de interfaces m.b.t. habituatie, NS.SCR frequentie, amplitude en SCL gevonden. Verder correleerden deze psychofysiologische maten slecht met subjectieve maten van werkbelasting. Taken met hoge fout-frequentie toonden vaak hoge EDA parameters, hoewel deze relatie niet significant was. Wij concludeerden dat het nieuwe interface gebruiksvriendelijker is, ook als EDA parameters niet tussen de interfaces konden onderscheiden. Dit is echter al een positieve uitkomst op zich gezien de korte ontwerpcyclus van de nieuwe interface. Gezien de innovaties en veranderingen die in het nieuwe interface zitten waren ook significant hogere
EDA waardes volledig begrijpelijk geweest.

Contents

1. Introduction....3
1.1. Prototype interface....5
1.2. Usability....9
1.3. Workload....9
1.4. Electrodermal activity....13
1.5. Research questions....20
2. Methods....21
2.1. Participants....21
2.2. Apparatus and recordings....21
2.3. Tasks....22
2.4. Procedure....23
2.5. Measures....25
2.6. Analysis....25
3. Results....26
3.1. Differences in NS.SCRs....27
3.1.1. Differences between measures....27
3.1.2. Differences between tasks....28
3.2. Differences in amplitude....29
3.2.1. Differences between tasks....30
3.3. Differences in SCL....30
3.3.1. Differences between measures....30
3.3.2. Differences between tasks....31
3.4. Differences between user groups....31
3.5. Correlation between objective and subjective measures of workload....31
3.5.1. Correlation between NS.SCRs per 5 seconds and BSMI....32
3.5.2. Correlation between amplitude per 5 seconds and BSMI....33
3.6. EDA measures and errors....33
3.7. Analysis with more conservative/lenient definitions of NS.SCRs....35
3.7.1. Applying new criterions for NS.SCRs regarding change in skin conductance....35
3.7.2. Changing the criterion for NS.SCRs regarding speed changes....36
4. Discussion....37
4.1. Differences between the interfaces....37
4.2. Habituation....39

4.3. Objective and subjective workload 40
4.4. Electrodermal activity and error rates 42
4.5. Decrease in SCL 43
4.6. Limitations 43
5. Conclusion 44
References 46
APPENDIX A: BSMI 52
APPENDIX B: Tasks and scenarios for both user groups used in experiment 53
APPENDIX C: Pre-questionnaire, welcome & instruction, informed consent 65
APPENDIX D: Post interview questions 68
APPENDIX E: SPSS Syntax 70

1. Introduction

In healthcare today, infusion pumps are an irreplaceable part of treating patients because they enable physicians to infuse fluids, medication or nutrients into a patient's circulatory system in a manner which would be impractical, expensive or unreliable if performed manually by nursing staff. These uses of infusion pumps include very small injections, injections every minute, injections with repeated boluses (medication administered during a running infusion to raise its concentration in blood to an effective level) requested by the patient, up to maximum number per hour or fluids whose volumes vary by the time of day. Yet, despite all these advantages infusion pumps were associated with more than 56,000 adverse event reports from 2005 to 2009, which include a minimum of 500 deaths (FDA, 2011). Moreover, infusion pumps are involved in about 30% of all reported (irreversible) incidents in the ICU (Intensive Care Unit) and OR (Operation Room) (Bogner, 1994). The U.S. Food and Drug Administration (FDA) concluded that the most common causes for errors with infusion pumps are:

Software malfunctions like failing to activate pre-programmed alarms when problems occur, while others activate an alarm in the absence of a problem.

Mechanical or electrical failures like components breaking under routine use and premature battery failures.

User interface (Human factors) issues such as confusing or unclear on-screen user instructions, which may lead to improper programming of medication doses or infusion rates.

According to the FDA, however, it is hard to discern the actual causes for pump failure and use error. Yet, according to Division of Electrical and Software Engineering Director Al Taylor (2010), FDA-run investigations have reached the conclusion that “Many adverse events are caused by design deficiencies that were foreseeable and preventable. Pump deficiencies place an undue burden on users, caregivers, and support staff, adding to an already stressful environment”.

These stressful environments (such as ICU and OR) are referred to as high risk areas because of their dynamic and complex nature with high activity levels, mental load and extensive use of technology and time stress (Bogner, 1994). It is quite apparent that good user-centered design of medical equipment is imperative in such environments in order to avoid a host of interface issues with infusion pumps, as poorly designed human-machine interfaces in medical equipment increase the risk of human error (Hyman, 1994; Obradovich and Woods, 1996). Prominent interface issues that were reported to the FDA during their *Infusion Pump Improvement Initiative* (2010) include

1. Confusing screens or faulty response during inappropriate data entry
2. Not making clear which unit of measurement the user is expected to enter
3. Unspecific or unclear user instructions, which may lead to under-/over-infusion
4. Inadequately designed alarm functions and settings cause users to miss problems or respond too late.
5. The infusion pump screen design is clunky or confusing to users, causing a delay in therapy
6. Warning messages are unclear. In the example below, for instance, it is unclear whether the user is confirming the warning message or the infusion settings.

> Volume in the syringe is inadequate to deliver the programmed dose.
> PRESS CONFIRM.

Figure 1. Example of an ambiguous warning message

Such issues are quite disturbing in the light of the estimate that 90% of all hospitalized patients receive an infusion as part of their treatment (Husch et al., 2005).

Therefore, the project group "Safer interfacing" (which will be described in detail below) proposed an infusion pump interface paying special attention to human factors guidelines and principles in order to eliminate or at least lessen the impact of user interface issues on patient safety. As testing the usability of this prototype interface and comparing it to another commonly used one by means of workload measures is the primary subject of this thesis the prototype interface as well as the concepts of usability and workload will be discussed in the following sections.

1.1. Prototype interface

The project group "Safer interfacing" (part of the project ‘Patient safety’ which is funded through the “Pieken in de Delta”-program by the Ministry of Economic Affairs, Agriculture and Innovation, and the city of Utrecht and the province of Utrecht) developed the new prototype interface by investigating user requirements from literature and expert users. These included standards and guidelines from the FDA (Sawyer, 1996) and the Association for the Advancement of Medical Instrumentation (2009) which proposed solutions to the issues described above. Thereafter, iterative paper prototype testing with target user groups revealed possible usability problems which could lead to use-related hazards. Such a user-centered design has been shown to produce medical devices which are less prone to use error and require less training to use (Sawyer, 1996). Schmettow, Vos and Schraagen (2013) conducted a usability test of the prototype and adjustments to the latter were made in accordance to the usability problems and anticipated potential hazards which were found. By this approach two of three stages of User-centered Design (Gould & Lewis, 1985) have been covered, namely:

1. Iterative Development
    (a) Usability requirements are a moving target
    (b) Iterate between design and evaluation of design

2. Participation
    (a) Know your users, know their tasks
    (b) Involve users in design early

The third principle, however, has not yet been attended to. This is:

3. Empirical testing
    (a) Measure performance of interaction
    (b) Evaluate design via direct behavioral observation

In the present study, this last principle will be concentrated on. In order to do this the newly developed interface (Figure 2) will be simulated on a tablet computer and performance will be tested against a commonly used interface (Braun Perfusor® Space, see Figure 2). Hereby a set of critical scenarios, including critical tasks, will be repeatedly tested with a group of expert users (nurses). This approach is in accordance to user-centric design and will determine in how far a newly developed interface that is designed with Human Factors principles in mind is in fact more usable than a common contemporary interface.

Some of these principles become apparent when comparing both the Braun and the prototype interface in Figure 2:

(1) *Size and visibility of syringe*:

One of the prototypes new features is that the syringe is placed under the interface instead of behind it and is therefore visible to the user. Thus, in addition to being displayed on the screen (as is the case with the Braun interface) the type of medication is also visible on the label of the syringe. To further clarify the medication used, differently colored backgrounds are used for distinct groups of medication. For example, painkillers are presented with a green background coloring, while vasoactives show red coloring. This feature is not present in the Braun Interface. Presenting the medication on both the display and the syringe is in accordance to Wickens et al.'s (2004) fifth principle of display design, namely redundancy gain. Hereby, Wickens et al. claims that presenting a signal in more than one way increases the likelihood it will be interpreted correctly. In the case of the prototype, this is not only done by presenting the medication twice, but also by the different color of each medication.

(2) *Input mode:*

There are some striking differences regarding the operability of both interfaces. While the only mode of input of the Braun Interface is through same-sized physical buttons at the right-hand side of the pump, the newly developed Interface offers a greater variety of input modes for different functions. The most basic functions of the infusion pump such as starting and stopping the pump, switching it on and off and locking it are still handled by physical functions. The adjustment of infusion rate, volume and time (the black area of the Interface with the up/down arrows), as well as browsing the Information display (the two grey buttons at the top-right of the Interface) are operated via a touchscreen.

(3) *Menu space*:

Another huge difference between both Interfaces is the menu structure. The Braun Interface requires one to navigate a deep menu structure (by means of the four arrow buttons) in order to access many functions of the pump. In this way only one mode can be accessed and adjusted at a time, requiring the user to navigate through the menu again, when another function is needed.

In contrast to this, the new Interface has a flat menu structure which is directly accessible after starting the pump. All relevant modes (rate, volume and time) are immediately presented and adjustable on the main screen. The bolus function, which is switched on by pressing either the automatic or manual bolus-button, is also directly accessible without any further navigation. Thus, Wickens et al.'s (2004) principle of minimizing information access cost is satisfied. It states that frequently accessed sources of information should be readily available and that certain information is always important and should not require anything but minimal effort to access (e.g. the infusion rate of an infusion pump). As mentioned above, important information is immediately visible and one does not even have to enter a menu to access it. Moreover, the principle "Replace memory with visual information: knowledge in the world" is satisfied by this approach. A user should not need to retain important information solely in working memory or retrieve it from long-term memory. A menu, checklist, or another display can aid the user by easing the use of their memory. As all information is either immediately accessible or retrievable though up to three clicks, users do not need to employ their working- or long-term memory in order to effectively use the infusion pump. Another principle has been incorporated into the prototype, namely proximity compatibility. Often, two or more sources of information are related to the same task. These sources must be mentally integrated and are defined to have close mental proximity. This is also true for the task of using an infusion pump. Rate, volume and time are interrelated variables, which need to be mentally integrated by an operator in order to use an infusion pump correctly. As all of these variables are presented directly adjacent to each other, information on these variables can be easily accessed and integrated. Nonetheless, care must be taken when applying this principle as close display proximity can be harmful by causing too much clutter. In the prototype this is prevented by a clear delineation between the three modes (see Figure 2).

(4) Buttons and functions:

As mentioned above, the Braun Interface utilizes buttons of uniform size and shape. These are labeled with distinct symbols and are partially colored. Hereby, some of the buttons are multifunctional and some buttons share functionality. An example of the former is the start/stop button, which is used both to initiate and terminate an infusion. The latter is exemplified by the left-arrow- and 'OK'-buttons, which can both be used to access the menu.

The prototype Interface on the other hand consists of buttons of different size and shape, distinctly colored and labelled. In further contrast to the reference interface, each button only serves one distinct, irreplaceable purpose. These different functionalities are clarified by the spatial dissociation of the buttons. For example, buttons for adjusting each of the main functionalities (adjusting infusion rate, time and volume) are placed directly above each of the digits (decimals to three-digit numbers) of the respective mode. Additionally, there are individual buttons for giving a bolus and separate ones for starting and stopping an infusion (which is both not the case in the reference interface). Furthermore, the function of locking/unlocking the interface has been added.

By means of giving buttons different shapes, the principle "use discriminable elements" is integrated into the prototype. This principle states that similar appearing signals are likely to be confused. As the Braun interface uses uniform buttons, it is likely much harder to confuse buttons on the prototype interface than it is on the Braun interface. Giving different functions, different shapes, colors and sizes is also related to Wickens et al.'s (2004) redundancy gain principle as functions are not only differentiated by form, but also by size and color. Again, this supports the likelihood of signals/buttons being interpreted correctly. Another principle that is satisfied in relation to button design is the principle of consistency. A user's long-term memory will trigger actions that are expected to be appropriate. Familiar icons, actions and procedures from other displays will easily transfer to support processing of new displays if they are designed in a consistent manner. A design must accept this fact and utilize consistency among different displays. This is exemplified by the "ON" and "OFF" buttons in the new interface. Both the color (red for off, green for on) and symbol (see Figure 2) are consistent with POWER buttons across a wide variety of technological devices. A further example of this is the question mark as identifier of the HELP function.

` Before we develop the research questions related to this experiment, however, it is important to discuss the concepts of "Human factors engineering", "usability", and specifically "workload" first, as they play a critical role in the evaluation of medical devices such as infusion pumps.

Figure 2. Scaled representations of both the Braun (top) and the prototype Interface (bottom) developed with human factors principles in mind

1.2. Usability

Human factors engineering is a multi-disciplinary science which seeks to improve the ease of use with which technologies can be employed by end-users by designing them to fit operators' cognitive abilities and needs. One principal concept in human factors engineering is **usability**. Usability is commonly defined as "the *effectiveness*, *efficiency* and *satisfaction* with which specified users achieve specified goals in particular environments"(ISO, 1994). Hereby, effectiveness is defined as the degree to which users can attain these goals. Indicators of effectiveness are task completion and error rates. Efficiency on the other hand is the degree of effort users have to invest to reach said goals. This dimension of usability is often measured as task completion time or learning rate. Satisfaction is simply the user's subjective experience when/after using the product. Satisfaction is typically measured through validated questionnaires.

1.3. Workload

Workload is a concept strongly related to the concept of usability.
There are various definitions of workload, but one that suits the purposes of this paper well is Hart and Staveland's (1988) definition of workload as "the perceived relationship between the amount of mental processing capability or resources and the amount required by the task".

According to this definition, a low workload while using a product should indicate high usability of said product.

However, the issue of determining the mental resources available for processing information and defining their nature is no easy one and throughout the course of the last 50 years scientists have struggled to find fitting ways in modeling workload. One of the first theories which tried to model workload was Kahneman's single resource theory (1973), which proposed that there is but one resource for distributing mental processing capacity. According to this theory the workload of different cognitive processing tasks simply stacks up additively until mental workload becomes too big to handle. Workload was often operationalized as two-task performance, where performance on a primary task is measured with and without a secondary task. However, it was soon found out that while some dual-task pairings show a *decrement* compared to single-task performance, others can be performed concurrently as well as they can in isolation (e.g., reading music and playing in skilled pianists). This gave rise to one of the most popular models of workload: Wickens' (1984) multiple-resource theory. This theory originally proposed information processing as dependent on the 3 different stages of processing (perception, processing, action), two codes of processing (spatial, verbal) and two modalities of encoding (visual, auditory). Response modalities (manual, vocal) were naturally also thought to influence workload during concurrent tasks. This model explains quite well why some tasks can be performed in parallel without a decrement to performance, while others have to be performed serially. Over the course of the last 30 years this model has steadily become more sophisticated, including tactual and olfactory input dimensions, different modes of reasoning (subconscious, symbolic, linguistic) and splitting visual processing into focal and ambient processing. By means of this sophistication the multiple resource model (see Figure 3) has become able to predict, which tasks can be conducted concurrently, when task-interference occurs and when increases in difficulty of one task induce a performance loss in the second task.

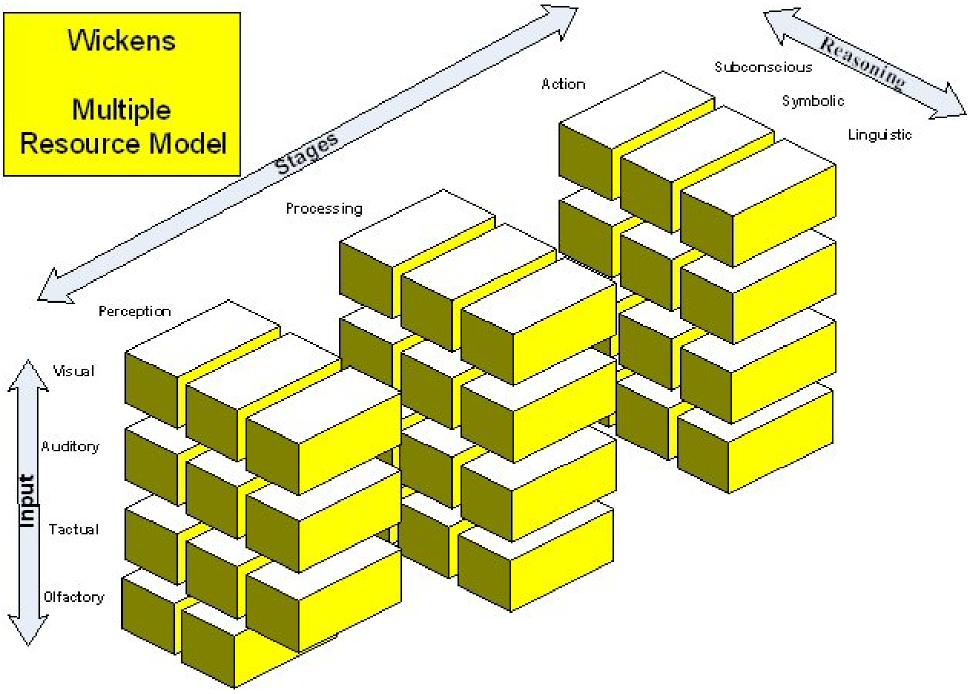


Figure 3. A recent version of Wickens' Multiple Resource Model

However, despite general scientific consensus of task performance being multi-modal, there are many different operationalisations of workload, which produce quite distinct measures. Aside from the mentioned performance in single- and dual-task situations, workload is often assessed in the form of questionnaires containing Likert-scales or simple scales that help evaluate workload. Due to this similarity in measurement to satisfaction ratings and its strong correlation with such measures, it is often categorized as a satisfaction-measure (Hornbaek, 2006). Yet, other authors consider workload to be an efficiency-measure (Hornbaek, 2006). This inconsistency in classifying workload may be due to still another kind of measuring workload, namely objective/psychophysiological measures. Before going into detail about these measures (and specifically electrodermal activity) one has to state that these measures generally operationalize workload as arousal. It is hereby assumed that a higher level of arousal indicates a higher level of workload. However, a low level of arousal (underload) seems to be as dangerous as a high one (overload) (see Figure 4). This is in accordance to the so-called Yerkes-Dodson Law of performance and arousal which predicts performance to be lowest when people are scarcely or overly aroused (Yerkes & Dodson, 1908).

Nowadays, there are a host of different psychophysiological measures, which are utilized to determine the level of one's workload/arousal. Some common measures include heart-rate variability and EEG (e.g. Izso and Lang, 2000). Among these automatic nervous system measures, however, tonic EDA parameters have been for a long time the most frequently used indicator of arousal in psychophysiological research (Duffy, 1972). These parameters will be discussed in detail in the following section.

As these purely psychophysiological measures of mental effort completely disregard the valence dimension of affective experience (see the circumplex model of emotion developed by James Russell, 1980) it is not hard to see why many authors view workload as an efficiency measure. That is, psychophysiological measures by themselves are unable to distinguish positive from negative valence, but can only measure whether a person is aroused or not.

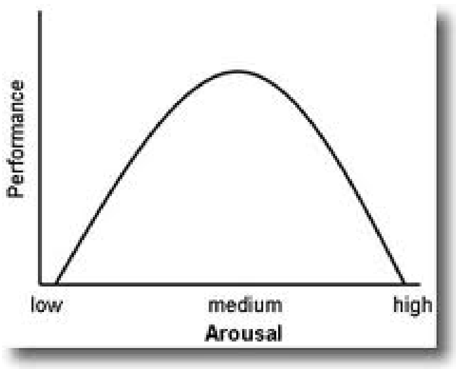


Figure 4. Simple graphic representation of the Yerkes-Dodson law

Because of the inconsistencies in measuring and categorizing workload and the inability to distinguish emotions during psychophysiological measures mentioned above, it will be interesting to see if and how subjective workload scales (such as the BSMI scale, Zijlstra & Van. Doorn, 1985) correlate with objective workload measures, such as EDA. For example, Novak et al. (2010) found that psychophysiological measures of workload do not always agree with participants' subjective workload or performance. One explanation could be that subjective workload scales might be influenced by the valence dimension of affective experience. This could be in such a way that positive valence (high satisfaction scores) for one interface might decrease subjective workload for that interface. Objective workload, however, would not be influenced by the valence dimension of affective experience, creating a disparity between subjective and objective workload measures. One indication for such a phenomenon is Cárdenas et al.'s findings (2013) that perceived exertion (physical effort) is strongly correlated with reported emotional/hedonic valence of a task, but not so tightly with reported arousal. Therefore, perceived mental effort might not correlate with arousal (as measured by EDA) that well, either.

1.4. Electrodermal activity

'Skin conductance' or 'electrodermal activity' is a measure of the electrical conductance of the skin (De Waard & Brookhuis, 1993). Variability in this conductance is due to the moisture level of the skin. What has made this measurement interesting for psychological research is that sweat glands are controlled by the sympathetic nervous system. Therefore, electrodermal activity is used as an indication of psychological or physiological arousal; when the sympathetic branch of the autonomic nervous system is aroused, sweat gland activity will also increase, which in turn increases skin conductance. Thus, electrodermal activity can be operationalized to measure emotional and cognitive agitation. Electrodermal activity, however, is not one unified measure but consists of a host of parameters that help us understand the kind and amount of cognitive and physiological strain put on people. An overview of prominent EDA parameters, their definition and typical values can be found in Table 1 and Figure 5.

Table 1. Electrodermal measures, definitions, and typical values (taken from Dawson et al., 2007)

Measure	Definition	Typical values
Skin conductance level (SCL)	Tonic level of electrical conductivity of skin	2-20 µS
Change in SCL	Gradual changes in SCL measured at two or more points in time	1-3 µS
Frequency of NS.SCRs	Number of SCRs in absence of identifiable eliciting stimulus	1-5 per min. during rest, over 20 in high arousal situations (Braithwaite et al., 2013)
SCR amplitude	Phasic increase of conductance shortly following stimulus onset	0.1-1.0 µS
SCR latency	Temporal interval between stimulus onset and SCR initiation	1-3 s
SCR rise time	Temporal interval between stimulus initiation and SCR	1-3 s

	peak	
SCR half recovery time	Temporal interval between SCR peak and point of 50% recovery of SCR amplitude	2-10 s
SCR habituation (trials to habituation)	Number of stimulus presentations before two or three trials with no response	2-8 stimulus presentations
SCR habituation (slope)	Rate of change of ER-SCR amplitude	0.01-0.5 µS per trial

For example, when analyzing electrodermal reactions to discrete stimuli (and thereby measuring for example novelty, surprise, intensity, arousal content or significance), phasic parameters such as SCR (skin conductance response) amplitude, latency and half-time are critical. A Skin conductance response is the phenomenon that the skin momentarily becomes a better conductor of electricity when either external or internal stimuli occur that are physiologically arousing. Hereby, high SCR amplitudes are correlated with a high significance of the presented stimulus (Dawson, Shell & Fillion, 2007). Furthermore, it has been found that a large SCR amplitude, high SCL, frequent NS.SCRs, short rise time, short latency and short recovery usually cluster together (Dawson, Shell & Fillion, 2007). When evaluating continuous stimuli, tonic parameters such as SCL and frequency of NS.SCRs become most important. Hereby, a high SCL and a high frequency of NS.SCRs are correlated to a high effect of the presented continuous stimuli (e.g. longitudinal tasks). Tonic and phasic EDA measures thus differentiate smooth underlying slowly-changing levels in EDA (tonic) from rapidly changing peaks in the EDA signal (phasic).

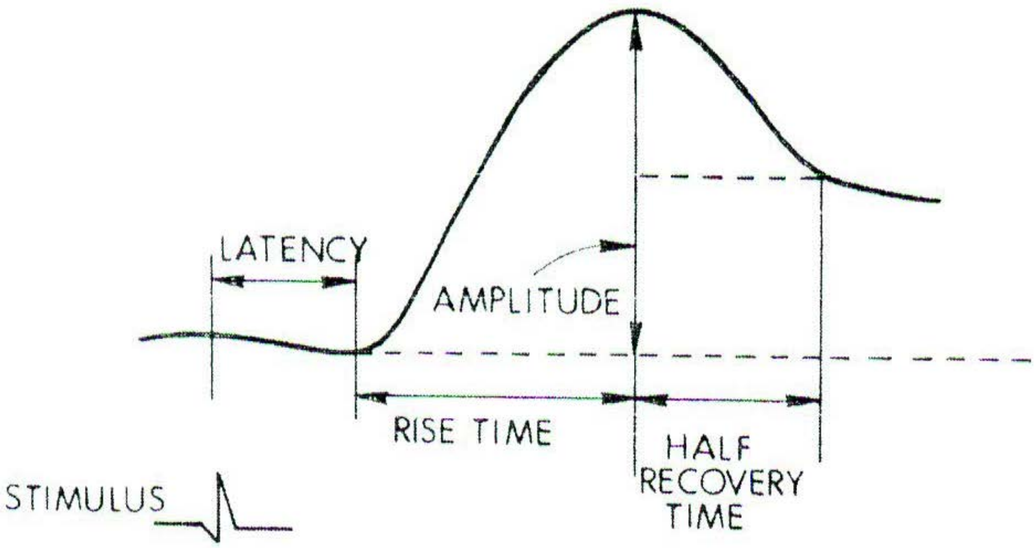


Figure 5. Graphical representation of principal EDA components (taken from Dawson et al., 2001)

Because electrodermal activity is highly correlated to sympathetic reactions of the nervous system it has a long history of application within the field of psychology and is said to be one of the most used response systems in the history of psychophysiology (Dawson, Shell & Fillion, 2007). Throughout the course of the last century it has found use in tackling a large variety of research questions and psychological applications such as the assessment of Anxiety, Psychopathy and Depression (e.g. Lader and Wing 1964, Lader and Wing 1966), Schizophrenia Research (Dawson & Shell, 2002) and the detection of deception (e.g. Lykken, 1981). Another prominent application of EDA is research on habituation and the orienting reflex. According to the classical definition given by Humphrey (1933) and Harris (1943), habituation is characterized by decreasing response intensity with repeated stimulation.

Typically, habituation in EDA is related to the amplitude of an SCR following a specific stimulus (Boucsein, 2012). This specific SCR is called an orienting response. When people get used to the stimulus, the amplitude of SCRs to that stimulus decreases. This has often been interpreted as a basic form of learning (e.g., Thorpe, 1969). The present study, however, does not present specific but rather continuous stimuli (in the form of tasks). Therefore, frequency of NS.SCRs and their amplitude will be used as an indicator of habituation/learning. Frequency of NS.SCRs has shown a quite stable correlation of r= 0.41 and r= 0.56 to the habituation index of mean amplitude in earlier studies (Bull & Gale, 1973; Martin & Rust, 1976). Changes in the NS.SCR frequency (and the skin conductance level) due to a stimulus are called tonic orienting responses in contrast to phasic ORs, which are characterized by changes in SCR amplitude (Sokolov, 1963).

Yet, while research on EDA components related to phasic SC-ORs is abundant, investigations aiming at the usability of EDA measures for tonic SC-OR components are sparse (Boucsein, 2012). In one of those instances, Wilson (2001) employed tonic EDA measures (SCL, NS.SCR frequency and amplitude of these) in his analysis of mental workload of pilots during flight. To my knowledge, this was one of the few cases where EDA in pilots during flight was measured outside of a laboratory setting. Hereby, it was found that the VFR (visual flight rules) takeoff, touch and go and the final landing exhibited the most NS.SCRs. The pre- and postflight baselines showed the fewest responses. All other tested segments (17) showed no significant differences in EDA (in contrast to heart rate, which found more differences). EDA amplitudes showed a similar pattern, while SCL produced a linear decline throughout the experiment. The tonic level does, however, show significant increases associated with VFR takeoff, VFR touch and go, and the final landing. During subjective assessments of workload, however, these tasks were not rated as highly demanding. This was concluded to be due to the practice pilots had with these situations and the missing practice they had with higher rated segments. Other researchers (Collet et al., 2003) have used other parameters of EDA as an indicator of workload during bus driving. These were skin resistance response and ohmic perturbation duration. They concluded that electrodermal activity recordings have been shown to be reliable tools in evaluating mental workload in the field. Yet, it was found again that physiological data seem to be inconsistent with the drivers' subjective responses. However, measures of skin resistance are not common nowadays and are often transformed into measures of skin conductance instead (Boucsein, 2012).

Other experiments have found EDA measures to correlate with subjective measures. One of these is Baldauf et al.'s (2009) measure of driving performance under different road conditions. Here –for both subjective and objective measures- driving in the city differed significantly from driving on a straight road, which in turn differed significantly from driving with oncoming traffic. In contrast to these findings, Seitz et al. (2012) found that while subjective measures of workload were sensitive to road conditions and maneuvers, electrodermal activity (measured as amplitude per second) did not lead to statistical significances between the investigated situations. Instead, the most demanding activity with regard to electrodermal activity during the performed study was making phone calls, whereby the participants were confronted with a planning task. Paying attention to an approaching car was also a high demanding task with anticipatory requirements with respect to the car driver's behavior. Seitz et al. therefore concluded that EDA is unfit to distinguish between routine tasks, but applicable when trying to detect cognitively high demanding planning or anticipation tasks. In another study regarding electrodermal activity in traffic and automation, De Waard found that when workload increases, so does skin conductance (1996). An early experiment related to driving and EDA, which asked participants to drive 20 different routes, found that SCRs seemed to be correlated to the experience of the driver rather than route conditions. Moreover, distribution of SCRs per km was comparable to the distribution of accidents per km (Taylor, 1964). Hence, EDA seems to be correlated to risk of accident, which is useful in a lot of fields of applied psychology.

Yet, performance and workload do not always correlate. For example, Mehler, Reimer and Coughlin (2012) found that SCL rose significantly with each difficulty level of an auditory presentation–verbal response working memory task during driving. However, driving performance measures did not provide incremental discrimination. In another study by Shimomura et al. (2008), subjects were asked to memorize target letters, detect them within a 4X4 alphabet arrangement, and answer whether the number of targets contained in the arrangement corresponded to a randomly displayed number. Hereby, the score on the card sort NASA Task Load Index (CSTLX) increased in correspondence to task performance. NS.SCR frequency and their amplitudes, however, did not show any significant effect of task difficulty.

A correlation between EDA and workload was also found for arithmetic and reading tasks (Nourbakhsh et al., 2012), which hints at a general applicability of EDA as a measure of workload. Additionally, EDA has been used in Human-Computer Interaction to measure arousal and emotional reaction (e.g. Drachen, Nacke, Yannakakis & Pedersen, 2010). In a study by Laufer and Németh (2008) EDA has even been applied in order to predict user action during gaming. They concluded that in the tested gaming situation the user actions can be concluded from the skin conductance level of the player. To them this is because in a game like the tested YetiSports, where the player has to choose the best time for clicking, the user anticipates the emotional stress of the click. This process of preparation for the action has a reliable pattern, from which the exact timing of the click can be inferred.

Using medical devices such as infusion pumps is in many respects similar to all types of tasks discussed earlier. It is similar to tasks related to HCI as nurses are interacting with interfaces when operating infusion pumps. Cognitively, this essentially is a Human-Computer Interaction. Yet, it is similar to traffic and automation tasks due to the risks associated to erroneous use of both vehicles and infusion pumps. Both of these environments (traffic, ICU/OR) constitute high risk environments due to their complex and dynamic nature. Therefore, arousal during infusion pump operation should be generally higher than during simple HCI. It remains to be seen if this is the case when simulating an infusion pump operation without a real patient. Regardless of the level of arousal/workload, though, EDA is a valid means of identifying reactions to novel stimuli, as discussed above. Identifying situations that produce high tonic ORs in this way is very helpful to the whole design process of technical devices, as weaknesses of the new product can be spotted and eliminated more easily.

Regarding workload and EDA, one has to keep in mind, that stress and workload are two distinct concepts. While workload is generally valence-neutral, stress is accompanied by either positive or negative valence creating either eustress (positive) or distress (negative) (Selye, 1956). This is exemplified in Setz et al.'s (2010) paper on discriminating stress from workload using a wearable EDA device. In this study stress was produced in the form of mental stress induced by solving arithmetic problems under time pressure and psychosocial stress induced by social-evaluative threat. It was concluded that the monitoring of EDA allows discrimination between cognitive load and stress with accuracy larger than 80% with leave-one-person-out cross validation. Hereby, EDA peak height and the instantaneous peak rate carry information about the stress level of a person. Specific values, however, were not given. Further investigations of stress and workload by Conwey (2012) underlined the distinction between cognitive load and stress. Subjects were asked to solve math problems at three different levels of difficulty. After solving the problems, the procedure was repeated while subjects were told that their performance would now be monitored, time-limits were introduced and a screen was switched on with researchers apparently observing the subjects (which actually was a pre-recorded video). SCL was shown to be significantly different between CL levels in the 'no-stress' condition, but not in the 'stress' condition. Conwey therefore concluded that the stress response overshadowed the signal-variation owing to cognitive load. This interaction between stress and workload measures might account for some of the different findings related to subjective/objective workload and EDA, as both concepts are scarcely differentiated. Because of the influence of stress on EDA parameters, it is imperative to safeguard participants from stress-inducing situations.

Moreover, Peters (1974) noted that while electrodermal changes appear mainly appear during mental tasks, EDA was highest when test participants spoke during his observation of 11 female phonotypists. This high amount of EDA must to his mind be regarded as being mainly due to an artifact. This is supported by Boucsein (2012) who argues that movement is the most important physiological source of artifacts in EDA recording. This includes not only skin movements beneath the electrodes, but also muscular activity being exerted not directly underneath electrodermal recording sites. Thus, to ensure an artifact-free EDA recording, gross body movements should be avoided during recording. Boucsein (2012) advises to tell the study participant to sit or lie quietly, to relax and to try to avoid movements, especially those of the limbs from which EDA is recorded.

Furthermore, Quantitative relationships between skin stretching at the volar side of the forearm and elicited EDA artifacts have been established by Burbank and Webster (1978). Yet, exosomatic measures of EDA (such as the one employed during this study) are not as likely to be influenced by these skin strechings, as endosomatic ones (Boucsein, 2012). For clarification, endosomatic measurements involve the application of tiny electrodes directly onto the 'sympathetic' skin neurons. This yields a direct measurement of the electrical activity of the skin's neurons. The exosomatic measurement employs two electrodes that are placed on the skin's surface and an electrical signal of tiny magnitude is passed over this surface between the two electrodes. Additionally, study participants may elicit voluntary EDRs by a deep inhalation and subsequent holding of their breath (e.g., Hygge & Hugdahl, 1985).

Another source of physiological artifacts is the influence of temperature on EDA recordings. It is therefore advisable to pay close attention to physiological artifacts when measuring EDA due to the wide range of disturbances that can occur.

As has become apparent, there are still a lot of inconsistencies regarding the relationship of measures of EDA, subjective/objective workload, performance and stress. This paper, however, addresses all of these concepts and tries to establish underlying relationships between them. Therefore, it fits perfectly into the line of contemporary research on problems regarding the measurement of EDA as an indicator of workload. The most common measurements of tonic EDA –namely SCL, NS.SCR frequency and their amplitude- are employed, which is consistent with state-of-the-art research on EDA. Under these premises some research questions will be tackled, which are discussed in the next section.

1.5. Research questions

Firstly, the main goal of this paper is to assess in how far the novel interface is more usable than an older, commonly used one. While two fellow students addressed the effectiveness and satisfaction dimension of usability, this paper will be focusing on an efficiency measure, namely objective workload. In order to tackle this question NS.SCR frequency, as well as their amplitudes and the SCL will be compared throughout the 3 measures in order to determine, whether the new interface is as efficient as the reference interface or can even surpass it with regard to efficiency of use. Yet, considering the earlier discussion on orienting responses even higher EDA values are to be expected. This is due to the novelty of the new interface. A lot of new features have been introduced with the prototype, which make it prone to higher EDA values than a more established device would generate.

Secondly, this experiment will also compare objective and subjective workload and see whether they correlate or if they measure two different concepts entirely. As seen before, EDA measures and measures of objective workload do not always correlate and it will be interesting to see, how their relationship is in regard to the use of infusion pumps. If EDA data and BSMI scores do not correlate, usability tests should distinguish the two concepts (objective and subjective workload) more rigidly.

Thirdly, this paper will focus on the learnability of both interfaces by means of comparing habituation processes. Habituation –as mentioned above- will be measured as decreasing NS.SCR frequency and amplitude with increasing exposure to tasks.

Lastly, it will be evaluated if a difference in objective workload is in fact correlated to more/less user errors. Again, a lot of controversial findings have been reported on this point with regard to EDA measures. Some studies found a correlation, while others did not. However, in the domain of working with medical products, no relationship between objective workload and error rate has ever been established (at least to my knowledge). Yet, as mentioned in the beginning, adverse advents related to infusion pumps are in the ten-thousands every year in the U.S. alone. Despite the similarities to other tasks discussed earlier, working with medical equipment differs from for example driving a car or piloting a plane as there often is no immediate external feedback when operating the former. This is reason enough to thoroughly investigate whether objective workload characteristics can predict error-proneness of a medical device in order to ultimately protect people from harm.

2. Methods

2.1. Participants

The participants sampled for this study consisted of employees of two Dutch University Medical Centres (UMCs). In total, 25 subjects participated of whom 5 were male and 20 female. 16 of the participants (64%) were employed at the UMC Groningen, while 9 participants worked at the UMC Leiden. Of the study sample, 13 subjects worked at the nursing department while 12 were employees of the intensive care unit. These two groups differ mainly in the pressure they work under. Decisions at the Intensive care unit have to be taken and followed through in much less time than at the nursing unit. According to FDA guidelines (2011), the groups therefore constitute two different populations, which will be analyzed separately. Prior experience with the Braun Perfusor Space pump was checked before testing and denied by all participants. The infusion pumps in use at the time were the Alaris syringe pump for the UMC Groningen and the Syramed syringe pump for the UMC Leiden. Usage of infusion pumps during working hours varied from zero to more than four times per day. The majority of participants stated a usage of either one or two times or more than four times a day (36 % in both cases). Further 20 % of the subjects stated to use infusion pumps three to four times a day, while the remaining 8% of the sample never using an infusion pump within the scope of their current employment.

Years of experience in using infusion pumps ranged from zero to 31 years (*M*= 15.2, s.d.= 1.917). The highest degree of education obtained for most subjects was HBO (70.8%). Further 16.7% held a MBO-degree, while 12.5% had obtained a WO-degree. For one participant the highest degree of education was left undisclosed. Lastly, at the moment of testing, the majority of subjects (10/40%) had already finished their shift, 7 (28%) participated before starting their shift, while 12% participated during their shift and 20% took part in the experiment on their day off.

2.2. Apparatus and recordings

For all participants testing took place in a silent and separated room in the respective UMCs they were employed at. This was done to ensure easy access and prevent distraction.

For presentation of the infusion pump interfaces a *Fujitsu* tablet was used. This tablet was connected to a *Dell* laptop via a router. By operating a task manager on the laptop, tasks could be sent to the tablet and recordings of button presses by the participants could be started/stopped. Furthermore, videos of all user actions were recorded by a *Sony Handycam* camcorder. EDA data was measured via a *Q-sensor curve* (Poh et al., 2010), which is an unobtrusive, portable device for measuring skin conductance. The *Q-sensor* was mounted on the wrist of the non-dominant hand in order to minimize the influence of movement on the EDA-data. Although uncommon, measures of EDA at the wrist are highly correlated to measures of EDA at the more traditional locations of the finger or palm (e.g. Poh et al., 2010; van Dooren et al., 2012). Even though we aimed to reduce the impact of movement, choosing the wrist also retains the possibility of using both hands during tasks, which would be inhibited with sensors at the palm/finger, enhancing naturalness of the tasks.

2.3. Tasks

In order to account for the different fields of employment of the participants two subsets of tasks were created to attain as realistic tasks as possible. The difference between these sets was that the task of giving a bolus (giving an extra dose of medicine to the patient in imminent need) was presented verbally to the intensive care group in order to account for the pressure they work under. Moreover, task descriptions differed slightly between the two groups considering their different work situations (different medication/different scenarios). Tasks for both groups were each created in such a way that everyday functions of a syringe infusion pump were included. This yielded seven different types of tasks (see Appendix B). From these tasks eight different scenarios were created per user group (see Appendix B). Scenarios were more sophisticated task descriptions. For the task “adjusting an infusion” one scenario was: “A patient of 61 years has undergone knee surgery but is still in critical condition and is therefore moved to the IC. In order to prevent pain as a consequence of the surgery Morphine (0.08 mg/ml) must be administered.”

Hereby, the task of adjusting an infusion was included twice per set of tasks as adjusting an already running infusion is a crucial function often used by nurses. For the Braun Perfusor infusion pump the "give a manual bolus" task was given twice per set of tasks, as the otherwise presented "automatic bolus" functionality was not a feature of this syringe infusion pump. Two further variations were created per set of tasks in order to ensure that participants did not have to perform the exact same task twice. Hereby, only numeric values were changed in order to ensure a similar level of difficulty. In this way, subjects performed each set of tasks exactly once for each type of interface (the Braun and novel interface). Face validity and realism of the tasks/scenarios was judged by some experts prior to conducting the experiment and was found to be sufficient. The sequence of 6 of the total 8 tasks was changed per variation of the task sets in order to prevent order and learning effects and thereby biased performance. The first and last task of each task-set (starting/shutting down the infusion pump) was held constant with the purpose of ensuring realism.

2.4. Procedure

Prior to testing and after welcoming the participants, the Q-sensor was mounted on the participants' wrist and all subjects were asked to walk along the hallway a few times. By doing this roughly 5 minutes before testing we aimed to ensure skin conductance by activating sweat glands. Thereafter, participants were seated in front of the tablet computer, were given general written instructions about the experiment and filled in a form of informed consent and a questionnaire asking for demographic data (for the instructions, see Appendix B; for the forms, see Appendix C). Upon completion, a tutorial showing the basic functions of the first interface was shown. The interface being selected first depended on the number assigned to the subject. Subjects assigned an odd number started the experiment by using the Braun Perfusor pump interface, whereas subjects with an even number started using the prototype interface. By choosing such a counterbalanced design, order effects were tried to be mitigated. Following the tutorial video, the first set of tasks had to be completed using the Fujitsu tablet computer. After completing this set of eight tasks for one interface, a tutorial video for the other interface was presented and the second set of tasks had to be completed. This process was repeated twice, thereby providing measures of all three variations of tasks per interface per participant (see Table 2). By repeatedly measuring each interface for each participant, a learning curve per interface could be established and compared.

Table 2. Procedure of testing all variations for both interfaces

Start with Braun Perfusor interface
Measure 1:
Variation 1 : Braun Perfusor interface

Variation 2 : New interface
Measure 2:
Variation 3 : Braun Perfusor interface Variation 1 : New interface
Measure 3:
Variation 2 : Braun Perfusor interface Variation 3 : New interface
Start with New interface
Measure 1:
Variation 1 : New interface Variation 2 : Braun Perfusor interface
Measure 2:
Variation 3 : New interface Variation 1 : Braun Perfusor interface
Measure 3:
Variation 2 : New interface Variation 3 : Braun Perfusor interface

Moreover, upon completion of each separate task, subjects were asked to give an estimate of their mental workload, using the BSMI scale (see Appendix A).

The Belasting Schaal Mentale Inspanning (BSMI) is a unidimensional subjective rating scale for mental effort (Zijlstra & Van Doorn, 1985).

After completion of all task-variations with both interfaces, a brief post-test-interview -as recommended by the FDA (2011)- was conducted. Hereby, opinions and preferences concerning the separate functions of both pumps and both pumps in general were asked for. Upon completion of the whole experiment, all participants received a 50 Euro gift voucher. During each experimental session at least two experimenters were present. One would give instructions during the experiment, while the other was responsible for starting and stopping the tasks on the task-controller, thereby saving them in log-files. Moreover, all button presses (and their respective times) were also saved in these files. Additionally, audio- and video recordings of all sessions were made in order to trace use errors more effectively. Each session took approximately 90 minutes to complete.

2.5. Measures

For later analysis, a number of measures were taken during the experiment. These included time-to-completion, error rate (number of incorrectly completed or uncompleted tasks), deviation from the optimal path (number and order of steps deviating from the optimal solution), subjective metal workload (via BSMI) and electrodermal activity as a measure of objective workload. The focus of this paper lies on electrodermal activity and its relationship to measures of subjective workload. Hereby, skin conductance (mainly number of NS.SCRs and amplitude) during usage of both interfaces will be compared. A higher number of NS.SCRs and higher amplitude of these skin conductance responses would hereby imply a higher workload, meaning a possible threat to patient safety. For the analysis, changes in skin conductance above the threshold of 0.03 µS and speed changes of .000009 µS were counted as a SCR. Furthermore, the minimum gap between peaks was 700 ms. Two fellow "Human factors and Media" students conducted analyses of the other measures taken, which fall outside of the scope and focus of the present thesis.

2.6. Analysis

Analysis of the data was conducted with tonic EDA measures such as number of non-specific skin conductance responses and total amplitude per task as dependent variables. To account for the influence of time-on-task, time segments of 5 seconds were formed and chosen as basis for the analysis. This helped reducing over-dispersion and guaranteed that different lengths of tasks had no direct influence on the findings. A common technique for analyzing statistical data, ANOVA, was not applicable due to its assumptions. These are the independence, homogeneity and normality of the variances (which also assumes normality of values on the dependent variable) of the residuals (Eisenhart, 1947). That is, the variance of the residuals is independent of values of the predictor variables, is equal between groups and normally distributed. Yet, the normality of residuals is not given, as EDA data was heavily skewed (see Figure 6). Therefore, the more recent **generalized linear models** (**GLM**) were used, which are flexible generalizations of linear regression that allow for response variables that have error distribution models other than a normal distribution (McCullagh & Nelder, 1989). As a repeated measures design was chosen for this experiment, the repeated measures form of generalized linear models – generalized estimated equations (GEE) - was used for analyses, too. For analysis of the number of NS.SCRs a Poisson GEE analysis was chosen. Analysis of maximal amplitude was done by using a gamma GEE as was the analysis of skin conductance level.

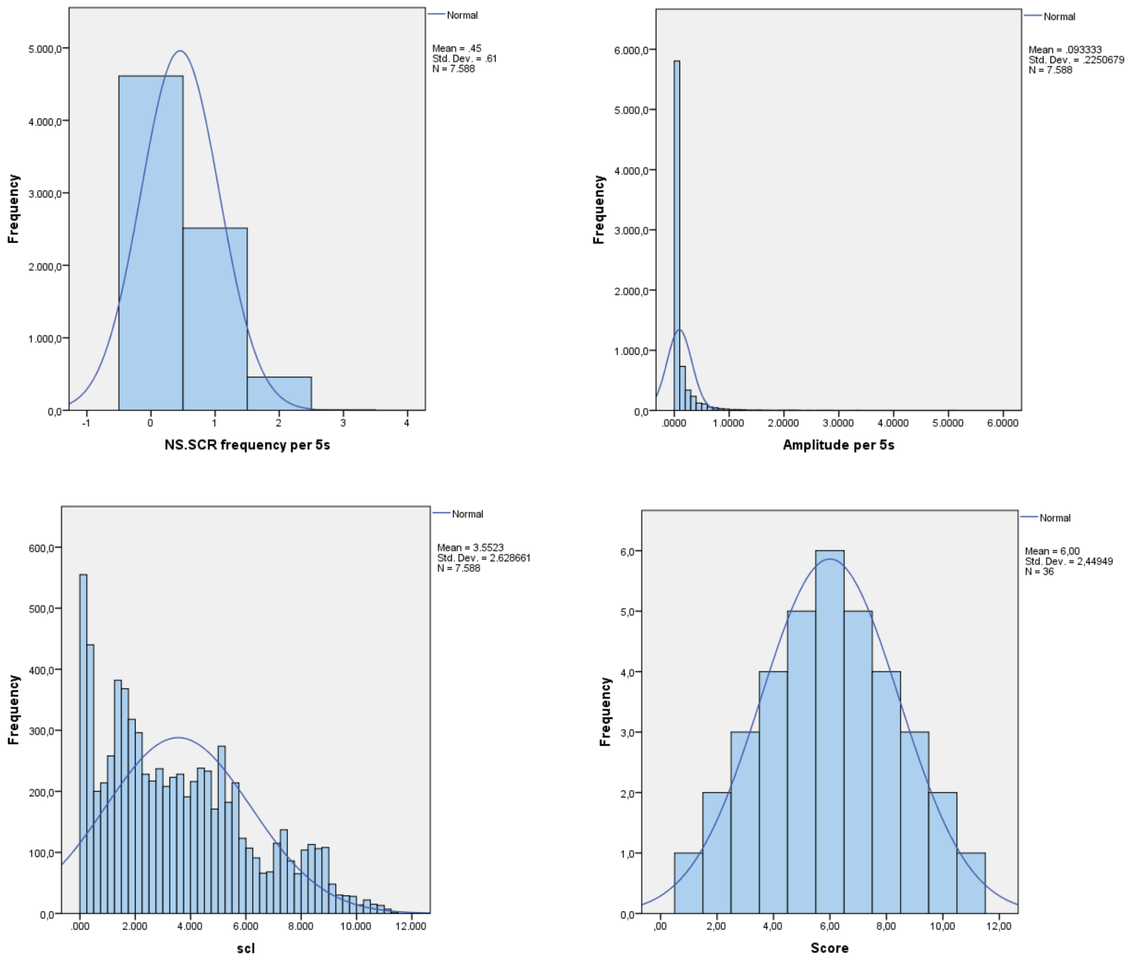


Figure 6. Distribution of the dependent variable in the experiment (top, bottom left) and an example of an approximate normal distribution (bottom left)

3. Results

In the following section, analyses of NS.SCRs, amplitudes of these and the SCL will be reported. Thereafter, these indicators of objective workload will be compared to a subjective measure (BSMI). However, one general finding, which will be reported now, was that standard deviations on all variables were quite high. This was due to the huge individual differences between people with regard to skin conductance. Yet, this is expected when conducting experiments related to electrodermal activity.

3.1. Differences in NS.SCRs

A Generalized estimated equations analysis of the data set with type of interface, task, measure and movement as predictors and number of non-specific skin conductance responses (henceforth NS.SCR frequency) as dependent variable showed no significant main effect for the type of interface (p= 0.398) or bodily movement (p= 0.111). These do therefore exert no significant influence on the number of NS.SCRs produced. However, other significant differences have been found with regard to NS.SCR frequency.

3.1.1. Differences between measures

When analyzing general differences between measures one observers that both measure one (p< 0.01) and two (p= 0.042) differ significantly from measure 3. Moreover, the difference between the first two measures is significant, too (p< 0.01) (see Figure 7). This was in such a way that measure 1 produced the highest mean frequency of NS.SCR per five second segment (M= .53), followed by measure 2 (M= .46) and measure 3 (M= .36). As one can see in figure 7, values were slightly lower, yet insignificantly so, for the Braun interface. When only assessing the difference between measures for the Braun pump one comes to the conclusion that only measure 1 differs significantly from measures 2 and 3 (p < 0.01). Measure 2 and three, however do not differ (p = 0.059).

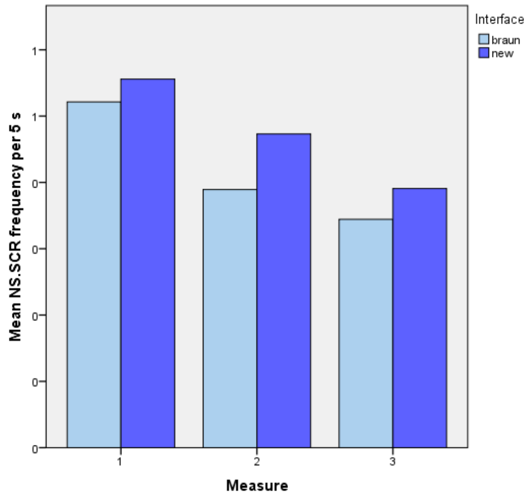


Figure 7. Mean NS.SCR frequency per measure for both interfaces

3.1.2. Differences between tasks

Although differences between tasks in general are not the main focus of this paper, analysis is still worthwhile, as different interfaces may be differently fit to handle certain tasks. Also, differences in EDA between tasks lets one deduce conclusions about the EDA measures themselves.

In general, task 3 (giving a bolus) was performed with the highest NS.SCR frequency (*M*= .46), with task 7 (retrieving information) (*M*= .45) being a close second. In contrast tasks 1 (starting the pump) and 4 (adjusting the infusion) exerted the lowest mean amount of NS.SCR frequency (*M*= .40 respectively). These two tasks differed significantly from the task of giving a bolus in the NS.SCR frequency they produced ($p < 0.01$). Differences between all other tasks remained insignificant.

For measure 1 task 1 (*M*= .42) differed significantly from all other tasks but task 4. Task 4, however, only differed significantly from task 3 ($p = 0.021$), task 7 ($p = 0.049$) and task 8 (giving an automatic bolus) ($p = 0.32$).

During the second measure tasks 8 (*M*= .35) and 10 (shutting the pump down) (*M*= .33) (having the lowest scores) both differed significantly from task 2 ($p < 0.05$) and task 10 differed significantly from task 3, too ($p < 0.05$).

For the third measure no tasks differed significantly from each other. A representation of NS.SCR frequency per task can be found in Figure 8. These values were averaged out over the three measures. Here, one directly sees the stark contrast between the interfaces on task 1. Yet, the difference between the interfaces on task 1 was insignificant across all measures.

Possible reasons for the difference will be given in the discussion.

To sum up, giving a bolus and searching for/extracting information produced the highest frequency of NS.SCRs, while starting the pump, adjusting the infusion and shutting the pump down produced the lowest frequencies. However, differences between tasks become insignificant with sufficient practice.

Moreover, the insignificance of type of interface on NS.SCR frequency remained stable throughout all separate measures.

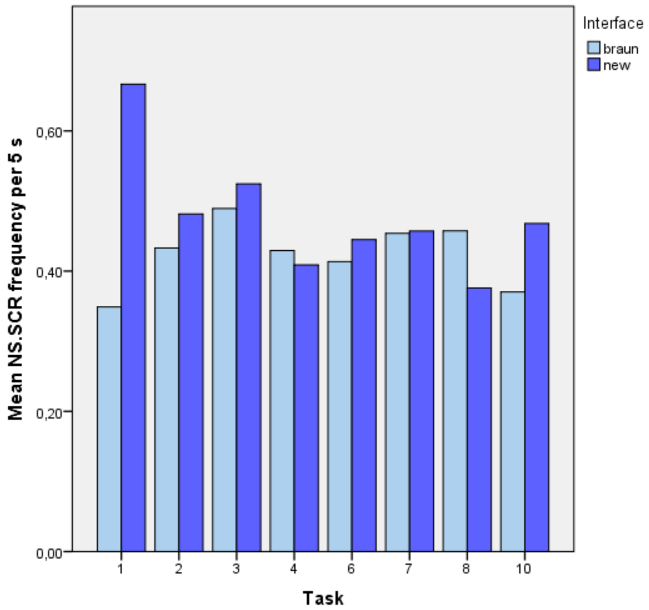


Figure 8. Representation of mean NS.SCR frequency per 5 seconds for both interfaces (averaged out over the three measures)

3.2. Differences in amplitude

Analysis concerning amplitude per 5 second segment was conducted with task, type of interface, measure and movement as independent and amplitude as dependent variable.

Regarding amplitude no significant difference between the two interfaces was found. Arm movement, however exerted significant influence ($p < .01$) in such a way that the fewer the movement, the smaller the amplitude. When analyzing the effect of measure on amplitude of the SCRs one observes that both the first and second measure differed significantly from the third ($p < 0.01$). There is, however, no significant difference between measure 1 and 2 with respect to amplitude although a slightly higher mean amplitude per 5 second segment was observed for measure 1 (*M*= .226/.223 µS respectively). This pattern was evident for both interfaces.

3.2.1. Differences between tasks

With regard to amplitude of SCRs the task which scored highest was task 10 (shutting the pump down) (*M*= .24 µS), followed by task 4 (adjusting the infusion) (*M*= .22 µS). The lowest amplitude of SCRs was found for tasks 8 (giving an automatic bolus) (*M*= .19 µS) and task 6 (*M*= .16 µS). Amplitude of SCRs for all other tasks per 5 second segment lay between a value of .2 and .21. None of the differences in amplitude were significant, however. This stayed true when only comparing tasks for the Braun Interface. Yet, when comparing amplitudes between tasks for the new interface one sees that task 6 (adjusting an infusion, with the lowest mean score of .16 for this interface) differs significantly from all other tasks except tasks 7 (extracting information) and 8 (giving an automatic bolus).

3.3. Differences in SCL

Similar to the other parts of the analysis, initial comparisons of Skin conductance level were conducted with task, type of interface, measure and arm movement as predictors and SCL as outcome variable. One of the first general findings was that arm movement exerted no significance at all (p = .919). Again, the difference between the two types of interfaces was insignificant as well (p = .932).

3.3.1. Differences between measures

A general finding in regard to measures is that both the first and the second ($p < .05$) and the second and third measure ($p < .01$) differ significantly from each other. Hereby, mean SCL was highest for measure one (*M*= 4.07 µS), followed by measure two (*M*= 3.4 µS) and three (*M*= 2.9 µS). When looking at SCL per type of interface one observes the same pattern for the Braun interface. However, measures two and three did not differ significantly for the new interface (p = .116).

3.3.2. Differences between tasks

With respect to SCL, task 2 (adjusting the infusion) exerted the lowest mean SCL (*M*= 3.23 µS), while task 1 (starting the pump) produced the highest SCL scores (*M*= 3.63 µS). Pairwise comparison revealed that task 1 differed significantly from all tasks but 6, 7 and 8. All of those differences were significant at the .05 level with the exception of the difference to task 2 ($p < .01$). In contrast, task 2 differed significantly from all other tasks but tasks 4 and 10, which exerted quite low mean SCLs as well (*M*= 3.3 µS respectively).

This pattern was quite different when only measure 1 was regarded. Here, the tasks producing the highest mean SCLs were tasks 8 and 10 (*M*= 4.2 µS respectively), followed by tasks 7 (*M*= 4.1 µS) and 6 (*M*= 4.0 µS). Tasks 8 and 10 differed significantly from all tasks but the tasks mentioned just now. Task 6 exerted a significant difference to tasks 2, 3 and 4 (*M*= 3.68/3.94/3.98 µS) whereas task 7 only showed significance to the task with the lowest mean SCL, namely task 2 (*M*= 3.68 µS).

During the second measure tasks 1 and 6 produced the highest mean SCL scores (*M*= 3.6 µS respectively) while tasks 8 (*M*= 3.12 µS) and 10 (*M*= 3.06 µS) showed the lowest scores. Tasks 8 and 10 hereby differed significantly from all other tasks but 6 and themselves. The insignificant difference to task 6 is due to the low amount of measure point at this task (the task was on average completed within the second lowest amount of time; task 10 being the shortest task on average). During measure 3 the same pattern could be observed.

3.4. Differences between user groups

When comparing nurses of the nursing department to nurses of the intensive care unit with regard to EDA measures some differences were observed. While the SCL ($p< .01$) and NS.SCR frequency ($p< .01$) differed significantly between all three measures for nurses of the nursing department, this was not the case for nurses employed at the intensive care unit. For this population no significant difference between measure two and three could be found on both variables. In contrast to this, no significant difference in amplitude could be found between measure 1 and 2, and measure 2 and 3 for employees of the nursing department.

3.5. Correlation between objective and subjective measures of workload

For the comparison of objective and subjective measures of workload a Pearson correlation between the NS.SCRs/amplitude per 5 seconds and the BSMI rating of the tasks was established for each participant separately.

3.5.1. Correlation between NS.SCRs per 5 seconds and BSMI

The correlation between NS.SCRs per 5 seconds and BSMI scores remained largely insignificant. Significant correlations were only found for 4 subjects (subjects 13, 61, 65, 74). Except for participant 61, all of these correlations were significant at the 0.01 level. The correlation for subject 61 was significant at the 0.05 level. The significant correlations themselves exerted values ranging from 0.357 to 0.491. Correlations for all other participants ranged from -0.277 to 0.243 (see Figure 9). When analyzing NS.SCRs per 5 second segment with task, type of interface and repetition as predictors for these subjects, the difference between the interfaces becomes significant at the 0.05 level with a mean of 0.31 NS.SCRs per segment for the new interface and a mean of 0.41 NS.SCRs per segment for the Braun Interface.

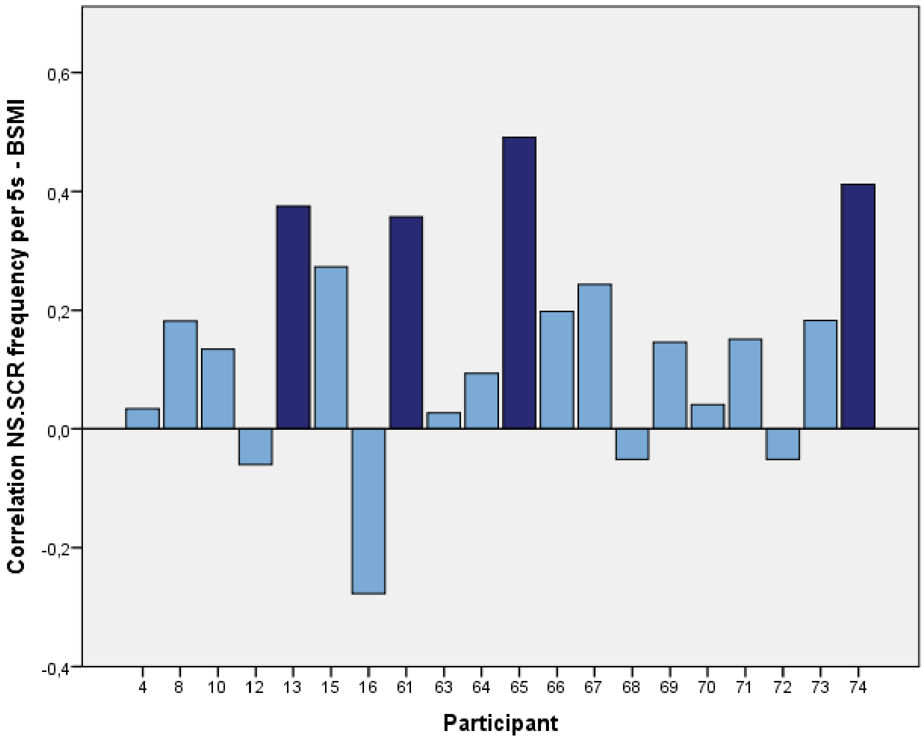


Figure 9. Pearson correlations between NS.SCR frequency per 5 seconds and BSMI scores per participant (significant correlations are colored dark blue)

3.5.2. Correlation between amplitude per 5 seconds and BSMI

Similar to the correlation between NS.SCRs per 5 seconds and BSMI scores, the correlation between amplitude per 5 seconds and BSMI remained predominantly insignificant; the only exception being the four subjects who also showed significant correlations between NS.SCRs per 5 seconds and BSMI scores. Significant correlations ranged from 0.315 to 0.501. All other correlations ranged from -0.175 to 0.229 (see Figure 10). Here, however, the amplitude of NS.SCRs per 5 second segment remained insignificant between interfaces.

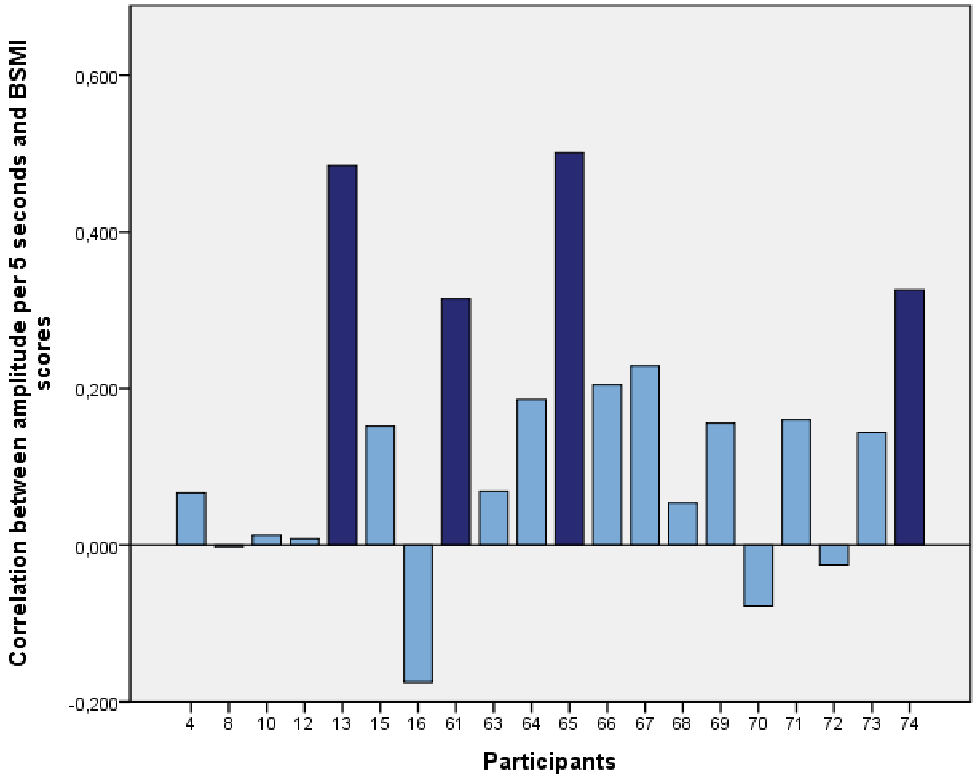


Figure 10. Correlations between amplitude per 5 seconds and BSMI scores (significant correlations are colored dark blue)

3.6. EDA measures and errors

When conducting Poisson GEE analysis with outcome as dependent variable and NS.SCR frequency, amplitude and SCL as predictors, it becomes apparent that none of these predictors exerts a significant influence on the outcome (correct or erroneous). The only variable which comes close to significance is SCL with p= .071. Distribution of NS.SCR frequency, amplitude and SCL values per outcome condition is very similar as can be seen in figures 11-13.

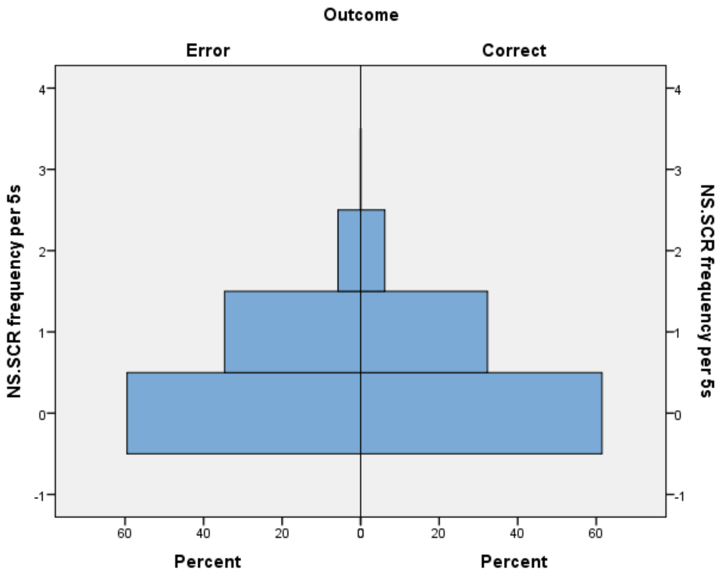


Figure 11. Histograms of NS.SCR frequency per 5 seconds in percentages for erroneous and correct trials

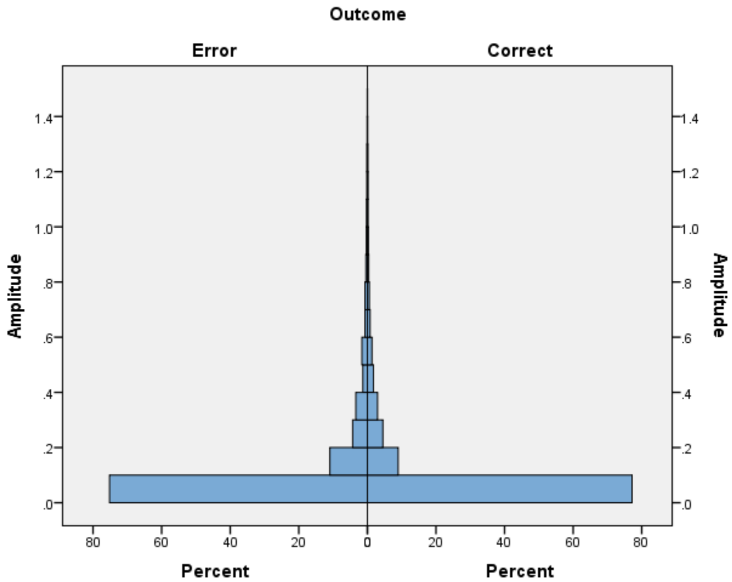


Figure 12. Histograms of amplitudes per 5 seconds (measured in µS) in percentages for erroneous and correct trials

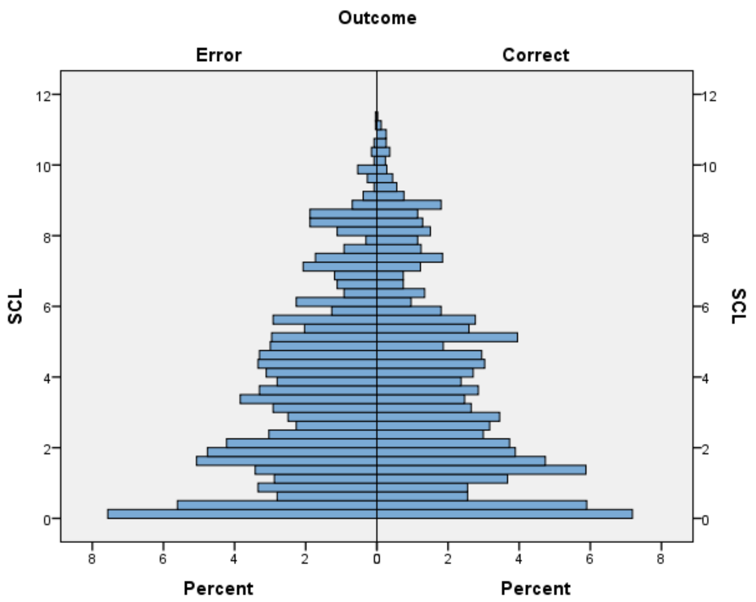


Figure 13. Histograms of SCL (measured in µS) in percentages for erroneous and correct trials

This shows that one cannot distinguish between erroneous and correct trials in this experiment based on the EDA parameters which were analyzed.

3.7. Analysis with more conservative/lenient definitions of NS.SCRs

Finally, analysis was conducted with more conservative or lenient definitions of what constitutes a NS.SCR. Therefore, either the threshold of change in skin conductance (originally 0.03 µS) was made more lenient (new threshold of 0.01 µS) or conservative (new threshold of 0.05 µS) or the criterion for speed changes (originally 0.000009 µS) was either made more lenient (new threshold of 0.0000008 µS) or conservative (new threshold of 0.00007 µS).

3.7.1. Applying new criterions for NS.SCRs regarding change in skin conductance

Applying a more lenient threshold for what constitutes a NS.SCR with regard to change in skin conductance (0.01 µS) produced roughly the same results as applying the original threshold. One difference was that measures two and three did not differ significantly anymore and that task 10 now exerted the least amount of NS.SCRs (*M*= 0.58 NS.SCRs per 5 second segment). Task 7 now produced the highest amount (*M*= 0.68 NS.SCRs per 5 second segment).

When applying a more conservative criterion on what constitutes a NS.SCR (0.05 µS) all measures became significantly different from each other again in relation to the NS.SCR frequency they produced. Similarly to analysis using the original criterion, task 3 produced the highest number of NS.SCRs (*M*= 0.38 NS.SCRs per 5 second segment), followed by task 7 (*M*= 0.36 NS.SCRs per 5 second segment). Task 4 in contrast revealed the lowest amount of NS.SCRs (*M*= 0.31 NS.SCRs per 5 second segment).

When applying the conservative criterion of change in skin conductance (0.05 µS) to the analysis of amplitude of SCRs per 5 second segment no striking differences to the original analysis were found.

3.7.2. Changing the criterion for NS.SCRs regarding speed changes

When applying a more conservative threshold of what constitutes a NS.SCR with respect to speed changes (0.00007 µS in contrast to 0.000009 µS) one observes no difference to the original analysis concerning NS.SCR frequency. A more lenient threshold (0.0000008 µS) led to the exact same results as well. Application of a more conservative criterion of what constitutes a NS.SCR with respect to speed changes (0.00007 µS in contrast to 0.000009 µS) to the analysis of amplitude underlined the pattern of measure 1 and 2 differing from measure 3 (*p*= 0.015/0.01 respectively), but not from each other. In contrast to the original analysis, however, measure 2 (*M*= 0.31 µS) exhibited a higher amplitude per 5 second segment than measure 1 (*M*= 0.3 µS). Tasks which exhibited the highest amplitude under this criterion were task 7 (*M*= 0.35 µS) and task 4 (*M*= 0.32 µS) while task 6 (*M*= 0.26 µS) produced the lowest mean amplitude.

When considering analysis with a more lenient criterion of what constitutes a NS.SCR with regard to speed changes (0.0000008 µS) the pattern of significance between measures remained intact (measures 1 and 2 differing significantly from measure 3 ($p < 0.01$ respectively), but not from each other), with measure 1 producing the highest mean amplitudes (*M*= 0.23 µS), again. Tasks 7 (*M*= 0.24 µS) and 4 (*M*= 0.22) exhibited the highest amplitudes with tasks 2 and 8 (*M*= 0.19 µS respectively) scoring lowest.

4. Discussion

During the course of the experiment a lot of interesting findings have been made; some of which are known from existing in literature, while others were quite surprising.
First of all, EDA measures were successful in that they produced typical, natural values which showed no (or at least little) influence from artifacts due to for example bodily movements. Thus, some common phenomena with regard to EDA measures could be established. Before we tend to these, however, the differences between the interfaces will be discussed.

4.1. Differences between the interfaces

No differences could be found between the two tested interfaces in regard to NS.SCR frequency as well as amplitude and SCL. Although this might surprising given the fact that the BSMI ratings for the new interface were found to be significantly lower than the BSMI scores for the Braun interface, one has to take into account that BSMI scores in general were quite low (Braun: M=21.17, Prototype: M=13.11). This is an indication that subjects in general were hardly aroused. However, not finding significantly higher values for the new interface on EDA parameters is a testimony to the quality of the interface itself. That is, due to the novelty of the prototype interface recognizably higher EDA values might also have been reasonable outcome. Novel features often induce higher electrodermal responses in the form of higher orienting responses. The prototype interface included many novelties for infusion pumps (such as the lock-button). Higher values on EDA parameters would thus have been perfectly fine.

Therefore, it is concluded that the prototype interface's design on basis of human factors guidelines is a success, as it seems to be well structured and intuitive to use. This intuitiveness is proven by the consistently lower error-rate and lower BSMI scores in comparison to the Braun interface, which has had a much longer design cycle and has been improved throughout the years repeatedly (the first Braun Perfusor pump was introduced in 1951).

Yet, as was mentioned before, subjective measures of workload do not always converge with objective measures such as EDA (e.g. Collet et al., 2003). The low BSMI values indicate that the overall arousal the tasks produced was quite low (although error rates might indicate otherwise, being 28.2% in general). This may have led to little opportunity for physiological measures to really distinguish between interfaces. This view is supported by the analysis of NS.SCRs. While 1-5 NS.SCRs per minute are common during rest, over 20 NS.SCRs signify a high arousal situation (Braithwaite et al., 2013). For the first measure mean NS.SCR frequency per minute was 6.36, which is nowhere near a high arousal situation. During the course of measure 2 and 3 this mean declined to 5.52 and 4.32 NS.SCRs per minute, respectively. These values are barely above or even at resting level. This may indicate that usage of both infusion pump interfaces in the current experimenter controlled scenarios is barely demanding. Error rates indicate that there is still a lot to learn until usage is nearly flawless, as ¼ of the tasks were still executed erroneously during measure 3.

However, it might be the case that participants were not really engaged in the tasks and were indifferent to their performance, rather than the tasks being undemanding. This may have been due to the low stakes they were tested under. In contrast to a real-life situation there was no risk of injuring or even killing a patient through misuse of the infusion pump. Importantly, during the experiment subjects were specifically told that not their performance, but rather the interfaces were tested. This situation is in stark contrast to a real life situation using infusion pumps, where the slightest inattention can have serious consequences for the patient. Thereby, the high error margin should be explained in the light of low BSMI and EDA scores. This is in accordance to the Yerkes-Dodson law (Yerkes & Dodson, 1908) mentioned in the introduction. This law claims that performance is low, when people are either scarcely or overly aroused. As participants' EDA values were generally at or around a resting level, it is no surprise that performance often was erroneous. A test design which uses a more cognitively demanding setting with higher-stakes (maybe a reward/punishment related to performance) might therefore be fitter to distinguish learnability and objective workload between interfaces. This is supported by Setz et al.'s (2012) notion that EDA cannot distinguish between routine tasks, but is rather fit to determine cognitively demanding planning and anticipation tasks. Such a more demanding test design could incorporate operating multiple infusion pumps/medical devices, background noise and working under time pressure. However, the comparability of both interfaces does not necessarily suffer from the low-stakes situation of the present experiment. In order to support this claim, the highest-scoring third of subjects with regard to SCL was compared to the lowest scoring third. SCL was chosen as discerning variables it is seen as a good indicator of general arousal/ engagement. The highest scoring third exhibited mean SCL values of at least 4.5 µS, while the lowest third produced mean SCL values of less than 1.5 µS.

Analysis underlined that the pattern of NS.SCR frequency regarding differences between measures, tasks and interfaces is exactly the same for the highest and the lowest scoring third (or the rest of the sample for that matter). There are significant habituation effects throughout the three measures, yet no differences between the interfaces were found. Thus, it is concluded that the low-stakes situation had no influence on the general discriminability of both interfaces. Another finding was that subjects produced considerably higher scores on task 1 for the prototype interface (although the difference was insignificant still). This is suspected to be mainly due to the addition of a "(un)lock device" button, which many subjects did not directly discover, as it seemed to be a novelty for infusion pump interfaces. Subjects, who found it often tried to repeatedly sweep it (which is the ultimate intention of the designer, see Figure 12). This, however, was not possible due to technical limitations of the prototype. One had to directly press the button instead of sweeping over it. Sadly, the button was not very responsive as well (mostly due to the *Fujitsu* tablet). Therefore, this (insignificant) difference is regarded as due to an artifact together with the fact that the button was a novelty in infusion pump interface design. Therefore, the higher NS.SCR frequency and their accompanying amplitudes for the prototype interface on task 1 were not surprising. Further research with this prototype should incorporate the possibility to sweep over the "lock" button and make it generally more responsive.

Figure 14. The "(un)lock" button in the new interface

4.2. Habituation

One hypothesis, which was confirmed, was that scores for NS.SCR, amplitude and SCL should decline per measure. This was indeed the case for almost all different measures. Therefore, a clear habituation with the use of both interfaces could be established. As stated above the difference in NS.SCR frequency and their amplitudes between the first and last measure were employed as indicators for habituation. Differences in SCL were disregarded as indicators for habituation, as these indicate a more general level of arousal and are not task specific. Differences in mean amplitude between the first and last trial block have long been employed as a measure for habituation in EDA (e.g. Koriat, Averill, & Malmstrom, 1973).

Comparison between both interfaces showed little difference in habituation. Differences in NS.SCR frequency from the first to the last measure were 0.18 SCRs per 5 seconds for the Braun and 0.17 SCRs per 5 seconds for the prototype interface. Differences in amplitude habituation (measured as difference in mean amplitude per 5 second segment between measure 1 and 3) were only observable at the third decimal place with 0.0478 µS for the Braun and 0.0445 µS for the new interface. To establish differences in the course of habituation between both interfaces with regard to EDA, parameters per measure were evaluated.

Analysis revealed that mean values of all EDA parameters never differed significantly between interfaces throughout all three measures. It is therefore concluded that habituation in regard to the usage of both interfaces did not, either. This habituation in EDA parameters is in line with a diminishing error-frequency, even though error rates are still high (1/4 of all tasks are still executed erroneously during measure 3). Habituation of EDA parameters does not necessarily go along with a diminished error-frequency, though, and is in fact not always correlated to it at all (e.g. Mehler et al., 2012) . Habituation only implies that participants get used to the task they are conducting/ the product they are using, which often, but not always goes hand in hand with a reduced error frequency/ better performance, as we have seen in the introduction.

4.3. Objective and subjective workload

Another intriguing finding was that objective and subjective workload did rarely correlate. A significant correlation could only be established for four subjects. For these four subjects the new interface produced significantly less subjective as well as objective workload (as measured by EDA). Yet, the significant correlations ranged from approximately 0.3 to 0.5 on both tested parameters (NS.SCR frequency and amplitude). According to Crewson (2006), these are still low correlations, even though they are significant. These weak correlations thus indicate that objective and subjective workload may indeed be two fundamentally different concepts, which rely on fundamentally different mechanisms. On the other hand, the disparity between objective and subjective measures of workload could indicate that people are in fact poor at estimating their physiological arousal/workload.

One explanation for this, as mentioned in the introduction, could be that positive valence (satisfaction) of a product moderates subjective workload, while obviously being unable to moderate EDA scores. One indicator for this explanation is that after the experiment most participants (76%) preferred the prototype interface while only 12% preferred the Braun pump (the remaining 12% had no preference at all). This was also true for all the different functions of the interfaces. Post-experiment interviews revealed that this was mostly due to the clear structure/easy overview of the interface and the direct and easy use. These points are ease-of-use issues that do not generally point to a more difficult use of the reference interface but rather a more bothersome use. When asked to assign BSMI scores to different tasks on both interfaces, however, BSMI scores given to the reference interface were significantly higher than for the prototype interface ($p < 0.01$). Yet, the dissociation of subjective measures of workload and EDA is not unknown in literature, as was pointed out in the introduction. For example, Seitz et al. (2012) found that while subjective measures of workload could distinguish driving situations, EDA scores could not. It was therefore concluded that EDA is unfit to differentiate between routine tasks. This view is supported by our own findings as all tasks given were essentially routine tasks, which are performed on a day to day basis. Scores of the different EDA measures (SCL, Frequency of NS.SCRs and amplitude) scarcely differed significantly between tasks. With regard to NS.SCR frequency, task 3 produced the highest values, being significantly different from the lowest scoring tasks during measure one and two. This task was giving a bolus, which is the rarest of all tasks employed. Anecdotal evidence suggests that some nurses (especially nurses from the nursing department) do not routinely or rarely give boluses. This further reinforces Seitz et al.'s claim that routine tasks are hard to differentiate using EDA, non-routine tasks, however, are not. This view is supported by Wilson (2001), who noted that during his study on workload in pilots during flight, EDA could not differentiate between most tasks. `

Heart rate on the other hand showed greater sensitivity with regard to differences between these routine flight tasks. Therefore, heart rate might be a better indicator of workload in routine tasks and should be employed when novel or high workload conditions are not expected. However, in the scope of this paper, EDA remains a fitting measure, as two completely different interfaces were compared and differences between the tasks were only of secondary importance. Furthermore, influences of novelty were expected considering the originality of the prototype interface.

Yet, there is another explanation, why EDA measures generally could find so little difference between tasks/interfaces, although other measures (error rates, BSMI) did, namely stress. As pointed out in the introduction, different levels of workload can only be distinguished in a no-stress setting (Conwey, 2012). There were always at least 2 researchers present, one of which had the purpose of handling the task controller and one who would give the tasks and ask for an evaluation of each task/ hold the post-experiment interview. Therefore, the participants might have been put under psychosocial stress. Subjects might have gotten the feeling that they were observed (which they were by means of an experimenter and a camcorder) and evaluated (which they were not). Yet, participants were explicitly told, that not they, but rather the interfaces were tested. Therefore stress can be discharged as source of indiscriminability between different tasks/interfaces. This is supported by the low EDA scores in general, which make sociopsychological stress seems highly unlikely and a low engagement of the subjects much more probable.

4.4. Electrodermal activity and error rates

Considering differences between tasks one observes that tasks that elicited high numbers of NS.SCRs (task 3 and 7) were also two of the three most error-prone tasks. Even though the relationship between tasks and error rates was not significant this is still an interesting finding. While task 3 had an error rate of 42.7%, task 7 even produced an error-rate of 63.3%. This is in accordance to Taylor's findings (1964) mentioned above, who found that SCRs per kilometer in a car-driving task corresponded to errors made during these kilometers. This finding therefore seems to be applicable not only to the driving task, but also to other tasks involving technological devices. Thus, number of NS.SCRs may be used as a "quick and dirty" predictor of error-rate, where extensive testing of error-rates is unfeasible or too costly. However, the exact relationship between errors and EDA as an indicator of workload could not be established and further research should try to unveil the circumstances under which EDA and error-rate correlate, as this study could -in contrast to Taylor (1964)- not establish a significant relationship between both. Yet, this is not unprecedented in literature as for example Mehler et al. (2012) found that SCL rose significantly with each difficulty level of an auditory presentation–verbal response working memory task during driving, while driving performance measures did not provide incremental discrimination. The exact pattern of findings as in the present study was made by Shimomura et al. (2008), who found that the score on the card sort NASA Task Load Index (CSTLX) increased in correspondence to task performance, while NS.SCR frequency and their amplitudes did not show any significant effect of task difficulty.

4.5. Decrease in SCL

Lastly, analysis could establish a decline in SCL throughout the course of the experiment. This is quite the common finding in EDA, indicating natural, successful measures of our experiment.

Consistently, analysis showed the highest SCL for task 1 (starting the pump). This might on first sight be a surprise, considering that task one was generally found to be quite undemanding. However, literature (e.g. Boucsein, 2012) indicates that SCL decreases over time during tasks. Task 1 was both the first task in each sequence of tasks and was also often completed within a small amount of time, explaining the high SCL-scores. On the other hand, task 2 revealed the lowest SCL-scores. This may be explained by the fact that task 2 was by far the task taking the longest time to complete. While all other tasks taken together had a mean time-till-completion of 31.12 seconds, task 2 produced a time-on-task of 79.06; more than twice the amount of all other tasks taken together. Another indicator for this phenomenon is task 4, which scored second highest on time-till-completion (59.4s) as well as second lowest on SCL. Surprisingly, task 10 – having the lowest mean duration (11.7s) - produced the same low SCL. However, task 10 was always the last task given for each set of tasks. Therefore, both duration of a task and placement of a task in a perceived sequence of tasks may exert different influences on SCL as these findings indicate.

4.6. Limitations

Naturally, this study has some limitations regarding the generalizability of its findings, some of which have been mentioned before, some of which have not been attended to yet. Firstly, in contrast to real-life situations there was no real time pressure or risk of injury. Subjects were urged to conduct the tasks as complete and fast as possible, but of course this is no comparison to a situation where a patient lies before them, urgently needing medication. Another limitation related to EDA measures are the differing task times. Obviously, if one does not match the time periods one creates a situation where more or fewer peaks can occur in certain conditions, which has more to do with unequal time periods not allowing for matched levels to be quantified rather than being due to any real differences in the frequency of NS-SCRs.

Yet, because of the multi-facetted approach of this experiment with time-on-task being a dependent variable itself, this was unfeasible. Another limitation was that participants were allowed to talk during the experiment, which might have had an effect on EDA measures. For example Peters (1974) noted that while electrodermal changes appear mainly during mental tasks, EDA was highest when test participants spoke during his observation of 11 female phonotypists. In our experiment we did not prohibit participants from speaking, which they frequently did ("This is the right thing to do, isn't it?", "Now I'm completely lost"). However, most other sources for artifacts were ruled out as test subjects were seated, and tests were taken in a silent room at constant temperature and it is therefore concluded that the influence of artifacts on our results is marginal. Movement of the arm onto which the Q-sensor was mounted was recorded and accounted for during the analysis.

5. Conclusion

All in all, the comparison between both tested interfaces found no significant difference in EDA measures. Yet, subjective workload as well as task performance was smaller/better for the new interface. Therefore, satisfaction and effectiveness measures clearly indicate a better usability of the new interface. The only true efficiency measure (effort measured by means of EDA), however, showed no difference between both interfaces. Besides this note, the new Interface does come out as the more usable product, even when objective measures of workload do not underline this finding. Yet, it speaks for the quality of the new interface that no differences between the interfaces could be found. Considering the short design-cycle of the prototype, worse results regarding EDA parameters would also have been perfectly reasonable. Therefore, this experiment paints a bright picture for the future of the further design-cycle and the interface itself. It further reinforces the need to develop technological devices with human factors guidelines in mind. All in all, the prototype interface produced fewer errors and was subjectively less loading than the reference one. A new test-design, which specifically focusses on EDA might be able to find differences in objective workload, as well. Such a test-design would feature fixed times-on-task and a more challenging set of tasks (including multiple interfaces). Yet, analysis showed that discriminability of the interfaces was independent of actual physiological arousal and it is thus concluded that the results are valid and reliable.

To sum up, involving human factors principles into interface design has been a full success. Two of the three dimensions of usability (effectiveness, satisfaction) already show better results for the new prototype interface. EDA as a measure of objective workload and thereby efficiency could not establish such a clear superiority of the prototype. This, however, would have been surprising given the novelty and the number of changes (in comparison to established devices) incorporated in the prototype. Thus, insignificance between both interfaces is quite the positive outcome and hints at the potential of the new interface design. This result, taken together with the findings of my fellow students, underlines that designing interfaces with human factors guidelines in mind stands a good chance of reducing the shocking amount of harmful incidents related to infusion pumps.

References

Association for the Advancement of Medical Instrumentation (2009). *Human factors engineering –Design of medical devices.* Retrieved January 04, 2014 from http://marketplace.aami.org/eseries/scriptcontent/docs/Preview%20Files/HE750910_preview.pdf

Baldauf, D., Burgarda, E., Wittmann, M.(2009). Time perception as a workload measure in simulated car driving. *Applied Ergonomics*, 40(5):929-935, 2009

Bogner, M.S. (1994). Human Error in Medical devices: Lack of feedback. *FDA User Reporting Bulletin, 14*, 1-8.

Boucsein, W. (2012) *Electrodermal activity.* New York: Plenum Press. Pp. 141-142, 223, 261-262,

Braithwaite, J.J., Watson, D.G., Jones, R., Rowe, M. (2013) A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments, *Technical Report: Selective Attention & Awareness Laboratory (SAAL),* Behavioural Brain Sciences Centre, University of Birmingham, UK

Bull, R. H. C., & Gale, A. (1973). The reliability of and interrelationships between various measures of electrodermal activity. *Journal of Experimental Research in Personality*, 6, 300–306.

Burbank, D. P., Webster, J. G. (1978). Reducing skin potential motion artifact by skin abrasion. *Medical & Biological Engineering & Computing*, 16, 31–38.

Cárdenas, D., Perales, J. C., Chirosa, L. J., Conde-González, J., Aguilar-Martínez, D., Araya, S. (2013).The effect of mental workload on the intensity and emotional dynamics of perceived exertion. *Anales de Psicología*, 29

Collet, C., Petit, C., Champely, S., Dittmar, A. (2003) Assessing workload through physiological measurements in bus drivers using an automated system during docking. *Human Factors*, 45 (2003), pp. 539–548

Conway, D. (2012) Stress and cognitive load. *NICTA Summer Scholarship Report*

2011/2012.

Crewson, P. (2006). *Applied statistics handbook (AcaStat Software)*. P. 85

Dawson, M. E., Schell, A. M. (2002). What does electrodermal activity tell us about prognosis in the schizophrenia spectrum? *Schizophrenia Research*, 54, 87-93.

Dawson, M. E., Schell, A. M., Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (3rd ed., pp. 159-181). New York: Cambridge University Press.

De Waard, D. (1996). The measurement of drivers' mental workload. *Ph.D. thesis*, University of Groningen, Traffic Research Centre, Haren, The Netherlands.

De Waard, D., Brookhuis, K.A. (1993). The use of psychophysiology to assess driver status. *Ergonomics* 36 (9), 1099-1110

Drachen, A. Nacke, L.E., Yannakakis, G., Pedersen, A.L. (2010) Correlation between heart rate, electrodermal activity and player experience in first-person shooter games, *Proceedings of the 5th ACM SIGGRAPH Symposium on Video Games*, p.49-54, July 28-29, 2010, Los Angeles, California

Duffy, E. (1972). Activation. In N. S. Greenfield & R. A. Sternbach (Eds.), *Handbook of psychophysiology* (pp. 577–622). New York: Holt, Rinehart, & Winston.

Eisenhart, C. (1947). The Assumptions Underlying the Analysis of Variance. *Biometrics*. Vol. 3, 1-21.

Food and Drug Administration (2011). *Draft guidance for industry and food and drug administration staff - Applying human factors and usability engineering to optimize medical device design*. Retrieved June 19, 2013 from http://www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm259748.htm

Food and Drug Administration (2010). *Infusion pump improvement initiative*. Retrieved June 27, 2013 from http://www.fda.gov/downloads/MedicalDevices/ProductsandMedicalProcedures/GeneralHospitalDevicesandSupplies/InfusionPumps/UCM206189.pdf

Gould, J.D., Lewis, C. (1985) Designing for Usability: Key Principles and What Designers Think, *Communications of the ACM* (28:3), pp. 300-311.

Harris, M. D. (1943). Habituatory response decrement in the intact organism. *Psychological Bulletin*, 40, 385–422.

Hart, S. G., Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock and N. Meshkati (Eds.) *Human Mental Workload*. Amsterdam: North Holland Press.

Hornbæk, K. (2006) Current Practice in Measuring Usability: Challenges to Usability Studies and Research, *International Journal of Human-Computer Studies*, 64, 2 (2006), 79-102.

Humphrey, G. (1933). *The nature of learning*. New York: Harcourt Brace.

Husch, M., Sullivan, C., Rooney, D., Barnard, C., Fotis, M., Clarke, J., Noskin, G. (2005). Insights From the Sharp End of Intravenous Medication Errors: Implications for Infusion Pump Technology," *Quality & Safety in Health Care.*; 14.

Hygge, S., & Hugdahl, K. (1985). Skin conductance recordings and the NaCl concentration of the electrolyte. *Psychophysiology*, 22, 365–367.

Hyman, W.A. (1994). Errors in the use of medical equipment. *In Bogner (Ed.) Human Error in Medicine*, 327-347.

Izso, L., Lang, E. (2000). Heart period variability as mental effort monitor in human computer interaction. *Behaviour and Information Technology* 19 (4), 297–306 .

ISO (1998). *Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability.* Retrieved July 10, 2013 from http://www.it.uu.se/edu/course/homepage/acsd/vt09/ISO9241part11.pdf

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Koriat, A., Averill, J. R., Malmstrom, E. J. (1973). Individual differences in habituation: Some methodological and conceptual issues. *Journal of Research in Personality*, 7, 88–101.

Lader, M. H., & Wing, L. (1964). Habituation of the psycho-galvanic reflex in patients with Anxiety states and in normal subjects. *Journal of Neurology, Neurosurgery, and Psychiatry*, 27, 210–218.

Lader, M. H., Wing, L. (1966). *Physiological measures, sedative drugs, and morbid anxiety*. London: Oxford University Press.

Laufer, L., Németh, B. (2008). Predicting User Action from Skin Conductance. *Proceedings of the 13th conference on Intelligent user interfaces*, ACM Press, 2008, 357–360.

Lazarus, R.S., Folkman, S. (1984). *Stress appraisal and coping*. New York: Springer; pp.19

Lykken, D. T. (1981). *A tremor in the blood: Uses and abuses of the lie detector*. New York: McGraw-Hill.

McCullagh, P, Nelder, J.A. (1989). *Generalized Linear Models, Second Edition*. Boca Raton: Chapman and Hall/CRC

Martin, I., Rust, J. (1976). Habituation and the structure of the electrodermal system. *Psychophysiology*, 13, 554–562.

Mehler, B., Reimer, B., Coughlin, J. F. (2012). Sensitivity of physiological measures for detecting systematic variations in cognitive demand from a working memory task: an on-road study across three age groups. *Human Factors*, 54(3), 396-412.

Obradovich, J. H., Woods, D. D. (1996). Users as designers: How people cope with poor HCI design in computer-based medical devices. *Human Factors*, 38(4), 574-592.

Peters, T. (1974). Mentale Beanspruchung von B€uroangestellten im Schreibdienst und bei Vorzimmert€atigkeit. *Zentralblatt fuer Arbeitsmedizin und Arbeitsschutz*, 24, 197–207.

Poh, N.C. Swenson, R.W. Picard (2010). A Wearable Sensor for Unobtrusive, Long-term Assessment of Electrodermal Activity. *IEEE Transactions on Biomedical Engineering*, vol.57, no.5,pp.1243-1252, May 2010.

Russell, J. (1980). "A circumplex model of affect". *Journal of Personality and social Psychology* 39: 1161–1178.

Sawyer, D. (1996). Do it by design. An introduction to human factors in medical devices. *Food and Drug Administration.* Retrieved October 3, 2013 from http://www.fda.gov/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm094957.htm

Schmettow, M., Vos, W., & Schraagen, J. M. (2013). With how many users should you test a medical infusion pump? Sampling strategies for usability tests on high-risk systems. *Journal of Biomedical Informatics*, 46(4), 626-641.

Seitz, M., Daun, T. J., Zimmermann, A., Lienkamp, M. (2012). Measurement of Electrodermal Activity to Evaluate the Impact of Environmental Complexity on Driver Workload. *Proceedings of the FISITA 2012 World Automotive Congress*, Springer-Verlag

Selye, H. (1956). *The Stress of Life*. New York: McGraw-Hill.

Setz, C., Arnrich, B., Schumm, J., La Marca, R., Troster, G., Ehlert, U. (2010). Discriminating stress from cognitive load using a wearable EDA device. *IEEE Transactions on Information Technology in Biomedicine* 14(2), 410–417.

Shimomura, Y., Yoda, T., Sugiura, K., Horiguchi, A., Iwanaga, K., Katsuur, T. (2008). Use of frequency domain analysis of skin conductance for evaluation of mental workload. *Physiol Anthropol*., 27, pp. 173–177

Sokolov, E. N. (1963). *Perception and the conditioned reflex*. Oxford: Pergamon.

Taylor, A. (2010). Infusion Pump Design Deficiencies And Their Impact On Users and Caregivers. p. 4. Retrieved on 10.01.2014 from http://www.fda.gov/downloads/MedicalDevices/NewsEvents/WorkshopsConferences/UCM219687.pdf

Thorpe, W. M. (1969). *Learning and instinct in animals.* London: Methuen.

Van Dooren, M., de Vries, J.J.G., Janssen, J. H. (2012). Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology & Behavior*, Feb. 20120.

Wickens, C.D. (1984). "Processing resources in attention", in R. Parasuraman & D.R. Davies (Eds.), *Varieties of attention*, (pp. 63–102). New York: Academic Press.

Wickens, C. D., Lee, J. D., Liu, Y, Gordon Becker, S. E. (2004). *An Introduction to Human Factors Engineering*. Second ed,. Upper Saddle River, NJ: Pearson Prentice Hall, 2004. 185–193.

Wilson, G.F. (2001). An analysis of mental workload in pilots during flight using multiple psychophysiological measures. *Int. J. Aviat. Psychol.* 12, 3–18.

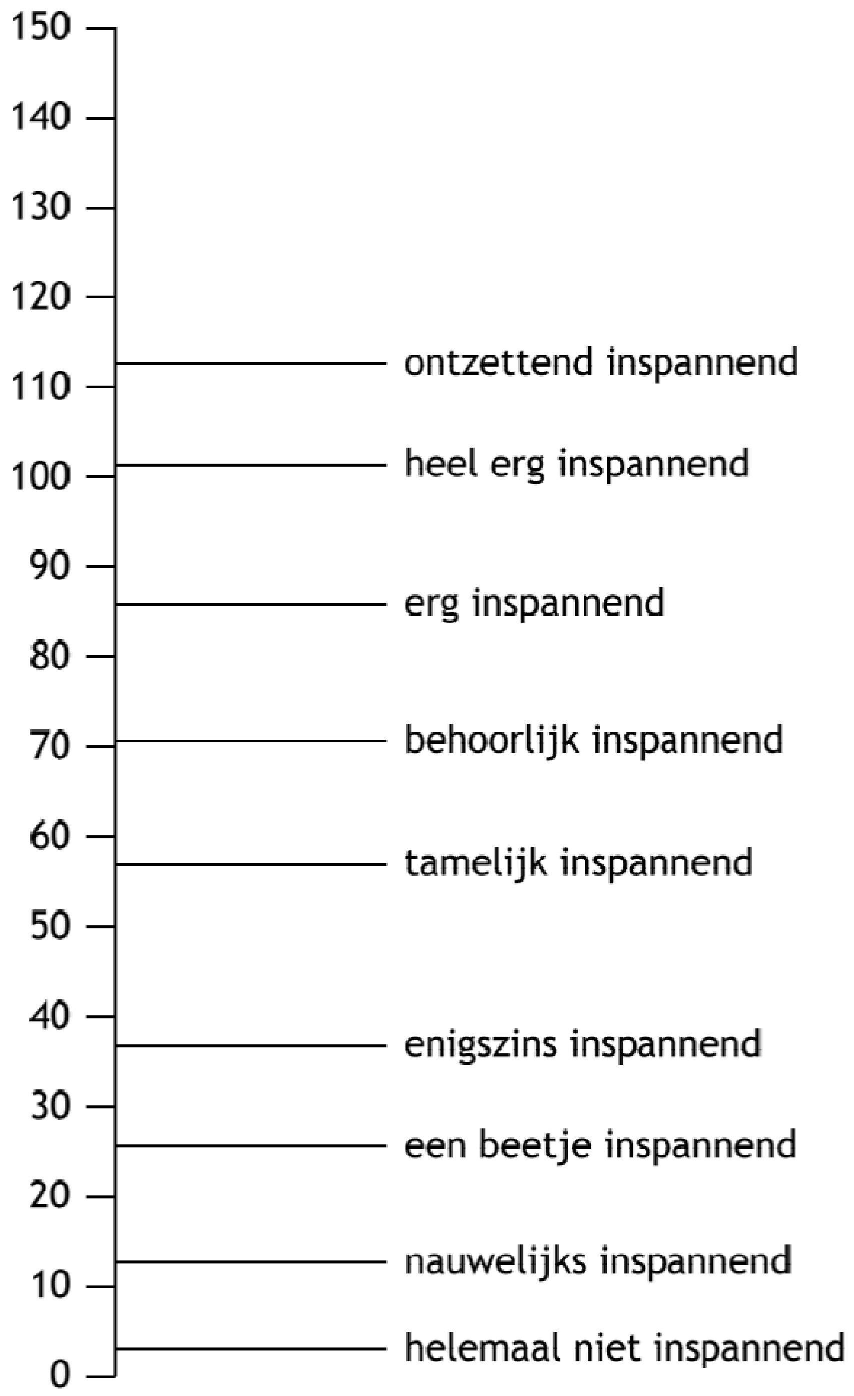
Yerkes, R.M., Dodson, J.D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*. **18**, pp. 459–482.

Zijlstra, F.R.H., Van Doorn, L. (1985). *The construction of a scale to measure perceived effort*. Delft, The Netherlands: Department of Philosophy and Social Sciences, Delft University of Technology.

APPENDIX A: BSMI

Belasting Schaal Mentale Inspanning

Instructie: Wilt u op de onderstaande lijn aangeven hoe inspannend u de taak vond die u net heeft uitgevoerd?



APPENDIX B: Tasks and scenarios for both user groups used in experiment

Taken IC, variatie 1

NB: Ga ervan uit dat de patiënten een gemiddeld gewicht van 75 kilo hebben

Taak 1)

Je wilt de spuitpomp gebruiken voor een infusie. De spuit is net geplaatst maar de pomp staat nog uit. Schakel de pomp in.

Taak 2)

Een patiënt van 61 jaar heeft zojuist een knieoperatie ondergaan maar bevind zich in een kritieke toestand en is daarom overgeplaatst naar de IC. Om pijn te voorkomen als gevolg van de operatie moet de pijn worden bestreden met Morfine (0.08 mg/ml).

Taak: Stel de juiste waarden in aan de hand van onderstaande gegevens en start hierna de infusie.

Medicatie:	Morfine (0.08 mg/ml)
Snelheid:	6,1 ml/uur
Tijdsduur:	3 uur en 30 minuten

Taak 3)

Je krijgt nu een taak mondeling aangeboden

"Snel, een patient heeft een bolus propofol nodig. Geef een (manuele) bolus van 5 ml".

Taak 4)

Een patiënt van 27 jaar heeft een liesbreukoperatie ondergaan en vanwege zijn kritieke toestand is hij op de IC geplaatst. Omdat de patiënt erg onrustig is wordt hij gekalmeerd door middel van propofol (1%) met een snelheid van 79 ml/uur. Het blijkt dat de patiënt niet voldoende kalmeert waardoor de dosering propofol verhoogd zal moeten worden. Je wordt gevraagd om de dosering aan te passen.

Taak: Verhoog de dosering van Propofol door de pompstand aan te passen op basis van onderstaande gegevens.

Snelheid: 96,4 ml/uur
Tijd: Resterende tijd

Taak 6)

Een 78 jaar oude patiënt is geopereerd aan zijn heup. Omdat de patiënt een te laag hartminuutvolume heeft wordt er dopamine toegediend (4 mg/ml) met 2.6 ml/uur en wordt de patiënt naar de IC gebracht. Omdat het hartminuutvolume niet voldoende toeneemt moet de dosering van dopamine verhoogd worden.

Taak: Verhoog de dosering van Dopamine door de pompstand aan te passen op basis van onderstaande gegevens.

Snelheid: 3,2 ml/uur
Tijd: Resterende tijd

Taak 8)

Je krijgt nu een taak mondeling aangeboden

"Snel, een patient heeft een bolus Labetolol nodig. Geef een (manuele) bolus van 10 ml".

Taak 7)

De arts wil weten hoeveel Amiodarone al is toegediend sinds de patiënt op de IC ligt. Omdat je niet zelf de Amiodarone hebt gegeven weet je niet hoeveel er al is toegediend. Je wilt daarom weten hoeveel Amiodarone (18 mg/ml) een patiënt al is toegediend.

Taak: Zoek deze informatie op in het interface en benoem de juiste hoeveelheid.

Taak 10)

De kuur is afgelopen en de patiënte hoeft dit medicijn niet meer toegediend te krijgen.

Taak: Stop de infusie en schakel de pomp uit.

Taken IC, variatie 2

NB: Ga ervan uit dat de patiënten een gemiddeld gewicht van 75 kilo hebben

Taak 1)

Je wilt de spuitpomp gebruiken voor een infusie. De spuit is net geplaatst maar de pomp staat nog uit. Schakel de pomp in.

Taak 4)

Een 25 jarige patiënte heeft een schouderoperatie ondergaan maar verkeert in kritieke toestand. Om haar te kalmeren is propofol (1%) toegediend. De snelheid waarmee de propofol wordt toegediend bedraagt 70 ml/uur. Omdat blijkt dat de sedatie langzaam afneemt wordt besloten dat de dosering propofol verhoogd moet worden naar 85,4 ml/uur.

Taak: Verhoog de dosering Propofol door de pompstand aan te passen op basis van onderstaande gegevens.

Snelheid:	85,4 ml/uur
Tijdsduur:	Resterende tijd

Taak 8)

Je krijgt nu een taak mondeling aangeboden

“Snel, een patient heeft een bolus Perlinganit nodig. Geef een (manuele) bolus van 1 ml”.

Taak 7)

De arts wil weten hoeveel Morfine (0,08 mg/ml) er al is gegeven sinds de patiënt op de IC ligt. Je weet niet precies hoeveel er al is gegeven en wilt dit daarom opzoeken in het interface.

Taak: Zoek deze informatie op in het interface en benoem de juiste hoeveelheid.

Taak 2)

Een mannelijke patiënt van 56 jaar is zojuist binnengebracht op de IC, hij is net geopereerd aan zijn hernia maar verkeert in kritieke toestand. Je moet het middel Remifentanil (50 mcg/ml) toedienen via *de infuuspomp.*

Taak: Stel de juiste waarden in aan de hand van onderstaande gegevens en start daarna de infusie.

Medicatie: Remifentanil (50 mcg/ml)
Snelheid: 47,4 ml/uur
Tijdsduur: 4 uur en 15 minuten

<u>Taak 6)</u>

Een 31 jaar oude patiënt is geopereerd aan zijn hart vanwege hartritmestoornissen. Om de hartfrequentie na de operatie te verhogen wordt Isoprenaline (40 mcg/ml) toegediend met 9,8 ml/uur. Omdat de hartfrequentie niet voldoende stijgt moet de dosering Isoprenaline verhoogd worden.

Taak: Verhoog de huidige dosering van Isoprenaline door de pompstand aan te passen. Hou je hierbij aan onderstaande gegevens.

Snelheid: 19,6 ml/uur
Tijdsduur: Resterende tijd

<u>Taak 3)</u>

Je krijgt nu een taak mondeling aangeboden

"Snel, een patient heeft een bolus Arterenol nodig. Geef een (manuele) bolus van 2 ml".

<u>Taak 10)</u>

De kuur is afgelopen en de patiënt hoeft dit medicijn niet meer toegediend te krijgen.

Taak: Stop de infusie en schakel de pomp uit.

Taken IC, variatie 3

NB: Ga ervan uit dat de patiënten een gemiddeld gewicht van 75 kilo hebben

Taak 1)

Je wilt de spuitpomp gebruiken voor een infusie. De spuit is net geplaatst maar de pomp staat nog uit. Schakel de pomp in.

Taak 3)

Je krijgt nu een taak mondeling aangeboden

“Snel, een patient heeft een bolus Suprarenin nodig. Geef een (manuele) bolus van 0.7 ml”.

Taak 7)

De arts wil graag weten hoeveel Reminfentanil (50 mcg/ml) er al is toegediend sinds de patiënt op de IC ligt. Omdat je de precieze hoeveelheid niet weet wil je dit opzoeken in het interface.

Taak: Zoek deze informatie op in het interface en benoem de juiste hoeveelheid.

Taak 6)

Een patiënte van 45 jaar heeft een operatie aan haar enkelband ondergaan en ligt nu op de IC om bij te komen. De patiente moet echter een tweede operatie ondergaan, als voorbereiding op deze operatie krijgt zij Atracurium (3 mg/ml) met 7,9 ml/uur toegediend. Omdat blijkt dat de spieren niet genoeg ontspannen moet de dosering Atracurium verhoogd worden.

Taak: Verhoog de dosering van Atracurium door de pompstand aan te passen. Hou je hierbij aan onderstaande gegevens.

Snelheid:	15,8 ml/uur
Tijd:	Resterende tijd

Taak 2)

Een patiënt van 43 jaar heeft een hartklepoperatie ondergaan en ligt op de IC. Om te voorkomen dat de patiënt last krijgt van atriumfibrilleren moet je preventief Amiodarone (18 mg/ml) toedienen om het fibrilleren te voorkomen.

Taak: Stel de juiste waarden in aan de hand van onderstaande gegevens en start hierna de infusie.

Medicatie: Amiodarone (18 mg/ml)
Snelheid: 3,2 ml/uur
Tijdsduur: 24 uur en 0 minuten

<u>Taak 4)</u>

Een patiënt van 36 heeft net een maagoperatie ondergaan maar verkeert in kritieke toestand. Om de patiënt te kalmeren is er propofol (1%) toegediend met een snelheid van 69 ml/uur. Het blijkt dat de patiënt niet genoeg kalmeert, waardoor je de huidige pompstand moet verhogen.

Taak: Verhoog de dosering van Propofol door de pompstand aan te passen. Hou je hierbij aan onderstaande gegevens.

Snelheid: 84,2 ml/uur
Tijd: Resterende tijd

<u>Taak 3)</u>

Je krijgt nu een taak mondeling aangeboden

“Snel, een patient heeft een bolus Dobutamin nodig. Geef een (manuele) bolus van 1 ml”.

<u>Taak 10:</u>

De kuur is afgelopen en de patiënt hoeft dit medicijn niet meer toegediend te krijgen.

Taak: Stop de infusie en schakel de pomp uit.

Scenarios Verpleegafdeling, variatie 1

NB: Ga ervan uit dat de patiënten een gemiddelde gewicht van 75 kilo hebben

<u>Taak 1)</u>

Je wilt de spuitpomp gebruiken voor een aanstaande infusie. De spuit is net geplaatst maar de pomp staat nog uit. Schakel de pomp in.

<u>Taak 2)</u>

Een patiënte van 58 jaar heeft een operatie ondergaan en werd overgeplaatst naar je afdeling. Ter preventie van trombose moet de patiënte Heparine (400 IE/ml) toegediend krijgen.

Taak: Stel de juiste waarden in aan de hand van onderstaande gegevens en start hierna de infusie.

Medicatie:	Heparine (400 IE/ml)
Snelheid:	2,0 ml/uur
Tijdsduur:	4 uur

<u>Taak 3)</u>

Een patiënt van 42 jaar heeft een buikoperatie ondergaan en krijgt de dag daarna nog Morfine toegediend. De spuit met het medicijn Morfine (0,4 mg/ml) is geplaatst en de pomp loopt met een snelheid van 5 ml/uur. De patiënt klaagt over plotselijke heftige pijn en het wordt daarom besloten om een bolus Morfine te geven.

Taak: Geef een manuele bolus met een volume van 5 ml .

<u>Taak 4)</u>

Een patiënte van 73 jaar wordt palliatief gesedeerd door middel van Midazolam (1 mg/ml) met een snelheid van 2 ml/uur. Het blijkt dat de patiënte nog steeds last van pijn heeft. Het wordt daarom besloten om de dosering te verhogen.

Taak: Verhoog de dosering van Midazolam door de pompstand aan te passen op basis van onderstaande gegevens.

Snelheid:	3,5 ml/uur
Tijd:	Resterende tijd

Taak 6)

Een 60 jaar oude patiënte, bekend met diabetis mellitus type 1, krijgt Insuline (1 IE/ml) toegediend met een snelheid van 2,5 ml per uur. Na een controle blijkt dat de bloedsuikerwaarde nog steeds veel te hoog te is. Daarom wordt besloten om de dosering te verhogen.

Taak: Verhoog de dosering van Insuline door de pompstand aan te passen op basis van onderstaande gegevens.

Snelheid:	4,8 ml/uur
Tijd:	Resterende tijd

Taak 7)

Een wisseling van de dienst heeft plaats gevonden en je moet de taken van een collega overnemen. Je wilt weten hoeveel Morfine (0,4 mg/ml) een patiënt al is toegediend.

Taak: Zoek deze informatie op in het interface en benoem de juiste hoeveelheid.

Taak 8)

Om bloedstolling tegen te gaan krijgt een patiënte van 67 jaar Heparine (400 IE/ml) toegediend met een snelheid van 1,5 ml/uur. De controle van de INR waarde laat zien dat deze waarde niet goed is en dat de patiënte onmiddelijk een bolus Heparine nodig heeft.

Taak: Geef een automatische Bolus van 1 ml.

Taak 10)

De kuur is afgelopen en de patiënte hoeft dit medicijn niet meer toegediend te krijgen.

Taak: Stop de infusie en schakel de pomp uit.

Scenarios Verpleegafdeling, variatie 2

NB: Ga ervan uit dat de patiënten een gemiddelde gewicht van 75 kilo hebben

Taak 1)

Je wilt de spuitpomp gebruiken voor een infusie. De spuit is net geplaatst maar de pomp staat nog uit. Schakel de pomp in.

Taak 4)

Een 25 oude patiënt wordt na het ondergaan van een schouderoperatie overgeplaatst naar de gewone afdeling. De patiënt krijgt nog voortdurend Morfine (0,4 mg/ml) toegediend met een snelheid van 2,6 ml per uur. De pijn blijkt na een tijd lager te worden en het wordt daarom besloten om de snelheid van de infusie te verlagen.

Taak: Verlaag de dosering Morfine door de pompstand aan te passen op basis van onderstaande gegevens.

Snelheid:	1,5 ml/uur
Tijdsduur:	Resterende tijd

Taak 8)

Een 80 jaar oude patiënt krijgt Midazolam (1 mg/ml) toegediend met een snelheid van 2 ml/uur. Het blijkt dat de patiënt plotseling hevige pijn krijgt waardoor besloten wordt om een bolus te geven.

Taak: Geef een automatische bolus van 2,7 ml.

Taak 7)

Een wisseling van de dienst heeft plaats gevonden en je moet de taken van een collega overnemen. Je wilt weten hoeveel Midazolam (1 mg/ml) een patiënt al is toegediend.

Taak: Zoek deze informatie op in het interface en benoem de juiste hoeveelheid.

Taak 2)

Na een ambulante ingreep ligt een patiënte van 70 jaar nog een dag op de afdeling. De patiënte is bekend met diabetis mellitus type 1 en moet daarom Insuline toegediend krijgen.

Taak: Stel de juiste waarden in aan de hand van onderstaande gegevens en start daarna de infusie.

Medicatie: Insuline (1 IE/ml)
Snelheid: 2 ml/uur
Tijdsduur: 4 uur

Taak 6)

Een 30 jaar oude patiënt werd geopereerd vanwege een gecompliceerde elleboogfractuur. Hij is nu overgeplaatst naar je afdeling en krijgt Dipidolor (0,4 mg/ml) toegediend met een snelheid van 1,8 ml/uur. De patiënt klaagt voortdurend over pijn waarop besloten wordt om de dosering van Dipidolor te verhogen.

Taak: Verhoog de huidige dosering van Dipidolor door de pompstand aan te passen. Hou je hierbij aan onderstaande gegevens.

Snelheid: 2,5 ml/uur
Tijdsduur: Resterende tijd

Taak 3)

Na een knie operatie te zijn ondergaan en na een verblijf op de IC wordt een patiënte van 65 jaar overgeplaatst naar de gewone afdeling. Om pijn te bestrijden krijgt de patiënte Dipidolor (0,4 mg/ml) toegediend met een snelheid van 2,1 ml per uur. Opeens klaagt de patiënte over ondraaglijke pijn in haar knie. Voor een onmiddelijke pijn verlichting wordt een bolus Dipidolor toegediend.

Taak: Geef een manuele bolus van 5 ml.

Taak 10)

De kuur is afgelopen en de patiënt hoeft dit medicijn niet meer toegediend te krijgen.

Taak: Stop de infusie en schakel de pomp uit.

Scenarios Verpleegafdeling, variatie 3

NB: Ga ervan uit dat de patiënten een gemiddelde gewicht van 75 kilo hebben

Taak 1)

Je wilt de spuitpomp gebruiken voor een infusie. De spuit is net geplaatst maar de pomp staat nog uit. Schakel de pomp in.

Taak 3)

Een 83 jaar oud patiënte krijgt Midazolam (1 mg/ml) toegediend met een snelheid van 3 ml/uur. De patiënte klaagt over hevige pijn.

Taak: Geef een manuele bolus van 2,0 ml.

Taak 7)

Een wisseling van de dienst heeft plaats gevonden en je moet de taken van een collega overnemen. Je wilt weten hoeveel Heparine (400 IE/ml) een patiënt al is toegediend.

Taak: Zoek deze informatie op in het interface en benoem de juiste hoeveelheid.

Taak 6)

Een patiënte van 45 jaar krijgt Midazolam (1 mg/ml) toegediend met een snelheid van 2 ml/uur. De patiënte klaagt over voortdurende pijn. Daarom wordt besloten om de dosering van Midazolam te verhogen.

Taak: Verhoog de dosering van Midazolam door de pompstand aan te passen. Hou je hierbij aan onderstaande gegevens.

Snelheid:	3,2 ml/uur
Tijd:	Resterende tijd

Taak 2)

Een patiënt van 58 jaar heeft een longoperatie ondergaan en is overgebracht naar de verpleegafdeling. Om de pijn te bestrijden zal de patiënt moeten worden geïnfuseerd met morfine (0.4 mg/ml).

Taak: Stel de juiste waarden in aan de hand van onderstaande gegevens en start hierna de infusie.

Medicatie: Morfine (0.4mg/ml)
Snelheid: 4,0 ml/uur
Tijdsduur: 24 uur

Taak 4)

Een patiënt van 79 heeft een hartinfarct gehad en is op dit moment stabiel waardoor hij naar de verpleegafdeling is overgeplaatst. Er wordt heparine (400 IE/ml) toegediend om de vorming van bloedstolsels tegen te gaan, met een dosering van 2,5 ml/uur gedurende 24 uur. Ongeveer twintig uur na de start van het toedienen van de heparine moet je de pompstand aanpassen omdat de hoeveelheid heparine die wordt toegediend niet meer correct is.

Taak: Verlaag de dosering van Heparine door de pompstand aan te passen. Hou je hierbij aan onderstaande gegevens.

Snelheid: 1,8 ml/uur
Tijd: Resterende tijd

Taak 8)

Een patiënt van 26 jaar is vanwege een gecompliceerde fractuur in het linkerbeen geopereerd . Hij is net overgeplaatst van de IC naar je afdeling. Zijn pijn wordt bestreden door middel van Dipidolor (0,4 mg/ml), geinfuseerd met een snelheid van 4,9 ml/uur. De patiënt roept naar hulp omdat hij plotseling hevige pijn heeft.

Taak: Geef een bolus Dipidolor van 3,6 ml.

Taak 10

De kuur is afgelopen en de patiënt hoeft dit medicijn niet meer toegediend te krijgen.

Taak: Stop de infusie en schakel de pomp uit.

APPENDIX C: Pre-questionnaire, welcome & instruction, informed consent

B1: Pre-questionnaire

Participantnummer (in te vullen door de onderzoeker): ______

Wat is je geslacht? Man ☐ Vrouw ☐

Wat is je hoogst voltooide opleiding?

Antwoord: ________________________

Wat is je beroep? Antwoord: ________________________

Hoeveel jaren ervaring heb je met het gebruiken van infuuspompen?

Antwoord: _______________________

Noem het aantal keren per week/dag dat je een spuitpomp gebruikt.

Antwoord: ______ / ______

Heb je ooit de Braun Perfusor® Space infuuspomp gebruikt? Ja ☐ Nee ☐
Als je de vooraangaande vraag met 'ja' hebt beantwoord:

• *Wanneer heb je de Braun Perfusor ® Space infuuspomp voor de laatste keer gebruikt ?*

Antwoord: ___________________
• *Hoe lang heb je met de Braun Perfusor® Space infuuspomp gewerkt?*

Antwoord: ____________________

Met welk merk spuitpomp werk je momenteel? ____________________________

Moet je na dit onderzoek nog beginnen met je dienst of ben je al klaar?

Ik moet nog beginnen ☐ Ik ben al klaar ☐

B2: Welcome & instruction

Beste deelnemer,

Allereerst willen wij u hartelijk bedanken dat u bereid bent om aan ons onderzoek mee te werken! Zonder u en de andere deelnemers zou het niet mogelijk zijn om dit onderzoek uit te voeren.

Tijdens dit onderzoek zullen wij u verscheidene taken uit laten voeren met twee verschillende interfaces van spuitpompen, gepresenteerd op een tablet. Wij zullen de tijd die u nodig heeft om deze taken te volbrengen daarbij meten, ook zullen wij de huidgeleiding meten door middel van de sensor om uw pols. Voordat wij met het onderzoek starten zijn er nog een aantal belangrijke dingen die u moet weten:

- wij testen de werking van de twee interface's en niet uw prestatie
- het is mogelijk dat wij u onderbreken tijdens het uitvoeren van één van de taken
- de instructie van de taken zal zowel schriftelijk als mondeling plaatsvinden
- het is belangrijk dat u de probeert om de taken **zo snel mogelijk** maar wel **zo accuraat mogelijk** uit te voeren
- na het uitvoeren van iedere taak vragen wij u op een schaal aan te geven hoe belastend u de taak vond
- we gaan tevens uw mentale werkbelasting meten met een pols-armband.
- de sessies worden met video opgenomen
- het is **belangrijk** dat u aangeeft wanneer u kunt beginnen met de taak en wanneer u klaar bent met de taak

Als u geen vragen meer heeft kunnen we nu starten met het onderzoek.

Met vriendelijke groeten,

Raphaela, Jan & Frauke

B3: Informed consent/non-disclosure agreement

Toestemmingsverklaring

voor deelname aan het wetenschappelijk onderzoek:

Ik stem ermee in deel te nemen aan dit onderzoek. Ik begrijp dat mijn deelname op vrijwillige basis plaatsvindt. Ik begrijp eveneens dat ik op elk moment kan beslissen de voortgang van mijn deelname stop te zetten indien ik overlast of ongemak ondervind, zonder dat ik daarvoor een reden hoef op te geven.

De volgende dingen zijn mij duidelijk:

- Het doel van dit onderzoek is om het design van twee verschillende spuitpomp interfaces te evalueren en deze met elkaar te vergelijken. We testen dus niet jou, maar de interfaces.

- Alle data die verzameld worden door de onderzoeker zullen volledig anoniem zijn en niet gekoppeld worden aan mijn deelname.

- Ik begrijp dat ik mijn opgenomen data niet in kan zien in verband met mogelijke misinterpretaties.

- Ik begrijp dat ik de inhoud van het onderzoek vertrouwelijk behandel en geen informatie aan derden verstrek.

Deelnemer: Ik stem toe met deelname aan het onderzoek.

Achternaam en voorletters:

Handtekening:

Datum:

Tijdstip:

Achternaam en voorletters onderzoeker:

Handtekening:

Datum:

APPENDIX D: Post interview questions

1) Voor welk interface heb je voorkeur? Waarom?

2) Voorkeur voor interface per specifieke functie:

2.1) Aan- en uitschakelen

2.2) kiezen/aanpassen van instellingen

2.3) starten/stoppen van infusie

2.4) opzoeken van informatie

2.5) Bolus

2.6) Alarm

3) Welke problemen had je bij het aan- en uitschakelen?

3.1) Bij de Braun? Suggesties voor verbeteringen?

3.2) Bij het nieuwe interface? Suggesties voor verbeteringen?

4) Welke problemen had je bij het kiezen/aanpassen van instellingen?

4.1) Bij het nieuwe interface? Suggesties voor verbeteringen?

4.2) Bij de Braun? Suggesties voor verbeteringen?

5) Welke problemen had je bij het starten/stoppen van infusie?

5.1) Bij de Braun? Suggesties voor verbeteringen?

5.2) Bij het nieuwe interface? Suggesties voor verbeteringen?

6) Welke problemen heb je bij het opzoeken van informatie?

6.1) Bij het nieuwe interface? Suggesties voor verbeteringen?

6.2) Bij de Braun? Suggesties voor verbeteringen?

7) Welke problemen had je bij de bolus functie?

7.1) Bij de Braun? Suggesties voor verbeteringen?

7.2) Bij het nieuwe interface? (Automatische bolus functie? Manuele?) Suggesties voor verbeteringen?

8) Welke problemen had je bij de alarm functie?

1) Bij het nieuwe interface? Suggesties voor verbeteringen?

2) Bij de Braun? Suggesties voor verbeteringen?

APPENDIX E: SPSS Syntax

Amplitudes

```
GET

FILE='C:\Users\MeeTy\Dokumente\master\masterthese\DATA_SPSS_scl_beweging_00
0009_03-1.sav'.
DATASET NAME DataSet1 WINDOW=FRONT.
* Generalized Estimating Equations.
GENLIN amplitude BY task type_of_pump repetition user_group
(ORDER=ASCENDING) WITH beweging
  /MODEL task type_of_pump repetition user_group INTERCEPT=YES
 DISTRIBUTION=GAMMA LINK=LOG
  /CRITERIA METHOD=FISHER(1) SCALE=MLE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).



* Generalized Estimating Equations.
GENLIN amplitude BY task type_of_pump repetition user_group
(ORDER=ASCENDING) WITH beweging
  /MODEL task type_of_pump repetition user_group INTERCEPT=YES
 DISTRIBUTION=GAMMA LINK=LOG
  /CRITERIA METHOD=FISHER(1) SCALE=MLE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=user_group SCALE=ORIGINAL COMPARE=user_group
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).


FILTER OFF.
USE ALL.
EXECUTE.
* Generalized Estimating Equations.
GENLIN amplitude BY task type_of_pump repetition (ORDER=ASCENDING) WITH
beweging
  /MODEL task type_of_pump repetition beweging INTERCEPT=YES
 DISTRIBUTION=GAMMA LINK=LOG
```

```
  /CRITERIA METHOD=FISHER(1) SCALE=MLE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).


USE ALL.
COMPUTE filter_$=(repetition = 1).
VARIABLE LABELS filter_$ 'repetition = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
* Generalized Estimating Equations.
GENLIN amplitude BY task type_of_pump repetition (ORDER=ASCENDING) WITH
beweging
  /MODEL task type_of_pump repetition INTERCEPT=YES
 DISTRIBUTION=GAMMA LINK=LOG
  /CRITERIA METHOD=FISHER(1) SCALE=MLE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).

USE ALL.
COMPUTE filter_$=(repetition = 2).
VARIABLE LABELS filter_$ 'repetition = 2 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
* Generalized Estimating Equations.
GENLIN amplitude BY task type_of_pump repetition (ORDER=ASCENDING) WITH
beweging
  /MODEL task type_of_pump repetition INTERCEPT=YES
 DISTRIBUTION=GAMMA LINK=LOG
  /CRITERIA METHOD=FISHER(1) SCALE=MLE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
```

```
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).


USE ALL.
COMPUTE filter_$=(repetition = 3).
VARIABLE LABELS filter_$ 'repetition = 3 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
* Generalized Estimating Equations.
GENLIN amplitude BY task type_of_pump repetition (ORDER=ASCENDING) WITH
beweging
  /MODEL task type_of_pump repetition INTERCEPT=YES
 DISTRIBUTION=GAMMA LINK=LOG
  /CRITERIA METHOD=FISHER(1) SCALE=MLE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).
```

SCL

```
GET

FILE='C:\Users\s1024833\AppData\Local\Temp\DATA_SPSS_scl_beweging_000009_03
-1.sav'.
DATASET NAME DataSet1 WINDOW=FRONT.
* Generalized Estimating Equations.
GENLIN scl BY task type_of_pump repetition (ORDER=ASCENDING) WITH beweging
  /MODEL task type_of_pump repetition beweging INTERCEPT=YES
 DISTRIBUTION=GAMMA LINK=LOG
  /CRITERIA METHOD=FISHER(1) SCALE=MLE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
```

```
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).



* Generalized Estimating Equations.
GENLIN scl BY task type_of_pump repetition (ORDER=ASCENDING) WITH beweging
  /MODEL task type_of_pump repetition INTERCEPT=YES
 DISTRIBUTION=GAMMA LINK=LOG
  /CRITERIA METHOD=FISHER(1) SCALE=MLE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).



USE ALL.
COMPUTE filter_$=(repetition = 1).
VARIABLE LABELS filter_$ 'repetition = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
* Generalized Estimating Equations.
GENLIN scl BY task type_of_pump repetition (ORDER=ASCENDING) WITH beweging
  /MODEL task type_of_pump repetition INTERCEPT=YES
 DISTRIBUTION=GAMMA LINK=LOG
  /CRITERIA METHOD=FISHER(1) SCALE=MLE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
```

```
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).



USE ALL.
COMPUTE filter_$=(repetition = 2).
VARIABLE LABELS filter_$ 'repetition = 2 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
* Generalized Estimating Equations.
GENLIN scl BY task type_of_pump repetition (ORDER=ASCENDING) WITH beweging
  /MODEL task type_of_pump repetition INTERCEPT=YES
 DISTRIBUTION=GAMMA LINK=LOG
  /CRITERIA METHOD=FISHER(1) SCALE=MLE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).



USE ALL.
COMPUTE filter_$=(repetition = 3).
VARIABLE LABELS filter_$ 'repetition = 3 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
* Generalized Estimating Equations.
GENLIN scl BY task type_of_pump repetition (ORDER=ASCENDING) WITH beweging
  /MODEL task type_of_pump repetition INTERCEPT=YES
 DISTRIBUTION=GAMMA LINK=LOG
  /CRITERIA METHOD=FISHER(1) SCALE=MLE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).
```

```
USE ALL.
COMPUTE filter_$=(type_of_pump = 1).
VARIABLE LABELS filter_$ 'type_of_pump = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
* Generalized Estimating Equations.
GENLIN scl BY task type_of_pump repetition (ORDER=ASCENDING) WITH beweging
  /MODEL task type_of_pump repetition INTERCEPT=YES
 DISTRIBUTION=GAMMA LINK=LOG
  /CRITERIA METHOD=FISHER(1) SCALE=MLE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).
```

```
USE ALL.
COMPUTE filter_$=(type_of_pump = 2).
VARIABLE LABELS filter_$ 'type_of_pump = 2 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
* Generalized Estimating Equations.
GENLIN scl BY task type_of_pump repetition (ORDER=ASCENDING) WITH beweging
  /MODEL task type_of_pump repetition INTERCEPT=YES
 DISTRIBUTION=GAMMA LINK=LOG
  /CRITERIA METHOD=FISHER(1) SCALE=MLE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).
```

NS.SCR frequency

```
GET

FILE='C:\Users\MeeTy\Dokumente\master\masterthese\DATA_SPSS_scl_beweging_00
0009_03-1.sav'.
DATASET NAME DataSet1 WINDOW=FRONT.
* Generalized Estimating Equations.
GENLIN no_scr BY task type_of_pump repetition user_group (ORDER=ASCENDING)
WITH beweging
  /MODEL task type_of_pump repetition user_group INTERCEPT=YES
 DISTRIBUTION=POISSON LINK=LOG
  /CRITERIA METHOD=FISHER SCALE=DEVIANCE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=user_group SCALE=ORIGINAL COMPARE=user_group
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).



USE ALL.
COMPUTE filter_$=(repetition = 1).
VARIABLE LABELS filter_$ 'repetition = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
* Generalized Estimating Equations.
GENLIN no_scr BY task type_of_pump repetition user_group (ORDER=ASCENDING)
WITH beweging
  /MODEL task type_of_pump repetition user_group INTERCEPT=YES
 DISTRIBUTION=POISSON LINK=LOG
  /CRITERIA METHOD=FISHER SCALE=DEVIANCE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=user_group SCALE=ORIGINAL COMPARE=user_group
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
```

```
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).



USE ALL.
COMPUTE filter_$=(repetition = 2).
VARIABLE LABELS filter_$ 'repetition = 2 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
* Generalized Estimating Equations.
GENLIN no_scr BY task type_of_pump repetition user_group (ORDER=ASCENDING)
WITH beweging
  /MODEL task type_of_pump repetition user_group INTERCEPT=YES
 DISTRIBUTION=POISSON LINK=LOG
  /CRITERIA METHOD=FISHER SCALE=DEVIANCE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=user_group SCALE=ORIGINAL COMPARE=user_group
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).



USE ALL.
COMPUTE filter_$=(repetition = 3).
VARIABLE LABELS filter_$ 'repetition = 3 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
* Generalized Estimating Equations.
GENLIN no_scr BY task type_of_pump repetition user_group (ORDER=ASCENDING)
WITH beweging
  /MODEL task type_of_pump repetition user_group INTERCEPT=YES
 DISTRIBUTION=POISSON LINK=LOG
  /CRITERIA METHOD=FISHER SCALE=DEVIANCE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=user_group SCALE=ORIGINAL COMPARE=user_group
CONTRAST=PAIRWISE PADJUST=LSD
```

```
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).



USE ALL.
COMPUTE filter_$=(task = 1).
VARIABLE LABELS filter_$ 'task = 1 (FILTER)'.
VALUE LABELS filter_$ 0 'Not Selected' 1 'Selected'.
FORMATS filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
* Generalized Estimating Equations.
GENLIN no_scr BY task type_of_pump repetition user_group (ORDER=ASCENDING)
WITH beweging
  /MODEL task type_of_pump repetition user_group INTERCEPT=YES
 DISTRIBUTION=POISSON LINK=LOG
  /CRITERIA METHOD=FISHER SCALE=DEVIANCE MAXITERATIONS=100 MAXSTEPHALVING=5
PCONVERGE=1E-006(ABSOLUTE) SINGULAR=1E-012 ANALYSISTYPE=3(WALD) CILEVEL=95
LIKELIHOOD=FULL
  /EMMEANS TABLES=task SCALE=ORIGINAL COMPARE=task CONTRAST=PAIRWISE
PADJUST=LSD
  /EMMEANS TABLES=type_of_pump SCALE=ORIGINAL COMPARE=type_of_pump
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=repetition SCALE=ORIGINAL COMPARE=repetition
CONTRAST=PAIRWISE PADJUST=LSD
  /EMMEANS TABLES=user_group SCALE=ORIGINAL COMPARE=user_group
CONTRAST=PAIRWISE PADJUST=LSD
  /REPEATED SUBJECT=Subject_number
WITHINSUBJECT=task*type_of_pump*repetition*five_second_segment SORT=YES
CORRTYPE=EXCHANGEABLE ADJUSTCORR=YES COVB=ROBUST MAXITERATIONS=100
PCONVERGE=1e-006(ABSOLUTE) UPDATECORR=1
  /MISSING CLASSMISSING=EXCLUDE
  /PRINT CPS DESCRIPTIVES MODELINFO FIT SUMMARY SOLUTION (EXPONENTIATED).



SORT CASES  BY Subject_number.
SPLIT FILE LAYERED BY Subject_number.
CORRELATIONS
  /VARIABLES=no_scr BSMI
  /PRINT=TWOTAIL NOSIG
  /MISSING=PAIRWISE.



SORT CASES  BY Subject_number.
SPLIT FILE LAYERED BY Subject_number.
CORRELATIONS
  /VARIABLES=amplitude BSMI
  /PRINT=TWOTAIL NOSIG
  /MISSING=PAIRWISE.
```