# Method generation for the evaluation of virtual buttons with multimodal feedback

vorgelegt von

M. Sc. Human Factors

Lisa Diwischek

geb. in Hildesheim

von der Fakultät V – Verkehrs- und Maschinensysteme

der Technischen Universität Berlin

zur Erlangung des akademischen Grades

Doktorin der Ingenieurwissenschaften

— Dr.-Ing. —

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr.-Ing. Henning J. Meyer

Gutachter: Prof. Dr.-Ing. Matthias Rötting

Gutachter: Prof. Dr. Mark Vollrath

Tag der wissenschaftlichen Aussprache: 30. August 2016

Berlin, 2016

## Abstract

The main topic of this thesis is a method generation for the evaluation of virtual buttons with multimodal feedback. Virtual switches or buttons are becoming more present in the automotive domain, and force-sensing for activation combined with vibrational and auditory feedback (so-called 3D switches) is the latest evolutionary step in switching technology. To evaluate this new technology, a new evaluation method needed to be developed. In order to evaluate preferences in concepts directly related to the users, the repertory grid method was employed for the generation of items. These items were accumulated to a fixed grid with variable addition possibilities through each testing phase. In total, three pre-studies and one main experiment were conducted. In the first study, 16 healthy US citizens conducted a pairwise comparison and a repertory grid in a lab study with different feedback settings in terms of frequency and amplitude in order to see whether additional information could be obtained through employing the grid. The second and third study took place in a driving simulator with German subjects, investigating different methodical aspects and in terms of switch characteristics, activation and deactivation forces, their optimal delta, and, for push-and-hold functions, the time between the different feedback actuations. Additional to the preference rating, workload and driving data was collected. 39 and 30 healthy participants were tested, respectively, and the fixed grid enhanced. In the main experiment, 40 healthy German participants compared mechanical, capacitive and 3D switches. Preference, workload, eye tracking and driving data was obtained for the analysis.

Concluding, optimal settings for 3D switches could be obtained in the three pre-studies and their automotive applicability could be substantiated in a driving simulator context. A new method was developed iteratively to measure the user preference for multimodal, vibrational and auditory feedback, and force interaction, which can be applied in other domains or steering wheel switch research, which is directly related to the users' concepts of interaction.

## Content

1. Introduction ... 1
2. Theoretical part ... 3
2.1 Human perception ... 3
2.1.1 Visual perception ... 4
2.1.2 Auditory perception ... 6
2.1.3 Haptic and tactile perception ... 8
2.1.3.1 Surface perception ... 9
2.1.3.2 Force perception ... 11
2.1.3.3 Neuronal basis ... 12
2.1.3.4 Confounding factors in tactile perception ... 13
2.2 Perception thresholds ... 14
2.3 Psychophysiological measurement of auditory perception ... 17
2.4 Psychophysiological measurement of tactile perception ... 19
2.5 Interaction principles ... 22
2.5.1 Visual interaction ... 22
2.5.2 Auditory interaction ... 23
2.5.3 Tactile interaction ... 23
2.6 Multimodality and its applications ... 25
2.6.1 Theories on attention and multimodal resources ... 25
2.6.2 Multimodal integration ... 29
2.6.3 Multimodal warning strategies ... 30
2.6.4 Multimodality in comfort functions ... 31
2.7 Usability and suitability ... 33
2.7.1 Definitions ... 33
2.7.2 Criteria of suitability ... 34
2.7.3 Tactile usability and suitability ... 35
2.8 Constructs of user experience ... 36
2.9 Interaction media ... 38
2.9.1 Conventional switches ... 39
2.9.1.1 Design guidelines for conventional switches ... 40

2.9.2 Resistive and capacitive touchscreens ..... 42
2.9.2.1 Design guidelines for touchscreens ..... 46
2.9.3 Virtual force sensitive switches ..... 48
2.10 Previous studies ..... 49
2.10.1 Expert evaluation ..... 49
2.10.1.1 Method ..... 49
2.10.1.2 Results ..... 50
2.10.2 Material design ..... 50
2.10.2.1 Method ..... 51
2.10.2.2 Results ..... 51
2.10.3 Single switch activation functionality ..... 51
2.10.3.1 Method ..... 52
2.10.3.2 Results ..... 52
2.10.4 Auditory settings ..... 52
2.10.4.1 Method ..... 53
2.10.4.2 Results ..... 53
2.10.5 Frequency and waveform ..... 54
2.10.5.1 Method ..... 54
2.10.5.2 Results ..... 55
2.10.5.3 Discussion ..... 58
3. Methods ..... 59
3.1 Methods of performance measurement ..... 59
3.2 Psychophysiological measurement methods ..... 60
3.3 Methods of subjective assessment ..... 60
3.3.1 Interviews in the usability context ..... 61
3.3.2 Pairwise comparisons ..... 62
3.3.3 Semantic Differential ..... 63
3.3.4 Similarity estimation ..... 63
3.3.5 Questionnaires ..... 64
3.3.6 Repertory Grid ..... 64
3.4 Method selection and adaptation ..... 68

3.5 Hardware ... 70
3.5.1 3D Switches steering wheel ... 70
3.5.2 Signal generation ... 71
3.5.3 Driving simulator ... 72
3.6 Preliminary studies ... 73
3.6.1 First preliminary experiment: frequency and amplitude of the tactile feedback signal 73
3.6.1.1 Method ... 73
3.6.1.2 Results of the first preliminary test ... 74
3.6.2 Second preliminary experiment: force settings of the skip and mode switches ... 77
3.6.2.1 Method ... 77
3.6.2.2 Results of the second preliminary test ... 80
3.6.3 Third preliminary experiment: Time interlock and force delta ... 88
3.6.3.1 Method ... 89
3.6.3.2 Results of the third preliminary test ... 91
3.6.3.3 Discussion ... 99
3.7 Methodology of the main experiment ... 100
3.7.1 Participants ... 100
3.7.2 Hypotheses ... 101
3.7.3 Stimuli preparation ... 102
3.7.4 Steering wheels ... 102
3.7.5 Eyetracking equipment ... 102
3.7.6 Final Grid design ... 103
3.7.7 Procedure ... 105
4. Results ... 107
4.1 Switch performance ... 107
4.2 Driving performance ... 108
4.3 Subjective preference data ... 112
4.4 Principal Component Analysis ... 115
4.5 Eyetracking data ... 120
5. Discussion ... 125
Literature ... 131

## Indexes

### *List of Figures*

Figure 1. Cross-Section of the Cochlea ........ 6
Figure 2. The Human Ear........ 6
Figure 3. Human Skin Structure ........ 8
Figure 4. Wickens' Multiple Resource Model. Adapted from Wickens (2008)........ 26
Figure 5. From physical to affective layer in tactile interaction. ........ 37
Figure 6. Multifunctional Steering Wheel. ........ 41
Figure 7. Force Time Diagram........ 42
Figure 8. Driving simulator........ 72
Figure 9. Acceleration measurement setup. ........ 78
Figure 10. NASA-TLX workload for the second preliminary test score........ 81
Figure 11. Mode Scree Plot. ........ 85
Figure 12. Skip Scree Plot. ........ 85
Figure 13. NASA-TLX results of the third preliminary study........ 93
Figure 14. Snap ratio distance to ideal configuration. ........ 95
Figure 15. Time between pulses distance to ideal configuration. ........ 96
Figure 16. Scree plot of third preliminary study. ........ 97
Figure 17. Interaction effect for time on task. ........ 108
Figure 18. Standard deviation of speed........ 109
Figure 19. Differences per switch type for steering wheel jerks. ........ 112
Figure 20. Profile of the distance to ideal. ........ 114
Figure 21. Sum Scores of the significant differences to the ideal configuration. ........ 115
Figure 22. Scree plot of main experiment. ........ 116
Figure 23. McDonald's omega determined factor structure. ........ 118
Figure 24. Final Factor Structure. ........ 120
Figure 25. Boxplots for steering wheel AOI. ........ 121
Figure 26. Boxplots for road AOI. Legend equivalent to Figure 25........ 121
Figure 27. Post-hoc results for road AOI. ........ 122
Figure 28. Median and mean of percent dwell time on steering wheel AOI. ........ 123
Figure 29. Mean glance duration on the steering wheel per switch type........ 124

## *List of Tables*

Table 1. Mechanoreceptor cells. .................................................................................................... 9
Table 2. Post-hoc comparisons using Tukey HSD. ..................................................................... 56
Table 3. Post-hoc comparisons using three different methods. .................................................. 57
Table 4. Factorial design of the first preliminary experiment. .................................................... 74
Table 5. Factorial design of the second preliminary experiment. ................................................ 78
Table 6. Correlations with good/bad scale. ................................................................................. 84
Table 7. Factor loadings of principal component analysis. ......................................................... 87
Table 8. Factorial design of the third preliminary experiment. .................................................... 90
Table 9. Factor analysis results of the third preliminary study. .................................................... 98
Table 10. Factorial Design for the main experiment. ................................................................ 102
Table 11. Lane keeping quality differences for driving condition. ............................................ 110
Table 12. Steering wheel velocity differences for driving condition. ......................................... 110
Table 13. Steering wheel acceleration differences for driving condition. ................................. 111
Table 14. Significant effects on driving maneuver in numbers of critical lane deviations. ....... 111
Table 15. Significant post-hoc results for the main effect on driving maneuver. ...................... 111
Table 16. Factorial structure of final grid. ................................................................................. 119

# 1. Introduction

In the field of Human-Machine-Interaction, the automotive domain owns a multiplicity of special characteristics: no highly trained subjects, a wide variety of users, and a certain amount of safety critical aspects to consider. When designing for a new interface, safety is the utmost crucial matter, but the user acceptance has its own entitlement; because what is the use of a bullet-proof concept that is never used? On the other hand, if a much preferred device endangers its user, the device cannot be deemed useful. All aspects of usability, efficiency, effectivity, and user acceptance have therefore to be taken into account. Performance as well as acceptance measures will therefore be evaluated in most of the studies conducted for the present thesis. The focus is on user acceptance because for efficiency and effectivity, satisfactory measurements do already exist.

In the area of user acceptance, a lot of different qualitative and quantitative methods have been designed and implemented. The gap between both methods is considered almost impossible to span: practitioners often focus on one or the other, reject the opposite approach, and discuss the quality criteria of both quite feverishly. This thesis aims at narrowing down this gap, providing users with the possibility to enlarge the known space of judgment criteria with their own words, and researchers nevertheless with a set of items usable for quantitative statistical analysis. This is conducted via an adaptation of the *Repertory Grid* method, which will be discussed in the methodological part. An instrument developed to measure certain switch aspects needs to be constructed around those.

As the human being with certain perceptional, procedural and attentional capabilities and resultant preferences is in the focus, the first chapters of the theoretical part, *Human perception* and *Perception thresholds*, will deal with these topics and how they have to be taken into account for the present technology. As the measurement of these aspects has a long tradition in user research, *Psychophysiological measurements* are discussed subsequently. The object of the interaction, technologically developed or rather basic, plays a crucial part in it; not only the user needs to be discussed, but the object and its design principles as well. These principles are covered in the chapter *Interaction principles*.

As two different types of interaction play the most important role, their interaction with each other and the existing applications are discussed in the chapter *Multimodality and its applications*.

Furthermore detailed are the requirements of the applications in the adjacent chapter, *Usability and suitability*.

The present work focuses on multimodal perception, mainly vibrotactile and auditory, as these are the primarily important feedback channels for *3D* switches, which are implemented this way in order to not compete for visual attention with the main driving task. 3D switches are to date steering wheel switches which have accurate force sensing and the capability of providing aforementioned feedback. These switches have been thoroughly investigated in the present work. They are discussed, among the more traditional technology of mechanical switches and the consumer-electronic oriented touchscreen technology, in the chapter *Interaction Media*. Studies conducted with the 3D Technology previously constitute the end of the theoretical part.

Although the instrument in the present work is developed around steering wheel switches, it is not necessary limited to this field of application. Vibrotactile and auditory feedback coupled with each other do occur on a variety of technologies and can be evaluated with this instrument. The focus is not on the functionality of a device, but the feedback itself. Interactions with a certain functionality may occur; this is not investigated in the present work. Further methodological aspects are discussed in the method chapter; among different approaches to quantitative and qualitative research, Repertory Grids in their original research domain and the adaptations which were conducted for the present work are discussed.

The experimental part is divided into three sections: the first preliminary experiment is still conducted in a lab setting, whereas the second and third take place in a driving simulator. The commonality of these preliminary studies is that the Repertory Grid technique is used iteratively to develop response categories, and that they all deal with certain aspects of 3D switches. The main experiment combines the previously developed response categories and compares a rough capacitive switch, a mechanical switch, and the 3D switches and their distraction potential. In order to make the development suitable for a user group as wide as possible, all age groups that currently possess a driving license have been included in the study. Both U.S. and German citizens were subjects in the studies. The final instrument can be found in appendix c) d *Final Instrument*, both in a German and an English version.

# 2. Theoretical part

## *2.1 Human perception*

The perception of humans is a mixture of passive and active mechanisms. For object identification, external stimuli like sound waves, light, or vibrations need to pass the sensory organs and are processed in the brain. On the other hand, past experiences and expectations shape the way in which we perceive the different stimuli. The difference between the object of perception and the reconstruction in every subjective brain has been a topic of disputation since the origins of psychology and can take up only a small part of this thesis. Whereas the more idealistic-driven perspective tends to stress the subjective part of the perception, thoroughly attributing every perception as subjective only leads to the relation of reality, the denying of the influence of objective factors (e.g. social norms or laws) and, in the last consequence, to a denial of an existing objective world as such. A purely deterministic view, on the other hand, which claims that the objects are speaking for themselves and may vary in their impact on human beings only, denies the complex individual processes which form the world we are experiencing. It is therefore necessary to balance subjective and objective influences on perception when designing an interface: physiological perception and information processes have to be considered as objective, determined mechanisms. Standards and norms for interfaces, which represent the culture individuals are used to in terms of interaction have to be taken into account as they are objective influences on the aforementioned processes - but also subjective ratings and preferences for an interface must be considered, inasmuch as that in interface should be designed in a way that the individual user is animated to perform the interaction.

The virtual (as they are not mechanical) switches of the technology examined in this thesis can stimulate the perception of haptic, tactile and auditory signals, but of course there is also a visual interface to it, so that the three senses touch, hearing and vision are of interest. The following section deals with the rough basics of visual, auditory and haptic as well as tactile perception. Sensual perception is prone to errors, but no knowledge of the outside world can be built up or exist without sensual perception.

The first subchapter of the perception chapter discusses the visual, the second subchapter the auditory part. The focus of this thesis is on the tactile design for an interface, so accordingly, the last and most detailed part of the chapter deals with haptic and tactile perception mechanisms. A distinguished chapter deals with perception thresholds for a specific subgroup and follows the chapter of perception mechanisms. The interplay between the modalities is modulated through the focus of attention, which will be discussed in the chapter *Theories on attention and multimodal resources* in more detail.

### 2.1.1 Visual perception

The visual perception has always played an important role in philosophy, science, and art. As early as in ancient Greece there were disputations about the structure of visual perception including contributions by Euclid, Ptolemy, Plato and Aristotle. The transformation from touching everything towards recognizing an object from merely looking at it plays an important role in the development of every child as well as for civilization as a whole (Elias, 1969). The following sector will provide an overview of basic visual perception processes, but not go into much detail, as the focus of this work lies on the tactile part.

The perceptional organ of vision, the eye, works similar to a camera. Through an objective, consisting of *Cornea*, *Lens* and *Anterior Chamber*, the reflection of light travels through and is projected on the background of the *Retina* (Birbaumer & Schmidt, 1996). The field of vision has a cone angle of 100°, for both eyes in sum 170° horizontal and 110° vertical. The visual perception is divided into *foveal* and *ambient* vision, which corresponds to different kinds of perception cells. The *foveal* vision is located at the most precise point of view, the *Fovea Centralis*. At this point, the *Retina* has its highest resolution, because there is no cover on the *Cones*; but as this area consists of *Cones* only, the perception at suboptimal light conditions may suffer in terms of accuracy (Kirschbaum, 2008). The most precise visual perception is possible only around 1° of the fixated point of view only (Joos, Rötting, & Velichkovsky, 2002). The farther objects are located in the peripheral field of vision, the more reduced is the resolution and the more the color intensity decreases. Already at 3° deviation from the object of fixation, visual acuity is reduced by half (Birbaumer & Schmidt, 1996). In addition, the distance between the *Cones* and the number of *Rods* increases with increasing distance from the *Fovea* (Schandry, 2006). In peripheral vision, only moving objects or brightness changes can be perceived due to the lack of *Cones*, but especially the first one very accurate, due to the high number of *Rods* (Birbaumer & Schmidt, 1996). The object which passes the lens of the eye is projected on the *Retina* upside down, due to optical refraction

(Goldstein, 2008). The *Iris* modulates the light intrusion in the eye (Birbaumer & Schmidt, 1996). *Cones* and *Rods* pass the information on to bipolar cells, which synapse with *Ganglion* cells. Those encode the light intrusion information into electrical impulses and transfer it through the *Optical Nerves*, of which every eye has two, corresponding to the right and the left side of the retina. They exchange parts of their information at the *Optical Chiasma*, where both right *Retina* information parts go to the right side of the brain, and both left side information to the left side of the brain. As the perception is upside down, this leaves each side of the brain processing the information from the opposite direction (Birbaumer & Schmidt, 1996).

Perceptual processes as object detection, spatial orientation, or attention are processed in different areas of the brain. The synthesis of this nerve cell information into a meaningful object (*percept*) happens in different areas in the brain as well as the cognitive assessment. As the information from the visual sense is prone to errors due to its construction in the brain, there are different visual perception biases and *Gestalt* principles in each of these domains which influence the visual perception. Some of them have been known, especially in the field of art (as so-called *Trompe-l'œil*, perspective paintings), from as early as paintings in ancient Pompeji, but the systematic documentation and the investigation on underlying cognitive processes is relatively new and origins in the field of cognitive psychology. Their characteristics and origins are not discussed here, as the focus lies on haptic and tactile processes. Attentional mechanisms, as an exception, will be discussed in the chapter *Theories on attention and multimodal resources*. Visual reaction times (the time it takes to create a response) to a stimulus vary between 180-200 ms (Kosinski, 2012) and the different stimuli take between 20-40 ms for processing (Marshall, Talbot, & Ades (1943), cited after Kosinski (2012)). The visual system is therefore the slowest system in processing information.

The next subchapter will focus on another perception modality, the auditory perception.

### 2.1.2 Auditory perception

When a sound wave travels through the external auditory canal and arrives at the eardrum (*Tympanic membrane*), the vibration of the membrane is transferred physically to the *Ossicles* (*Malleus, Incus and Stapes*). The *Stapes*, the innermost part of the *Ossicles*, transfers the signal to the *Cochlea* via the oval window. The *Cochlea* has a spiral form which normally contains of two and a half convolutions. Those convolutions are differentiated into *Scala Vestibuli*, *Scala Tympani*, and *Scala Media* (s. Figure 1. Cross-Section of the Cochlea). The whole *Cochlea* is filled with liquid (Birbaumer & Schmidt, 1996). This liquid differs for the regions of the *Cochlea*. In the *Scala Vestibuli* and the *Scala Tymphani*, the liquid is called *Perilymph*, which is rich in sodium (140 mM) and poor in potassium (5 mM) and calcium (1.2 mM). The *Scala Media* is filled with *Endolymph*, a product of *Perilymph*, which is rich in potassium and, compared to the *Perilymph*, has a positive potential (+80mV) (Delprat, 2013). This potential difference enables the transformation from a sound wave to a nerve impulse.

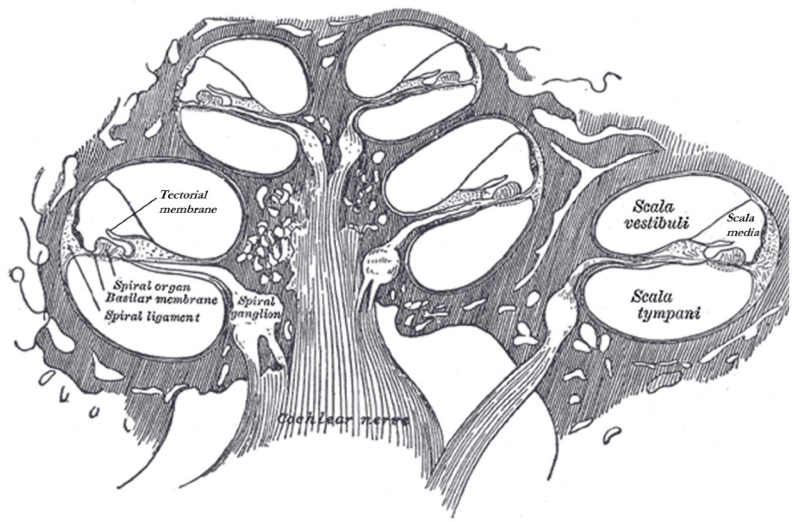



*Figure 2. Cross-Section of the Cochlea (https://commons.wikimedia.org/wiki/File:Gray928.png, 2016)*

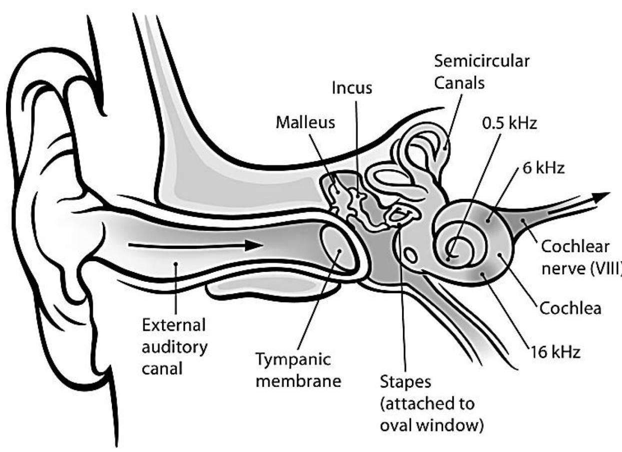



*Figure 1. The Human Ear (adapted from https://commons.wikimedia.org/wiki/File:10.1371_journal.pbio.0030137.g001-L-A.jpg, 2016).*

The *Cochlea* is vertically divided on the inside by the *Basilar Membrane*. This membrane has a certain resonance spot for each frequency which creates specific vibration patterns. High frequencies are located on the basis, whereas low frequencies are sensed at the upper part of the *Cochlea* (Goldstein, 2008). Embedded in the *Basilar Membrane* are the hair cells, which consist of inner and outer hair cells. They differ in location (the inner hair cells are inside the *Basilar Membrane*, whereas outer hair cells grow out of the *Basilar Membrane* into the *Scala Media* (Birbaumer & Schmidt, 1996)). The outer hair cells therefor dwell in the *Endolymph* fluid of the *Scala Media.* When sound is transferred to the *Cochlea* via the *Oval Window*, it travels upwards through the *Vestibule*, is transferred at the top (the *Helicotrema*), and travels downwards through the *Scala Tympani* back to the so-called *Round Window*, a pressure homeostasis membrane next to the *Oval Window* (Birbaumer & Schmidt, 1996). At the same time, the *Scala Vestibuli* and the *Scala Tympani* excite the *Scala Media* through each of their shared membrane. This excitement results into the local specific vibration patterns mentioned above. They also excite the *Tectorial Membrane*, a small membrane tissue inside of the *Scala Media.* The relative position of the *Tectorial Membrane* in regard to the *Basilar Membrane* changes due to the different vibration patterns. This shearing movement between both *Tectorial* and *Basilar Membrane* causes the excitement of the *Stereocilias* (small cell prolongings of the outer hair cell which look like hair) of the outer hair cells, and this leads to changes in their permeability, leading to the depolarization of the outer hair cells (Birbaumer & Schmidt, 1996). This depolarization causes the inward displacement of anions, causing the protein to shorten and leading to a contraction of the outer hair cells (Pujol, 2013). Instead of transmitting the auditory signal directly to the brain, the outer hair cells reinforce in this way the vibration patterns and increase not only frequency, but also sensitivity (Pujol, 2013) in the range of 50 dB to 60 dB (Birbaumer & Schmidt, 1996).

Contrary to the outer counterpart, the inner hair cells are not responsible for amplification, but for transmission of the signal. They are excited by the shearing movement between *Tectorial* and *Basilar Membrane*, and depolarize as well. This causes here the opening of voltage-sensitive calcium channels at this particular point and leads to the release of glutamate into the synaptic cleft, effectively activating the auditory nerve fibers and resulting in action potentials (Pujol, 2013).These action potentials are transmitted to the *Auditory Cortex*, but are pre-processed on their way through *Nucleus Cochlearis*, *Upper Olive*, *Colliculus Inferior* and *Corpus Geniculatum Mediale* (Goldstein, 2008).

The human auditory perception ranges between 20 and 16,000 Hz (Birbaumer & Schmidt, 1996). It has its highest sensitivity in the area between 2000-5000 Hz, the range in which human language is spoken (Birbaumer & Schmidt, 1996). Average reaction times to auditory stimuli vary between 140-160 ms (Kosinski, 2012). The processing takes about 8-10 ms ((Kemp & Bryan, 1973), cited after Kosinski (2012)). Auditory perception can therefore be viewed as the fastest perception modality.

The next subchapter discusses the main focus of this thesis, the tactile and, to some extent, the haptic perception modalities.

### 2.1.3 Haptic and tactile perception

Unlike with visual or auditory perception, haptic and tactile perception takes place over the whole human body, although the sensitivity and the types of perception vary across different locations. Most of these perception types origin in the human skin.

The skin is structured into the layers *Epidermis*, *Dermis*, *Subcutis/Hypodermis* and can be divided into hairless or *glaborous* skin and hairy skin (cp. Figure 3. *Human Skin Structure*). The different nerve fibers in the human skin are distributed in *Dermis* and *Hypodermis*, and vary in their density across the human body, which leads to the aforementioned different perceptual abilities of different body regions.

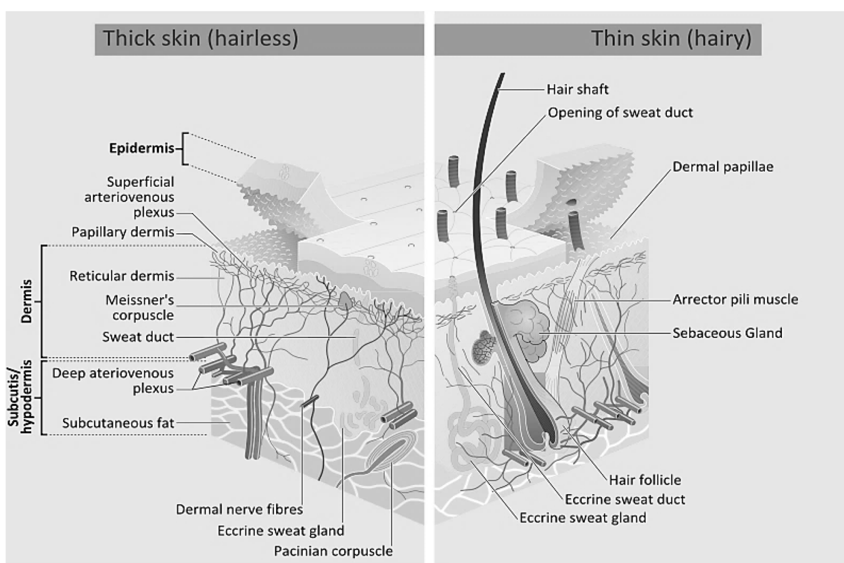



*Figure 3. Human Skin Structure (adapted from https://en.wikipedia.org/wiki/File:Skin_layers.png, 2016).*

Among haptic perception processes, there are two main variants: the perception of *tactile* and of *kinesthetic* information (Loomis & Lederman, 1986). Tactile perception is a cutaneous sense, including *Merkel-cells*, *Ruffini-corpuscles*, *Meissner- corpuscles*, *Vater-Pacini-corpuscles*, two kinds of *Nociceptors* (pain receptors) and two kinds of temperature receptors (Hsiao & Yau, 2008). The *kinesthetic* perception stems from the afferents in the muscles and provides information about static and dynamic body posture (Jones, 2000). Both processes together are, however, more than the sum of their parts: whereas passive tactile stimulation or isolated kinesthetic stimulation produces sensations, only the active and combined exploring enables object perception in space (Kaas, Stoeckel, & Goebel, 2008).

Tactile perception is also often referred to as surface perception, of which the subsequent first subchapter will be about. The second part, force perception, deals with the most relevant parts of kinesthetic information for the objective of this paper, which are sometimes used synonymously to the term haptic perception. Consecutively, the third part deals with the neuronal basis of tactile perception and finally discusses some confounding factors.

#### *2.1.3.1 Surface perception*

*Merkel-cells*, *Ruffini-corpuscles*, *Meissner-corpuscles*, and *Vater-Pacini-corpuscles* are *Mechanoreceptors* and part of the human surface perception. The following table (Table 1. *Mechanoreceptor cells*) lists their characteristics in terms of size of the perceptual cells, their domain, location in the skin, sensitivity to frequency and their type of adaptation.

*Table 1. Mechanoreceptor cells.*

| **Name** | **Size** | **Domain** | **Location** | **Frequency** | **Adaptation** |
|---|---|---|---|---|---|
| Merkel-cells | 80-120 nm(2) | Pressure(2) | Epithelium(2) | 0,3 - 3 Hz(1) | Slow(2)(SAI) |
| Ruffini-corpuscles | 5-10 µm(2) | Pressure(2) | Outer fibrous layer (2) | 100-500 Hz(1) | Slow(2)(SAII) |
| Meissner-corpuscles | 100-150 µm(2) | Touch(2) | Papillary layer(2) | 20-30 Hz(2) | Fast(1) (FAI) |
| Pacini-corpuscles | 1 mm (2) | Vibration(2) | Subcutis(2) | 20-1500 Hz(1) Maximum 200 Hz(2) | Fast(1)(FAII) |

*Note that these values may vary with the size and the frequency of a stimulus applied and should therefore provide rather an orientation or overview over the different types of mechanoreceptor cells. Additionally, the density values refer to the fingertips of a human and vary for other body areas.*

1 (“1-Consciousness-Sense-Touch-Anatomy-Receptors,” 2012)
2 (Halata & Baumann, 2008)

For haptic and tactile perception, the *glabrous* (hairless) skin on the fingertips is especially important. It is densely innervated by four kinds of myelinated axons which can be classified in two groups, *slow adapting* (SA) and *fast adapting* (FA) afferents. SA afferents do also fire in static conditions, whereas FA afferents are only active in dynamic conditions.

For each group, there are two categories, classified as Type I and II. Type I afferents have small receptive fields with sharp borders and a broad distribution of maximum sensitivity. Opposed to that, Type II afferents have larger receptive fields with less defined borders and do not have the broad distribution of maximum sensitivity zones. Sometimes, the channels are also classified into NP (*Non-Pacinian*) I, II, III and PC (*Pacinian Channel*); NP-I corresponds to FAI, NP-II to SA II, NP-III to SA-I and PC to the FAII channel (cp. Hatzfeld (2013)). Bolanowski, Gescheider, Verrillo, & Checkosky (1988) found evidence for these channels and indications for their relative independence from each other. There are some shared areas of perception nevertheless, and some frequencies may elicit several channels (cp. Table 1. *Mechanoreceptor Cells*). Overlaps are possible between the PC channel and NP-I or NP II. This might explain why the tactile perception abilities in psychophysiological tests of different channels are not as separated in terms of frequency as in the auditory domain, where the corresponding location for each frequency is quite specific (cp. subchapter *Auditory perception*).

SAI innervates *Merkel cells*, SAII *Ruffini corpuscles*, FAI *Meissner corpuscles*, and FAII *Pacini -corpuscles*. The innervation of Type I afferents has its highest density on the fingertips ($0.7 mm^{-2}$ for SAI and $1.4\ mm^{-2}$ for FAI). Type II afferents do not vary their innervation density in the glabrous skin ($0.09\ mm^{-2}$ for SAII and $0.33\ mm^{-2}$ for FAII) (cp. Goodwin & Wheat, 2008). The sensor distribution on a fingertip consists of 60% *Meissner-corpuscles*, 30% *Merkel-cells*, 5% *Ruffini-* and 5% *Pacini-corpuscles* (Hatzfeld, 2013).

Especially the fast-adapting *Pacini-corpuscles* are of interest as they react to vibration signals. Their high adaptation leads to good perception of signals with a sinus-waveform (Birbaumer & Schmidt, 1996). Other waveforms are normally not recommended for interaction design, as long as not only beginning and end of interaction is of interest, as the high adaptation of the *Pacini-corpuscles* leads to insensibility for a continuous signal. So far, there is no evidence of an effect of other waveforms on subjective quality perception for different sinusoidal signals (Altinsoy & Merchel, 2009).

#### *2.1.3.2 Force perception*

There are two different interpretations of force perception: one refers to the four surface perception cells from the subchapter *2.1.3.1 Surface* perception and is investigating the gripping or lifting an object. Accordingly, their main fields of investigation are named *grip force* and *load force*. They are both closely related to each other: *grip force* increases and decreases with *load force*, across different types of gripping and body movements (Flanagan & Tresilian, 1994). The underlying perception mechanisms are those of tactile perception.

The other refers to the *kinesthetic* sense. This means not *proprioception*, the perception of one's position in space, but rather the knowledge about the status of muscles and therefore the applied force on an object. Force is perceived through three kinds of cells apart from the *cutaneous* sense, of which two are located in muscle tissues and one in the juncture between *Muscle Tendons* and *extrafusal* muscle fibers: the *Golgi Tendon Organ* (Jones, 2000). The exact determination of the function of the *Golgi Tendon Organ* remains unclear yet, although there are some hints that it supports the avoidance of overly flexing or overstretching a muscle (Jones, 2000). The other two are primary and secondary *Spindle Receptors*, which are signaling "the velocity and direction of muscle stretch or limb movement" (Jones, 2000) and static muscle length or limb position (Jones, 2000). These cells transmit in a haptic interaction the kinesthetic information, e.g. in a conventional mechanical switch, the travel or displacement which takes place, and its velocity.

There are approximately 4,000 muscle spindles in each arm. The obvious smaller density than the one of tactile cells is explained by the fact that the number of muscle cells does not correspond to the ability of sensing kinesthetic information (Jones, 2000). Muscle spindles are innervated afferently by *Ia Sensory Fibers*, which are thick and myelinated (Birbaumer & Schmidt, 1996). The efferent innervation depends on the position of the muscle spindles. *Intrafusal Muscle Spindles* are innervated by Aγ-fibers which are 2-8μm in diameter and have an information transportation velocity of 15-30 m/s (Birbaumer & Schmidt, 1996). Aα-fibers innervate *Extrafusal Muscle spindles* with a velocity of 70-120 m/s and are 10-20 μm in diameter (Birbaumer & Schmidt, 1996).

*Golgi Tendon Organs* are even less frequent than spindle receptors and their number varies even more (Devanandan et al., 1983; cited after Jones, 2000). They are innervated by one to two myelinated, thick nerve fibers (Ib fibers), which are about 10-20 µm in diameter and their efferent innervation is via Aα-fibers (Birbaumer & Schmidt, 1996).

The next subchapter details the further processing of information derived from *tactile* or *kinaesthetic* perception.

#### *2.1.3.3 Neuronal basis*

A depolarization generates normally an action potential; a type of signal for which the nerve cell is sensitive is transmitted into an electric signal. When this signal is transmitted via afferent nerves, it passes the *Spinal Cord* in different ways. Myelinated nerves are called *White Matter* in the *Spinal Cord* in regard to their color distinction from the unmyelinated, *Grey Matter*. *Kinesthetic* information is transformed via the *Tracti Spinocerebellaris* (both anterior and posterior), which lay at the outer parts of the spinal cord, and some of it via the *Funiculus Posterior*, which is as implied by the name located in the rear part of the *Spinal Cord* and transfers additionally vibration, pressure, discrimination, and touch information. There are other channels which transfer pain and temperature sensing, but as they are not the main focus of this work, they are not considered further. Some information processing takes already place in the *Spinal Cord*, but only in terms of reflexes. The *Tractus Spinocerebellaris Posterior* ends in the *Cerebellum* whereas the *Funiculus Posterior* divides itself in *Fasciculus Gracilis* and *Fasciculus Cuneatus*, which are pre-processed in the *Nucleus Gracilis* and *Nucleus Cuneatus*. Before the *Thalamus*, the nerve fibers cross each other. Their potential is transmitted through the *Medulla Oblongata*, *Pons,* and *Mesencephalon* to the *Nucleus Ventralis Posterolateralis* (a part of the *Thalamus*), which transfers the signals through *Tractus Thalamocorticalis*, *Capsula Interna*, and *Corona Radiate* to the *Gyrus Postcentralis* (Bähr & Frotscher, 2009). Here, perceptional parts are represented in the *Primary Somatosensory Area*, which can be divided into *Brodmann* areas 1, 2, and 3 (Birbaumer & Schmidt, 1996). There are also some sensory fibers which lead from the *Thalamus* to the *Primary Motoric Cortex*; these areas sometimes overlap in their function (Bähr & Frotscher, 2009). When the brain transfers a signal towards executive organs, the *Gyrus Praecentralis* or *Primary Motor Cortex* sends it via the *Tractus Corticospinalis* through *Capsula Interna* and *Brainstem*, where most of left and right side cross each other, through the *Mesencephalon*, *Pons* and *Medulla Oblongata* into the *Spinal Cord*. The parts where left and right side cross each other form the *Tracticus Corticospinalis Lateralis*; the ones which remain on their original side the *Tractus Corticospinalis Anterior* (Bähr &

Frotscher, 2009). From the *Spinal Cord*, the signals are transferred to the *Motoric End Neurons*, which operate the muscles towards a certain action, e.g. activating a switch.

In experiments, the tactile reaction time is noted as 155ms (Robinson (1934), cited after Kosinski (2012)). Haptic and tactile reaction times in the automotive context may vary between 100 and 250 ms (Krüger et al., 2009), but these values do not refer to switches but to steering, a task which is way more complex than e.g. a mere muscular reflex. Concluding, the velocity is highly dependent on the type of perception cell and cannot be generalized for tactile or haptic perception.

#### *2.1.3.4 Confounding factors in tactile perception*

The sensitivity and accuracy of tactile perception varies between subjects, but there are also confounding factors which influence it. When designing for tactile interaction, they have to be taken into account to avoid bad usability design. The most prominent of these confounding factors are described in the following paragraphs.

First of all, the temperature of the areas which contact the interaction object is an important issue. Especially the channels processing information from the *Pacini-corpuscles* and *Ruffini-corpuscles* have the property of being highly sensitive to temperature changes (Bolanowski et al., 1988), whereas there seems to be a dependency on the frequency: from 125 Hz on, *Mechanoreceptor* cells in the cutaneous skin start to become sensitive towards temperature changes in their perception as well (Hatzfeld, 2013).

Second, there are also age effects. Generally, the perception abilities decrease for all modalities with age, whereas the ability to aim attention at a specific object deteriorates (Milham et al., 2002). Although some authors mention that the age effects for surface perception do not have an impact as high as other modalities (cp. Grane, 2012), especially the vibration perception (via *Pacini-corpuscles*) seems to be prone to age effects (Gescheider, Bolanowski, Hall, Hoffman, & Verrillo, 1994). The authors argue that there is not a change in the high level information processing besides of the changes in the skin and the *Pacini-Corpuscles*, which indicates that this is rather a perceptional than a cognitive issue. Additionally, the ability to discriminate different stimuli decreases on a perceptional level, which is also an indicator of loss of tactile sensitivity (Dinse, 2008).

Thirdly, there are different masking effects. Masking means that an induced tactile stimulus is overlapped by another tactile stimulus which origins in environmental noise. One will be able to find the occurrence of spatial, cognitive (Craig, 1974), and temporal (Gilson, 1969; Hoggan, Crossan, Brewster, & Kaaresoja, 2009; Verrillo, Gescheider, Calman, & Van Doren, 1983) masking effects. Spatial masking occurs through a stimulus in the same or a nearby location of the original stimulus, cognitive masking appears due to transfers of a signal in cognitive information processing, and temporal masking means an overlap in time between an original and a masking stimulus. *Pacini-Corpuscles* are especially prone to suffer from masking effects due to vibration noise, as they are primary vibration perception cells. For the automotive context, this means that the transfer from the lab environment to a more realistic setting like on a test track, needs to be designed carefully and several parameters need to be amplified when taking the testing from a driving simulator to a test track, as on a test track, an increase in vibrational noise is highly likely, e.g. due to road vibrations or in-car rattle and vibration.

Fourthly, the area of the body on which the vibration is induced influences the perception accuracy and differentiability, resulting in different required intensities and frequencies of a stimulus for each area (Burdea, 1996). This is due to the variation of the density of perception cells over different parts of the body. Another factor which influences intensity and frequency perception of a tactile signal is the contactor size; the larger the contactor, the smaller are absolute and relative perception (Verrillo, 1963; Verrillo & Chamberlain, 1972; Jones & Sarter, 2008).

All the aforementioned factors may influence the specific perception thresholds, so the definitions of these thresholds are to a certain degree dependent on the specific settings. Literature based orientation values will be discussed in the next chapter, but the application always requires separate evaluation.

### *2.2 Perception thresholds*

The tactile perception thresholds are highly dependent on the influence factors mentioned above, such as temperature, age, environmental vibration noise, contact area and the size of that area. Under these circumstances, the determination of an "absolute" perception threshold, as performed by Sherrick & Craig (1982) for force perception, in which they determinated 1.4715 mN for the *Thenar* (muscles at the base of the thumb) and 0.7848 mN for the fingertips, is open to discussion (Sherrick & Craig, 1982), but provides a good indicator for values below which it makes absolutely no sense to provide tactile feedback. Other authors have gathered a similar force perception threshold for the fingertip, 0.008 N (induced via a nylon monofilament (Lederman & Klatzky, 1998)).

The general temporal resolving capacity of the human skin for detecting a temporal gap between two stimuli in lab conditions is around ~5 ms (Lederman & Klatzky, 1998). In the field this value can be much higher. The pressure perception at the palm of a touch stimulus requires a displacement of ~40 µm (measured from absolute contact point with the skin).

Another approach in determining the perception thresholds is via the *Just Notable Differences* (JND) or the Weber-quotient. This means a factor by which feedback has to be increased or decreased to be perceived as a different kind of feedback and constitutes a relative instead of an absolute threshold. There is a broad discussion in the literature about the JND for discrimination values of tactile vibration perception. Some examined an exponential function from 21% at 0.5 N to 5.5% at 2.5 N (Doerrer & Werthschuetzky, 2002), in some experiments occurred a constant JND at 15% (Burdea, 1996; Tan, Srinivasan, Eberman, & Cheng, 1994), and some experienced values between 5% and 10% from 2.5 to 10 N (Tan, Pang & Durlach, 1992). In the present technology force differences can be modulated in increments of 0.1 N, which gives far more options than humanly perceivable.

Although very similar from an interaction perspective, other force-applying types of interaction differ in their JNDs from the values mentioned above. When investigating different force settings in regard to the force feedback given, at a minimum 10% and at a maximum 20% should be considered in an experimental design. The differential threshold for kinesthetic force perception averages 7-10% over a force range of 0.5-200 N, and appears to be consistent across a wide variety of muscle groups (Jones, 2000). When subjects are asked to apply a certain force to a mechanical switch, the error rate of force application is below 5% which shows that for mechanical switches, the accuracy of force application is quite high regarding kinesthetic force (Zeilinger, 2005). A direct transition of force thresholds from mechanical switches to vibrotactile force-sensitive switches is not possible, as this transition takes also place between different perception cells (from *kinesthetic* to *tactile*).

In terms of frequency for vibrotactile feedback, the recommendations differ highly, which may be caused by different amplifiers and contactor sizes used. Jones & Sarter (2008) recommend a design around 150-300 Hz, around the optimal sensitivity of *Pacini-corpuscles*. Other recommendations include frequencies at 100, 125 and 160 Hz (Brammer & Piercy, 1998) or 60 and 130 Hz, with lowest user preferences for the frequency value being at 200 Hz (Dabic, Navarro, Tissot, & Versace, 2013). For differentiation thresholds, the JNDs seem to vary depending on the frequency range. The JND increased from around 18% at 20 Hz to approximately 30% at 300 Hz in one study (Rothenberg, Verrillo, Zahorian, Brachman, & Bolanowski, 1977), whereas Mahns, Perkins, Sahai, Robinson, & Rowe (2006) reported, that it decreased from 30% at 20 Hz to 13% at 200 Hz. This indicates that Weber's Law does not apply for frequency discrimination (Jones & Sarter, 2008), but the direction of the change seems also to be dependent on other factors. Some interaction effects between frequency and amplitude of the signal occur as well. As the amplitude of vibration increases (at a constant frequency), the perceived frequency of the signal also increases, although the rate at which perceived frequency changes varies considerably across people (Jones & Sarter, 2008; Morley & Rowe, 1990; von Békésy, 1959). As a result, amplitudes cannot be evaluated without controlling for frequency and their common interaction effect.

The temporal resolution for vibrotactile feedback depends on the perception modality. Mechanical passive tactile perception can distinguish between two stimuli in a time as short as 5 ms, whereas for vibration stimulation, it takes 50 ms between two stimuli to be perceived as two separate stimuli and not as one stimulus (Jones & Lederman, 2006). This is important especially for skip or push-and-hold functions, where at more than one point vibrational feedback is provided, or multiple feedback cue design.

In conclusion there are different possible thresholds and JNDs which can be applied to the design of auditory and tactile feedback. Several possibilities have been discussed, although due to confounding variables and technology differences, there are always additional experiments with subjects necessary to make sure that a differentiable feedback set is evaluated. In which way these designed feedbacks may correspond to psychophysiological measurements of auditory and tactile perception, especially in terms of which components of a signal can be translated into which measurements of the different perception modalities, will be discussed in the next chapter.

Visual measurements are not discussed, as in the experimental settings, the visually perceivable surface of the device will be kept constant and visual design will not account for a major part of the present thesis. Eyetracking will be included only as a measure of distraction in the main experiment.

### *2.3 Psychophysiological measurement of auditory perception*

The most common measurements for human sound perception is the auditory field as a function of sound pressure level or sound intensity level over frequency, loudness, sharpness, and roughness (Fastl, 1997). Pitch and masking of auditory signals are additional topics which have to be considered.

The sound intensity level (SIL) describes rather the original sound at the source, but not the one which is perceived, and is therefore not an optimal measure for subjective perception. It is a logarithmic measure of sound intensity in comparison to a reference level and commonly described in dB.

The sound pressure level is the effective sound relative to frequency dB (SPL). As the frequency response (perceived loudness) changes with amplitude, there are different filtering indexes: A, B, and C (NA 001-01-03 GA Gemeinschaftsarbeitsausschuss NALS/DKE: Schallmessgeräte, 2003), whereas dB (A) is the most common measure. The drawback in using these filters is that there is no possibility of converting dB (A) in dB (SPL). Additionally, the A-filtering does work for low levels and small bandwidth sounds only (Fastl, 2006a). As most tones in nature are not pure and have a broader bandwidth, the applicability of this filtering to human loudness perception is debatable.

The loudness level (measured in *phon*) is the level of pure tones with the same loudness as function over frequency (Fastl, 1997). One *phon* is, at a reference frequency of 1 kHz, equal to the dB (SPL) scale (Raichel, 2006). For soft tones, the loudness level corresponds to A-filtering, but the latter systematically underestimates loudness for higher levels and broadband sounds compared to the *phon* measure (Fastl, 1997). Another way to measure loudness is the Bark-scale, which corresponds to frequency bundles in the human ear and is therefore closer to human perception. Bark can be calculated from Hz, but the inversion is only possible between 200-6400 Hz or 2-20.1 Bark (Hartmann, 1997). Perceived loudness is measured in *sone*, which has the advantage, that it is a linear measure and takes frequency, *phon*-values and masking into account (Fastl & Zwicker, 2007).

Sharpness of a tone refers to an attribute of the timbre (Fastl, 1997) and depends on the frequency combination of a tone. For high sharpness high frequencies are needed, and the bandwidth influences sharpness as well. If low frequency noise is added, the sharpness of a tone decreases (Fastl & Zwicker, 2007), which seems to enhance the subjectively perceived quality (Fastl, 1997). Sharpness is measured in *acum*.

Roughness is less dependent on sound pressure level, but more on carrier frequency, modulation frequency and degree of modulation. Its unit is *asper* (Müller & Möser, 2003). The higher both modulation measures, the rougher a sound is perceived (Fastl, 2006). At a modulation frequency up to 20 Hz, roughness can also be termed as fluctuation strength, which is basically the same: variation in modulation frequency and degree (Fastl, 1997).

Pitch can be distinguished in pitch strength and pitch height (Fastl, 2006). It depends on the frequency, but is not exactly the same, because it depends also on sound pressure and waveform (Hartmann, 1997). Pitch is measured in Hz, but does not necessarily correlate linearly with the frequency of a tone. There are high interindividual differences in the perception ability of pitches (Hartmann, 1997), which make it difficult to draw clear conclusions from research. Human auditory perception, on the other hand, is not a pure measurement instrument for one single quality of an auditory signal but a rather complex mechanism of integrating several qualities of a signal, which might explain the difficulties of the psychoacoustic differentiability of sound.

Masking of auditory signals can occur simultaneously in time or shortly before or after the stimuli. The location of the sound and the frequency range can also be subject to masking. Masking effects can occur additionally after the masking sound is switched off (Fastl, 1997), which is called post-masking.

In summary, assessing the perceived quality of a sound involves beside the *sone* measures also characteristics of sharpness, roughness and the pitch. In an automotive environment, the effects of masking have also to be taken into account, such as engine, wheel, or driving noises. As those parameters change with velocity, it is important to design high quality sounds which are adaptive to the environment or at least to the velocity driven, as the velocity directly influences road noise.

### *2.4 Psychophysiological measurement of tactile perception*

As there are a variety of tactile perception cells, there is also a huge amount of psychophysiological measurements for tactile perception.

Force as a kinesthetic factor (e.g. activating a switch) is measured in Newton (N) and is a product of mass and length divided by the squared time. Hand grip force is also measured in Newton. For pushing a mechanical switch this is the most important measure. This genuine haptic interaction is now substituted through virtual switches, which leads to the consideration how to measure the kind of feedback given via tactile feedback and how to measure the specific characteristics of this feedback when it is substituting for a genuine *kinesthetic* interaction. The activation and deactivation force guidelines may still apply; but the tactile feedback given needs special consideration. In other words, the applied force still needs to be taken into account when designing for tactile feedback, although the feedback is merely an electric signal. In the following, different approaches of calculating vibrational feedback parameters are discussed. The subsequent formulas are described only in their reference to sine waves, as those are the commonly recommended waveform (cp. chapter *2.1.3.1 Surface* perception) for tactile feedback signals.

The DIN EN ISO 16352:2005 states three parameters for coding a tactile signal: intensity, duration and location. Location is the part of the body at which the feedback is presented. In the present case this corresponds to the fingertips. Duration is the time span on which the feedback is given, usually expressed in milliseconds. There are different recommendations and implications for the intensity of vibrational feedback. The most important influence factors are amplitude and frequency.

The rate of change of acceleration (*a*; sometimes referred to as *jerk*) in haptic perception research is sometimes expressed in g/s and describes rather the change rate of acceleration of gravity over time. A positive jerk value is increasing, a negative decreasing force on the body. This is not easily translated into human perception, as a reference value is missing. Additionally, this is not transferable to non-kinesthetic feedback processes.

Frequency (in Hz) is considered the most important, but not the only feature which influences the perception of tactile feedback (Jones & Sarter, 2008). The optimal sensitivity of Pacini-corpuscles ranges between 150 and 250 Hz (Jones & Sarter, 2008). There are hints that this most precise sensation range is not optimal in terms of user preference, especially when it comes to 200 Hz (Dabic et al., 2013). Other authors recommend using values between 100 and 160 Hz (Brammer & Piercy, 1998). Frequency perception is dependent on the amplitude of the signal, which is harder to determine.

A common measure for vibration intensity or amplitude is decibel (dB), but unfortunately this is a rather simple logarithmic unit which describes the ratio between two values of a specific physical quantity. It can be used for any physical quantity and is therefore not specific enough for a determined measurement.

When the focus is on the energy content of vibration over time, it is often referred to as Root Mean Square (RMS., sometimes also referred to as quadratic mean) of vibration. It can be calculated via the formula

$$a = \sqrt{\sum_{n=1}^{i} (K_i \times a_i)^{\,2}}$$

In this context, $a$ means the total acceleration in m/s², $K_i$ the weight of vibration at specified frequency, and $a_i$ the acceleration at specific frequency in m/s². This procedure is standardized in the ISO 2631-1:2000 (Park, Fukuda, Kim, & Maeda, 2013). This method was developed to examine to which vibrations a human body may be exposed without putting the human into danger, so it may be too raw as a measurement for simple switch interactions, where the vibrational amplitudes are relatively small.

A more refined value for the vibrational impact on humans is the vibration dose value (VDV), which shows the total amount of vibration received by the human over a period of time, considering the magnitude, frequency and exposure duration:

$$VDV = \left( \int_0^T a(t^4) \times dt \right)^{\frac{1}{4}}$$

Here, $a$ means the frequency-weighted acceleration, and t the total period in seconds that the vibration occurred. If the vibration induced lasts for longer periods of time, the VDV should be calculated in order to determine whether an impact might be an overdose for the human operator.

In virtual switch technology, the signals are rather short and therefore no harm should be induced when the peak value is within a normal range. References may therefore not be applicable, similar to the RMS pain measure mentioned prior to the VDV.

Another factor which describes the extremity of vibrational peaks is the crest factor, which is also taking the RMS into account.

$$C = \frac{|x|_{peak}}{x_{rms}}.$$

To express a relationship, some authors deem it more useful to utilize the peak-to-average power ratio, which corresponds to the crest factor formula but squares both peak amplitude and root mean square (RMS) in order to give power to the value. This approach is not very commonly used in literature, which makes the comparability of results difficult.

Displacement is, opposed to all measures above, a frequently used measure in the area of psychometric evaluation of virtual switches. It is sometimes defined in mm or $\mu m$, but this term describes only the length in distance (or travel) of the vibration, whereas the ratio between intensity and position, which is also a more refined measure for displacement, is expressed in dB in reference to 1 $\mu m$. As this measure describes the intensity as a ratio between values in reference to 1 $\mu m$, it has the advantage that the intensity of the signal as well as its actual impact on the skin is contained in this measure (see for an example Gescheider, Bolanowski, Pope, & Verrillo (2002)). Other common measures for vibrational displacement are the vibrational velocity and the corresponding change rate (acceleration in $m/s^2$), but those values do not take actual peak height into account and are therefore less preferable, whereas displacement in reference to 1 µm actually takes these measures into account:

Displacement referred to 1 $\mu m$ can be calculated from the amplitude of the signal, its frequency and its acceleration. There are different approaches for amplitude determination; the amplitude can be calculated either as the highest peak from the zero axis, the difference between highest positive and lowest negative peak of the signal (peak-to-peak amplitude, cp. Craig, 1974) or as an average value consisting of the different peaks of the signal. As some peak values are negative in a sine curve, the RMS or quadratic mean is another possibility to describe feedback intensity. The latter has, like peak-to-peak amplitude, the benefit of giving more impact to the highest peak, which will influence intensity perception of a signal the most.[3]

Summarizing this one can see, that in order to determine the characteristics of a vibrational tactile feedback feedback, frequency, acceleration, and velocity need to be measured and calculated. The displacement in reference to 1µm is a common measure which integrates these values into one reference metric.

Whereas the underlying mechanisms of perception in different modalities have been outlined above, perception is one crucial, but not the only component of human-machine-interaction. Cognition and the response reactions need to be taken into account as well. The next chapter deals with interaction principles and assumed theories in order to take these aspects of an interaction into account as well.

## *2.5 Interaction principles*

### 2.5.1 Visual interaction

Generally, visual interaction is normally one-directional and mainly contains only perception in most common interactions. There are different gaze interaction possibilities in the other direction, for example for people who cannot use their hands for the interaction with a device. In these cases special technology is needed. The only natural response possible to a visual cue is to keep it in the focus of attention. Visual interaction is therefore very limited in one way, but on the other hand one of the most selective and deliberately controllable of all perception channels.

There are different mechanisms which allow perceiving depth and size of objects, as well as their relation to each other. Some of them work monocular- like masking, relative height and size, perspective convergence, familiar size, aerial perspective, texture, kinetic depth effect (via shadows), and movement induced depth perception. Other mechanisms require both eyes to participate, such as stereopsis and convergence. Those mechanisms will not be discussed in detail here, for an overview of the underlying mechanisms see Goldstein (2008).

[3] Note that although the input signal is held at a constant value (e.g. pure sine wave), the perceivable signal of it (output signal) decreases over time. When using the RMS or other measures to calculate the displacement in reference to 1µm, this refers to the output acceleration data of an exciter and not the original input signal.

Color perception, which is usually a major concern in interface design, is not discussed here as the device in question uses a strictly black-and-white surface and the comparison with other steering wheel switch technologies will be conducted with similarly designed surfaces, in order to ensure the comparability of the different devices.

For this thesis, visual interaction will be discussed as a unidirectional channel, as the user group for the evaluated device consists of healthy persons without requirements for gaze interaction. Its main role in the present work lies in the distraction potential of the interaction itself with the switches. This will be discussed in the section *Eyetracking data*

### 2.5.2 Auditory interaction

Auditory interaction is determined unidirectional, too, as a response in the auditory channel is performed either by a subject providing a verbal or a motoric reaction. Verbal output is in some ways a mechanical, muscular interaction as well, although voice interaction is sometimes discussed as direct auditory interaction. For this thesis, auditory interaction will be discussed as a unidirectional channel. The response reaction to an auditory cue can be unintentional, as it is the case with other primary attention reactions (for example described in Birbaumer & Schmidt, 1996).

Localization of an auditory cue can work, similar to the organ of vision, via both ears (*binaural*) or with one ear (*monaural*). Binaural localization can take place either over the time difference a cue reaches both left and right ear (*Interaural Time Difference*) or via the sound pressure difference (*Interaural Level Difference*) of a cue reaching both left and right ear. The *Interaural Level Difference* does appear on high frequencies only, as the pressure is modulated by the head position of a person. The ear on the far side of the head is in an auditory shadow and perceives different sound pressure levels. *Monaural* position cues are perceived via cue position in space related to the head (head related transfer function). Basically frequency patterns are reflected differently in regard to their position in space by the *Auricle* and in this way give information about the position of the cue (Goldstein, 2008).

### 2.5.3 Tactile interaction

Tactile interaction is generally bidirectional; the efferent and afferent fibers enable interaction in both directions. Motoric reflexes work without intentional initiation of a response; tactile interaction can therefore be unintentional as well.

The sensation of touch can also be differentiated into active and passive touch. With active touch, observers freely explore their environment, receiving sensory input from both *cutaneous* and *kinesthetic* systems. With passive touch, the observer remains stationary with contact produced by an external agent (e.g., the experimenter). Purely passive touch offers only cutaneous information (Lederman & Jones, 2011). Active touch involves in nature normally some sort of passive touch.

Passive touch relies mainly on perception mechanisms and does work only in one direction. In to that intentional response is rather complex and requires muscle activation mainly. This active touch is important for object identification and involves these intentional responses. Object identification incorporates the interplay between different *tactile* and *kinesthetic* perceptions, intentional motoric responses and the cognitive integration (Goldstein, 2008). This makes haptic interaction a very complex procedure: iterative feedback loops are not modulated by one perceptional channel, but the interplay of different perception cells and modalities. And the small feedback loops are so rapidly changing that standardization in a more realistic setting is not easily accomplished. Another classification for touch modalities was proposed by Loomis & Lederman (1986), which distinguishes between control and no control. Control refers to active controlled movements, whereas no control means that an outside person (e.g. an experimenter) moves or fixes the hand of a subject. This would not differ as much from the active/passive touch definition originally discovered by Gibson (1962), if the authors did not add a second classification referring to the involved sensory systems, *kinesthetic* and *tactile* (*cutaneous*) with various sub classifications. They subsume *cutaneous* information, *afferent kinesthesis*, and *cutaneous information* and its interaction with *afferent kinesthesis* under no control and classify it respectively as *tactile perception*, *passive kinesthetic perception*, and *passive haptic perception*. *Afferent kinesthesis* together with *efference copy* constitutes by this definition *active kinesthetic* perception, and *cutaneous* information with afferent *kinesthesis* and *efference copy* as *active haptic perception*. The interesting implication from this classification is that that there is no isolation for purely *Pacinian* stimulated perception when pressing a virtual switch, although there is barely a *kinesthetic* movement. Providing tactile feedback only would therefore be in this classification a passive touch. As user intentions and active operations, such as force application, are involved, the interaction with a virtual switch is still more than *passive kinesthetic perception*.

In order to investigate different settings of stimuli related to vibration perception, the control of factors such as material or surface is necessary to avoid interactions as much as possible. By keeping those factors preferably constant, the perception modality of interest can be isolated in an imperfect, but sufficient way for thorough investigation of an automotive application.

Whereas even the basic perception of tactile and kinesthetic cues is complex, the interplay between the different haptic perception modalities increases this complexity, and the interaction with other perception modalities such as visual and auditory exceeds this in complexity even further. The latter topic will be discussed in the next chapter.

## *2.6 Multimodality and its applications*

In the automotive context, different approaches in the research on the integration of multimodal functions in the driving context have been taken. Two main categories of comfort and safety functions, such as in-vehicle-infotainment systems (IVIS) and advanced driver assistance systems (ADAS), evolved from this. The following paragraphs are structured to give first an overview on multimodal resource theories and attention research in the multimodal domain first, to then diversify between the two main applications. It is noteworthy that there seems to be a trend to blur the line between ADAS and IVIS due to new developments in terms of personal assistance systems for drivers, which are aimed at detecting the driver state, adjust car settings accordingly and store and choose infotainment settings for the driver based on his state. The reasoning follows the line that there is no possibility anymore to distinguish between IVIS and ADAS, when safety and comfort functions are mixed in these systems. For feedback and interaction, however, different design implications are applicable depending on criticality and urgency of information presented. The terms are used strictly in the following paragraphs in their infotainment respective warning functionality and do not refer to system design and system boundaries.

### 2.6.1 Theories on attention and multimodal resources

There have been a lot of discussions about the nature of attention processes, especially between two researcher groups. One approach tries to explain attention via passive, object related characteristics, such as colour, salience, contrast, and the environment, which is called the bottom-up theory. The other approach stresses the internal factors of an individual such as tasks, goals, motivation, as well as spatial and feature characteristics important to the individual, which is called

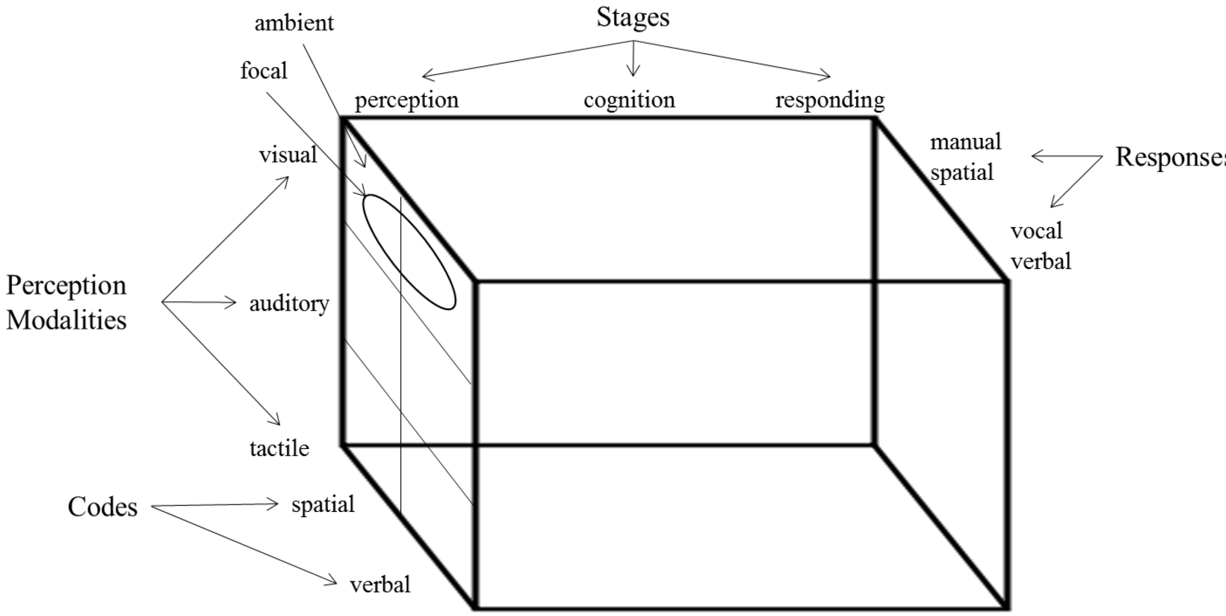



*Figure 4. Wickens' Multiple Resource Model. Adapted from Wickens (2008).*

the top-down approach. There is no pure bottom-up or top-down approach and most modern models assume a mixture of both procedures (Chun & Wolfe, 2000).

In the beginning of attention research, attention was viewed as a sequential process. The debate in research was focused on at which point in the perception process attention filters take place. Due to studies which showed that attention is not one single process, this view is outdated (Chun & Wolfe, 2000). As early as in 1972, Allport, Antonis, & Reynolds (1972) proved that attention is no single channel functionality, but occurs in several channels, corresponding to the different perception modalities (visual, auditory, haptic, olfactory and gustatory). Different models were developed in the following years, which intended to integrate multiple perception and response options in respect to their cognitive processing demands.

The Multiple Resource Model of Wickens (Wickens, 2002, 2007; 2008; Wickens & Hollands, 1999) is based on this assumption. This model is an approach to information processing, dealing with the perception modalities, processing stages, responses and memory codes (cp. Figure 4. *Wickens' Multiple Resource Model*).

The processing stages differentiate in perception, cognition and producing a response reaction. It is assumed that each perception modality (in this case visual, auditory and tactile) is processed differently. Nested in the visual perception modality are the visual channels, which are divided into focal and ambient, corresponding to the perception modes. The main task of the first channel is the focal object recognition, while the ambient channel is responsible for the perception of movement and spatial orientation. Tactile perception could be differentiated into the modes discussed in the chapter *Haptic and tactile perception*. The different codes refer to spatial and verbal information storage. The spatial memory is responsible for spatial orientation, colour and shape information, whereas the verbal memory refers to abstract constructs and language. The response reaction can be verbal or motoric - a mimic expression would be defined as a motoric response reaction in this model as well.

In the *Multiple Resource Model*, it is obvious that there are different sorts of attention. Any task and its associated workload can be represented by a specific combination. When two different tasks are carried out simultaneously, task performance depends on which resources are used. When different resources are used for each task only, both tasks can maybe be processed in parallel (e.g. simultaneously see some sort of information and speak). As soon as two or more tasks claim partly or completely the same resources, there will be trade-offs in task processing and the resources available to the different tasks are split (e.g. monitoring the road and, at the same time, perceiving and processing information on a navigation device). This model is in each case a simplifying model and can be used as a rule of thumb only (Spence & Ho, 2008). It can only give hints at the design of a human machine interface in terms of prediction of resource demands, overlapping and division between concurrent tasks. For the detailed evaluation of a user interface and its impact on task performance, one or more studies with human subjects will always be needed.

As an alternative to the *Multiple Resource Model*, McCracken & Aldrich (1984) classify workload in a way which looks similar at first. Differences include that they do not differentiate input and output, but just general workload, and do not differentiate between verbal and manual output. They only distinguish between motoric, visual and cognitive demands. It has to be kept in mind that this model was developed to classify pilot’s tasks via cognitive task analysis and did not focus on which channels can be used in an interaction like the approach of the Wicken’s cube. Despite its little granularity, this model is used quite frequently in research. The *Multiple Resource Model* seems to be more adequate for the present thesis due to its higher level of detail and the assumption of the possibility of performing several tasks in parallel. This is of course only reliable when different modalities are used and the main cognitive effort is assigned to the main task. The following paragraphs deal with attention mechanisms of the different perception modalities separately.

For foveal vision it is assumed that the attention is focused on the point at which it is looked at in that moment, the so called eye-mind-assumption (Carpenter, Just, & Rayner, 1983). This excludes phenomena as daydreaming or, in the automotive context, mind wandering behind the wheel (He, 2010). As the visual perception and visual information processing underlies certain biases, there are also some attention mechanisms. When distracted, people may not notice changes in settings which are identical in all depicted things except one (Simons & Levin, 1997), the so called change blindness phenomenon. There are also other possible influence factors and underlying mechanisms influencing change blindness, which will not be discussed here in detail; for an overview see Simons (2000).

It is also possible to deliberately direct attention in the auditory modality. A prominent example has been the *Cocktailparty-phenomenon*: in the presence of more than one auditory source, the human auditory perception is able to focus on one source, like listening to one person during a party, where a lot of people are talking (Cherry, 1953). However, this phenomenon in the original study as well as in a replica has only proofed to be true for one third of the participants (Wood & Cowan, 1995) and suggests that there are some attenuation filters which filter irrelevant information, when subjects have little attention resources. This would also be in line with Treisman's attenuation theory (Treisman, 1960, 1964; cited after Wood & Cowan, 1995). Filter performance depends on individual abilities as well to block out information (Conway, Cowan, & Bunting, 2001).

Although tactile attention focuses on stimuli which have direct impact onto the body, there seems to be an attention selection mechanism corresponding to visual and auditory attention (Müller & Giabbiconi, 2008). Corresponding to visual change blindness, there is also a kind of tactile change blindness: when distracted, up to 30% of participants in a study did not notice a change in the presented vibratory patterns (Gallace, Tan, & Spence, 2006). Given the fact that the subjects in this study were rather young and change blindness seems to increase with age, an impact of attention in the tactile modality seems rather likely. In a study with tactile feedback in a secondary task, it is possible that subjects are distracted by the primary task to an extent that they do not recognize changes and therefore encounter difficulties in judging different vibrotactile patterns.

The interaction between the different modalities and possible dominance effects will be subject to further discussion in the following. The devices on which this interaction takes place are of special interest here, as the context influences the subjective perception. There are two main application areas in the automotive context on which multimodal interaction usually does take place: warning functions and comfort functions, which will be discussed subsequently to the multimodal integration subchapter.

### 2.6.2 Multimodal integration

There are some mechanisms which enable the illusion of a feedback in one modality through the elicitation in another modality. Whereas this is sometimes referred to as Synaesthesia, this term has mostly been used in the interaction between auditory and visual stimuli (Haverkamp, 2012). In the aforementioned case, haptic feedback is considered only in terms of surface characteristics and resulting consequences for motoric actions, which are not in the focus of the present work.

In terms of tactile vibrational feedback, the term pseudohaptic feedback is used by some researchers (e.g. Lécuyer, 2007). The goal of this kind of feedback is to simulate surface characteristics via vibrational feedback and visual illusions, where subjects judge their proprioception via a visual feedback category. The visual feedback in general seems to override the tactile perception, which is not surprising when keeping in mind that the visual modality seems to be the utmost dominant one in human perception (cp. Posner & Nissen, 1976). For the visual feedback it is also possible to override either contrary auditory or tactile feedback. Tactile feedback or auditory feedback in combination do not have this effect on each other, whereas the combination of both against a visual stimulus erases the visual dominance (Hecht & Reiner, 2008). Even when subjects are instructed to focus on the other modality, vision still prevails, which seems to imply that there are more influencing factors than an exclusive attentional process (Spence, Parise, & Chen, 2012). In current research, there is no explanation for this phenomenon. For haptic or kinesthetic perception, not so distinct results are existing. The visual domain remains dominant, but seems to have larger effects on geometry than on material (Lederman & Klatzky, 2009). There are hints for the auditory modality to override the tactile (Altinsoy, 2008). The seamlessness of the auditory and tactile feedback combination must therefore be ensured for a safe and pleasant interaction. To be optimized for human perception and to enrich the multimodal interaction, multimodal stimuli should therefore rather be congruent than incongruent, and the congruency may even improve the interaction opposed deteriorating it when providing the subject with conflicting stimuli.

### 2.6.3 Multimodal warning strategies

The most common application of multimodality in the automotive context is the design of warning functions. The combination of different warning strategies takes into account that there are different processing resources for different perception modalities, as proposed by the Multiple Resource Model (see Figure 4. *Multiple Resource Model*). The most commonly combined modalities used in warning strategies are visual and auditory warning elements (see Sarter, 2006) but there have been also studies conducted on tactile interfaces as a warning. Theoretically an olfactory warning function (see Sarter, 2006) could be possible as well. Except for visual displays and verbal auditory warnings it is hard to transfer complex information (Fricke, 2009). An auditory tone warning without any semantic implication for the occasion may lead to orientational reactions, consuming irrevocably important milliseconds of a driver's reaction time. Through multimodal interfaces, various aspects of a situation can be presented, redundancies are created and a greater amount of information can be transferred (Oviatt, 2003). The focus of multimodal warning strategies is mostly on people in situations with high visual demand, often in safety-critical situations, which need support for attention allocation (cp. e.g. Latorella, 1999; Brickman, Hettinger, & Haas, 2000; Nikolic & Sarter, 2001; Spence & Ho, 2008). Driving a car is a highly visually demanding activity. Thus it is an area which is predestined for a multimodal warning strategy using additional perception channels.

Another important aspect in the design of multimodal warning strategies is the task-appropriate choice of location for the warning modalities and their adaptation to the sensory modalities. Following this it makes sense to display relatively complex information in the direct line of sight of the driver, while in the peripheral visual field movements, brightness changes and the appearance of new objects is perceived particularly well, whereas object and detail recognition is not supported (Sarter, 2006). While reducing distraction is already important in the design of a unimodal warning, it becomes even more important with an increasing number of combined modalities. The user should never become irritated by a higher number of warnings or, in the worst case, irritated to a point where the warning distracts and therefore harms rather that producing a benefit. The DIN EN ISO 14915, Part 3: 2002 provides a clear set of design principles by postulating that combinations of modalities should have a thematic congruence, a manageable information load, and mutually supplementing focuses, to support consistency and redundancy, but this is a standard which is concerned rather with software than with driving.

User acceptance also needs to be taken into account. Chang, Hwang, & Ji (2011) used a tactile warning device in the driver seat during a simulator study which performed better than auditory or visual warnings, because of the unfamiliarity with this kind of warning, their test subjects were not satisfied with it and user acceptance was not very high.

It can be summed up that multimodal warning strategies are used due to the fact that the saliency of these cues is increased, a disability in one modality (e.g. impaired hearing) is compensated and reaction times are improved in some cases. Additionally, some semantic information can be transmitted easier through one modality than through another.

#### 2.6.4 Multimodality in comfort functions

The research of comfort functions such as making a phone call (a typical IVIS task under investigation), shows that the driver distraction does not necessarily decrease with multimodal optimization or integration according to the *Multiple Resource Model.*

Shifting traditionally visual perception and interaction to another modality (e.g. a visual-motoric task to a voice interaction) seems to not always reduce the overall workload for a driver, but to come at a certain cost (Vollrath & Totzke, 2003). A shift in the response modality does also not necessarily improve the distraction of a driver: the motoric response task of dialing a number or holding a cellphone with one hand while talking is not the only factor influencing distraction. The high cognitive demand of talking on a cellphone has a rather crucial effect on driver attention and distraction (Spence & Ho, 2008). Similar patterns of eye glance behavior as a measure of attention shifting could be found between hand-held and hands-free devices by Harbluk, Noy, & Eizenman (2002). There are also influences of task difficulty (Metz & Krüger, 2011). For this reason it is not legitimate to conclude that a shift from visual modality to another will always result in a reduction of driver distraction. Research on comfort functions should therefore always investigate their distraction potential.

The studies cited above are conducted in the context of driving without assistance functions. Additional attention demands to a manual driving task are always a distraction in this situation. Complex distraction should be avoided as much as possible and every new device needs to be evaluated carefully in terms of its distraction potential. More complex comfort functions might be possible in the context of higher automation in passenger cars, and may be even necessary to keep the driver in the loop or prevent him from falling asleep. The vigilance of the driver must be maintained and the system state information must be easy to access. There are multiple design restrictions which may not apply in their current state, when automated driving has become the most frequent way of transportation for passenger cars. The focus of this thesis, however, is on the development of virtual multimodal switches for existing systems and those which will be introduced in the near future, so automation issues and their implications for multimodal feedback design are not discussed here.

The detailed design of multimodal comfort functions is dependent on the inner and outer environment of the vehicle and the urgency of a message (Cao, Van Der Sluis, Theune, & Nijholt, 2010). Vibration has a higher appropriateness when there is a lot of auditory noise, and there is less urgency to the message. Sound comes into play when a message is highly urgent and there is already road vibration, e.g. due to driving on cobblestones (Cao, Van Der Sluis, Theune, & Nijholt, 2010). Sensory concepts, which take into account the effects of masking and adapt the vibrational and auditory devices accordingly, are therefore necessary. An example for a vibrotactile comfort function would be the prototype developed by Hwang & Ryu (2010), which was built up as a steering wheel with 32 actuators providing directional navigation information via tactile cues. The addition of tactile feedback in IVIS systems on a touchscreen seems to reduce subjective task difficulty, error rates and input times (Pitts, Burnett, Williams & Wellings, 2010; Pitts, Skrypchuk, Wellings, Attridge, & Williams, 2012; Richter, Ecker, Deisler, & Butz, 2010), which corresponds to generally lowered error rates when tactile feedback is added (Altinsoy, 2008).

It can be argued that in comfort functions urgency should never be the matter as they are secondary or tertiary tasks, depending on which definition is taken into account. The microinteraction of pushing a switch should nevertheless be optimized regardless of which function is employed. In terms of safety-criticality, Grane & Bengtsson (2012) could show in a driving simulator study that although every interaction form resulted in an increase in lane deviation and also a delayed initiation of lane changes, the combination of tactile and visual cues led to the highest decrease of these effects on performance.

## *2.7 Usability and suitability*

There are several standards, guidelines and recommendations when designing a product, in the automotive or another context. Different aspects of usability are discussed in the ISO 9241 series, others taking place explicitly in the driving context are defined in the ISOs 15005-15008, ISO 17287, and many more. The following paragraphs deal with the specific characteristics of these standards, the resulting consequences and their application in the context of multimodal design in the automotive context.

### 2.7.1 Definitions

Usability is defined in the DIN EN ISO 9241-11 as the "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (DIN EN ISO 9241-210: 2010, p. 6). Effectiveness describes how good a user can fulfil a task, in other words: reaches his goals, measured in accuracy and completeness. Efficiency is defined by how much resources it takes to fulfill goals in an effective way for the user; and satisfaction as the absence of discomfort and a general positive attitude towards using the product (DIN EN ISO 9241-210: 2010).

This definition of usability is aiming at the market of consumer electronics. It was developed in the context of software use. In this context the priority of the task is a different one compared to the priority of a task in the driving context. When designing a product for a primary activity, the rules of this standard should be taken into account and should be applied as well. In a context with more than one task, a primary activity or task is the one which has the highest priority. A secondary task is a task with less priority. It is important to note that -at least in the driving context- this definition is based on objective criteria; there are situations where it may indeed be of higher subjective importance to a driver to write a text than to watch the road ahead. Nevertheless, due to the safety-critical aspects for the driver and other traffic participants, the objectively more important task with objectively greater impact possibilities is driving and is therefore defined as the primary task. Following this a primary task is any task that is directly connected with the driving task, such as planning, maneuvering and stabilizing, whereas a secondary task is every activity which is conducted independently from the driving task (Jürgensohn & Timpe, 2001, p. 14).

Bubb (2003) defined the primary tasks as those tasks which are crucial for driving, such as steering and lanekeeping, whereas a secondary task is everything which is connected to the driving task without being directly to steering or lanekeeping, such as using the windshield wiper or turning on the head beams. Tertiary task would be everything that is related to comfort or communication, such as turning on the radio or talking on the phone. Whereas the first definition is not as accurate and detailed as the second, the term "secondary task" is used far more frequently for the tasks described as tertiary by Bubb (2003) and therefore more common. Additionally, switches are usually not used for the primary driving task in both definitions[4], whereas it is not a priori defined if a switch is employed for a secondary or tertiary task according to the definition of Bubb (2003). On this account the term secondary task is used in the following passages.

Infotainment and display concept products in the driving context obviously need to be designed in a different way as they are not the primary, but the secondary task. The usability of a secondary task interface is defined as suitability in the DIN EN ISO 17287:2009. This standard defines the impact on the driving task, the controllability, the efficiency and the learnability of a secondary task system as the most important suitability aspects, which are also discussed for a warning system in the DIN EN ISO 16352: 2005. These criteria are discussed more detailed in the following subchapter.

### 2.7.2 Criteria of suitability

The impact on the driving task is a negative interference on the ability of the driver to adapt to the vehicle and the environment and to drive safely. The following definitions are based on the DIN EN ISO 17287:2009.

[4] An exception for this rule might be Start-Stop-Switches, but they are usually not employed while driving with a certain speed.

The controllability is the character and the extent to which the driver can influence functions and interaction speed. This includes to initiate, to end, to repeat, to skip, to resume, to manage and to adapt functionality.

Efficiency is defined as the amount of resources which are dedicated by the driver to fulfill goals, in regard to exactness and completeness. This refers to the situational awareness of a driver, the mental, physical and sensory effort and the workload. Resources include physical and mental capacities as well as perception abilities. This definition of efficiency is based on the aforementioned DIN EN ISO 9241-11.

Learnability is defined as the gain of knowledge and abilities.

A switch push is a lesser complex task which does normally not require much learning. Contrary to that, when designing a switch, the impact on the driving task, the controllability, and the efficiency are crucial and have to be taken into account definitely. Their influence on interaction speed is dependent on the underlying functions which are not criteria for the switch push itself; in contrast, the actions mentioned in the description of controllability are all subject to the function. Therefore, controllability needs to be taken into account as such as that the mere switch push duration during driving needs to be investigated. The efficiency in this definition includes situational awareness; perception, processing, and reaction effort, which looks similar to the *Multiple Resource Model*; and workload. A problem with this definition is that this, again, refers to a whole function-based interaction, but does not take the microinteraction of a switch push into account.

### 2.7.3 Tactile usability and suitability

When taking the benefits for comfort functions mentioned in *Multimodal warning strategies* and *Multimodality in comfort functions* into account, the implication is that a non-obtrusive signal which is adequate to the situation increases suitability while driving. This adequacy is highly dependent not only on external factors, but also on the technical setup and design.

A general implication for multimodal feedback in terms of suitability is that its components should be congruent and not contradictory. This accounts in terms of usability (Wilson, Reed, & Braida, 2010) as well as in terms of long-term user acceptance (Ludden, Schifferstein, & Miquelon, 2012).

The DIN EN ISO 16352: 2005 states additionally that within the frequency range between 100 and 300 Hz, which was suggested in the chapter *Psychophysiological measurement of tactile perception*, signals “are also unlikely to be masked by road-induced vibration of the vehicle, which generally occurs at a lower frequency” (p.70). As the tactile modality is highly dependent on actuator type, size and contact area, these rather vague descriptions have their merit, but user studies are still necessary for determining the exact settings.

### *2.8 Constructs of user experience*

The term *user experience* was coined by Norman, Miller, & Henderson (1995) and was intended to elicit a broader view on the acceptance part of the usability definition. The DIN EN ISO 9241-210: 2010 defines the components of user experience as “a person’s perceptions and responses that result from the use and/or anticipated use of a product, system or service” (p.7). In this definition, all emotions, beliefs, preferences, perceptions, physical and psychological responses, behaviors and accomplishments that occur previously to, during and after the interaction of a user with the system are relevant. User experience is a consequence of brand perception, design, functionality, system output, the interactive behavior and the supporting resources of the interactive system, the internal and physical state of the user based on his experiences, attitudes, abilities, character and the context of use. It is possible to subsume user experience under usability, as long as the latter is viewed from the point of user goals and criteria of interaction. Usability can be employed to measure user experience (DIN EN ISO 9241-210: 2010). As this is a rather broad definition, research took place to investigate the most important factors of positive user experience.

Provost & Robert (2013) conducted a literature research and separated user experience into a product and a user pole, each containing five dimensions: for the product pole, functional, usability, informational, physical and external characteristics are of importance. The user pole consists of perceptual, cognitive, psychological, social and physical attributes. The authors evaluated the importance of each dimension through a user evaluation with user experience stories; for a positive user experience, the usability and the functional dimension is most important from a product perspective, and the perceptual and the psychological aspect from a user perspective. Perceptual aspects refer to the importance of the perceivable aspects of the product, whereas the psychological aspects mainly focus on the emotional impact and the underlying values of a product on the user. These results point into the direction of the influence of physical aspects of different products on user experience and emotions. Chen et al. (2009) developed a model that describes the transition from haptic product characteristics to emotional and affective dimensions (cp. Figure 5. *From physical to affective layer in tactile interaction*). For tactile interaction, there is no corresponding model so far, although the amount of research in the area of user preference is noteworthy. Several attempts have been made to add valence estimation to purely tactile perception, such as interaction commands

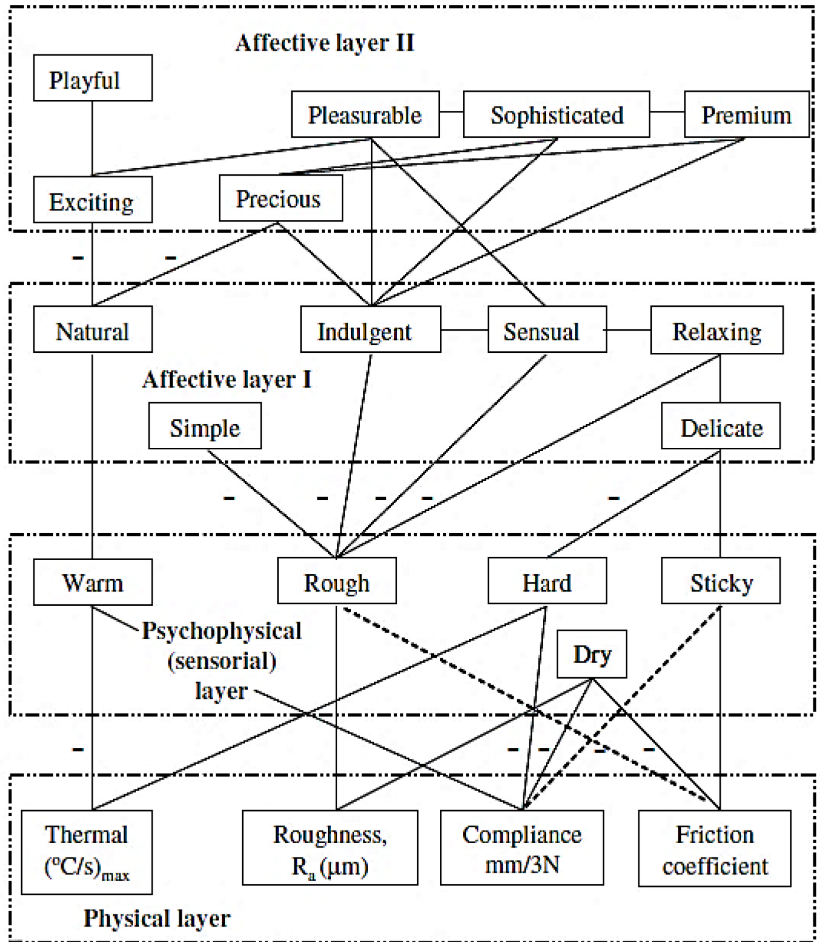



*Figure 5. From physical to affective layer in tactile interaction. Reprinted after Chen, Barnes, Childs, Henson, & Shao (2009), p. 4306.*

(Brewster & Brown, 2004), but this kind of valence addition usually requires the training of participants and is highly dependent on the context of the interaction (Jones & Sarter, 2008). A more accurate recognition without the requirement to previously learn specific meanings could so far only be obtained for the dimension importance, coded via location, characteristic frequency and duration (Hameed, Ferris, Jayaraman, & Sarter, 2006). This is rather an application area for warning functions, and not to be generalized over all feedback types and applications.

To elicit a positive experience, the usability criteria for an interaction need to be sufficient, and the perceptual and psychological impacts of a system need to be designed carefully. For research on characteristics of a system, its application context is important in terms of the appropriateness of the usability characteristics and in terms of its positive user experience. The descriptors with which the user experience can be measured are not fixed for tactile switches in a multimodal context, and on which dimensions physical aspects of an object relate to valence is not evaluated yet. In order to design a measurement instrument for user satisfaction, the characteristics and descriptors on which the judgment is based upon needs to be built up first.

### *2.9 Interaction media*

There are multiple possibilities for interactions in a modern car. Through the technology push coming from the consumer electronic market, access to mobile internet and apps is a consumer requirement and surfaces more and more frequently in cars. Different controls and control devices are distributed over the steering wheel as well as the center console, the inside of doors and seats. The following chapter gives an overview over the different variations of switches in the automotive domain.

It is generally possible to distinguish media which provide information to the driver and media which receive input from the driver. This general classification is however very rare, as most interaction media either directly give information back through tactile, auditory or visual feedback or request input from the driver with their information display. Another possible classification of the interaction media would be through the interaction channels proposed in Figure 4. *Wickens' Multiple Resource Model*. This is very appropriate when the focus is on human perception and action. In cases where the focus lies on the technological setup, another classification is needed. Bhise (2012) classifies the interaction technology mainly in displays and controls, which is very suitable as long as displays do not become control elements by themselves through using capacitive or resistive touchscreens. As most new displays imply some sort of touchscreen function, this separation cannot be kept up entirely. The separation in this chapter follows the evolvement in history for the automotive controls and therefore the underlying technical possibilities in order to give a framework for the environment of the user interaction. In describing the interaction media in this way, it is aiming at drawing a clearer picture of the requirements and test scenarios for the interaction. At first, conventional switches and controls are discussed, followed by the first examples of touch-screen interaction in the automotive domain. The chapter will close with the requirements for virtual switches in the automotive context. Specific design guidelines, which are developed in this context, will be discussed in the corresponding chapters.

### 2.9.1 Conventional switches

Haptic input devices are based on the advantage, that most of everyday interactions, such as opening a door or switching on an electric kettle, are in fact haptic. This means that a haptic interaction feels rather natural to a user, as it is performed regularly in a similar way. One form of this haptic interaction is the depression of a common mechanical switch. Due to their mechanical characteristics, mechanical switches create a multimodal, intuitive feedback, consisting of haptic, kinesthetic, visual and auditory information (Hoggan, Kaaresoja, Laitinen, & Brewster, 2008). Conventional switches were present in cars as early as in 1913 in Ford Model Ts[5].

In the design of these switches, reachability is the first concern. This is standardized within the DIN EN ISO 3958: 1978. The location of a switch on a steering wheel is of particular benefit as the driver does not have to take his hand off the steering wheel or visually search for a switch in the center stack, which may cause distraction. The DIN EN ISO 15005:2012 regulates the design and position of controls in the automotive context as well as the dialogue rules for input and output of a system. It states that “hand controls that are frequently used in conjunction with steering activities (e.g. cruise-control switches) will be positioned within fingertip reach of the steering-wheel rim“ (DIN EN ISO 15005:2012, p. 9). Functions which are frequently used should definitely be put on the steering wheel, but there are also different possibilities of integrating other functionalities than those on the steering wheel. The functions which are integrated vary highly between the original equipment manufacturers (OEMs). Common features include comfort

[5] http://www.mtfca.com/discus/messages/331880/366800.jpg

functions, such as push-to-talk, scrolling and selection functions, volume control and others, but also ADAS functions like the mentioned setting of the speed limit for adaptive cruise control (ACC). Complex interactions, such as a central controller computer unit, are sometimes performed via rotary-and-push knobs. The latest trends here are touch interactions with the possibility of handwriting recognition integrated into those knobs. Nowadays variations include push-button switches, rockers, dispatchers, toggle switches, rotary selector switches, thumb-wheels, levers and many more. Zeilinger (2005) classifies these interaction elements into push-button switches, switches, adjusters (in German Taster, Schalter and Steller) and further classifies these according to the movement which can be conducted with them: definitely (Anguelov, 2008). A mapping between psychometric and physiological parameters as well as an OEM's branding is aimed at.

Switch activation can be viewed as a microinteraction. „Microinteractions is about those critical details that make the difference between a friendly experience and traumatic anxiety" (Norman, 2013). For each application it is consequently necessary to pay close attention to the exact design of these microinteractions. When this is not the case, even a very sophisticated interaction design in terms of user experience and usability will inevitable fail. It is therefore necessary to pay close attention to the design of switches and their settings, optimizing it for user interaction regardless of which technology is employed.

#### *2.9.1.1 Design guidelines for conventional switches*

There are different kinds of requirements and standards depending on the kind of grip a human hand performs in the interaction. As the discussed technology focuses on multifunctional switches on a steering wheel, the main interaction finger of a hand is likely to be the thumb. The DIN EN 894 part 3 regulates the general requirements for controls and actuators. The manual contribution to the primary task, steering, has to fulfill the very important requirement of being able to be operated with a gloved hand, on a medium level of importance the ability to be cleaned and on a small importance level to avoid friction and accidental operation. Dealing with a push-switch, which is used in the secondary task (as usually employed with a multifunctional switch on the steering wheel), it is very important to fulfill the requirement of avoiding accidental operations, medium requirements for operation with a gloved hand, ability to clean and less importance requirements to avoid friction (DIN EN 894, 2000). Note that the latter originally referred to activation with the forefinger, not the thumb. The same requirements, however, can be and are generally used when pushing with a thumb as the nerve concentration of Pacinian corpuscles is not different. Depending on the angle of the thumb and the size of the hand, applied force may vary largely, although it feels like the same switch push over an arranged switchpack in the steering wheel spoke. Although the DIN EN ISO 3958 (1978) regulates the driver's hand-control reach, this standard refers to every hand control reach dimension but the hand-control reach on a steering wheel. Primary driving task related switch position is standardized in the DIN EN ISO 04040: 1986. As an increasing number of functions is integrated, more positions of the steering wheel are used for positioning switches (cp. Figure 6. *Multifunctional Steering Wheel*): the spokes at 3 and 9 o'clock position, the backside of the spokes (visible as pedals as well as hidden switches on the back) at the lower sides of the airbag cover or additional switches at the steering column switches. The activation and deactivation force and their snap ratio aimed at, vary for different switches across different cars. Bhise (2012) describes a value between 1.8 N and 5.3 N for typical push-switches. Although this is well established within the range of conventional switches on the steering wheel, there are different settings which are also commonly applied. Depending on the specification of the OEM, activation forces can vary between 3.0 and 6.7 N. Some offer tolerance values, ranging from 0.6 to 1.5 N for activation force. A study conducted by Anguelov (2008) suggests 3.1 N – 3.3 N or 2.45 N – 2.65 N for activation force. The snap ratio between activation and deactivation force varies as well: some OEMs use strictly 50%, others give a tolerance between 10% and 33% tolerance to 50%, depending on the OEM. Other OEMs use snap ratios of 70% or 35%. Surprisingly there are even snap ratios with a difference between activation and deactivation as

*Figure 6. Multifunctional Steering Wheel. From https://commons.wikimedia.org/wiki/File:Audi_R8_Lenkrad_20080225.jpg (2016)*

low as 25%, which are as low as some tolerance values of other OEMs. Other authors recommend a snap ratio between 35% and 50% (Rosenberg, 2003/2004; cited after Anguelov, 2008) or 40% - 55%. In Figure 7. *Force Time Diagram* different parameter which need to be taken into account for one switch push are displayed. These values depend highly on the location and function of the switch on the steering wheel.

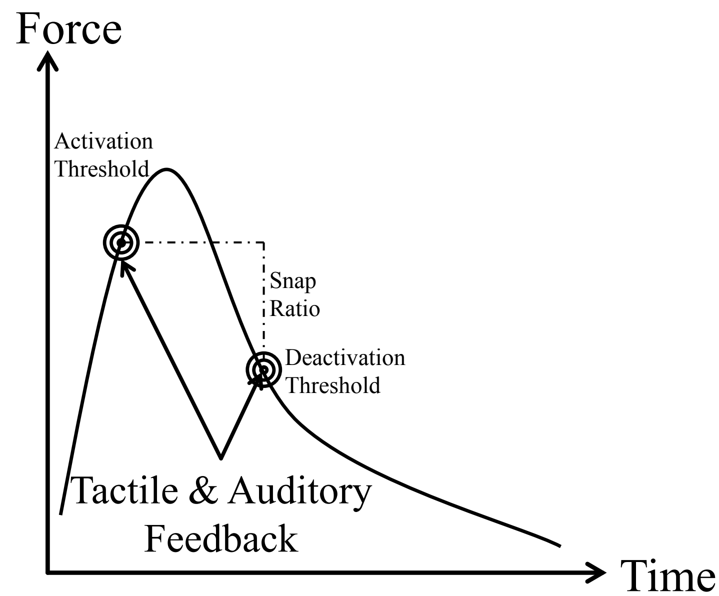


*Figure 7. Force Time Diagram.*

Mechanical switches solely rely on the application of force on the interaction surface. Their characterization is defined by the physical displacement, the spring mechanism, the overall mechanics of the system, the actuation sound, the actuation force and the actuation displacement. Immediate loss or application of force has immediate tactile implications.

### 2.9.2 Resistive and capacitive touchscreens

A present trend in the automotive industry is the substitution of mechanical switches through input devices from the consumer electronics market, resistive and, to a higher degree, capacitive touchscreens. Touchscreens are a technology which was first invented in 1965 (Johnson, 1965). Resistive and capacitive touchscreens are most common in consumer products nowadays. Surface auditory wave and infrared touchscreens are too cost- or space-intensive (Philipp, 2013). Because of their relevance in application, the characteristics of resistive and capacitive touchscreens will be explained in more detail.

Resistive touchscreens are made of two layers of resistive material or resistive coated material with an air gap between them. As a finger touches the surface, both layers are connected and pass a voltage, which allows for the allocation. The advantage of this technology is that it can be used with gloved hands or any non-capacitive material, but the response times are slow, as the deactivation of the resistive touchscreen is not instantaneously, but delayed through the inertia of the touching screens to get back in their pre-activation position. In addition to that the spatial resolution is not very exact, but sufficient for any non-portable device, such as train ticket machines. The visual perceptibility of resistive touchscreens is also not optimal (Lancet, 2013), especially when it comes to conditions with high luminous emittance, such as direct sunlight or bright daylight conditions. This is due to internal layer light reflection. As displays in the automotive context must be readable during all light conditions (DIN EN ISO 15008:2003), this makes resistive touchscreens not suitable for the use in this domain, although the first implementation of a resistive touchscreen into the automotive sector did take place in 1986 with the GM Riviera (Cox, n.d.), but only the following Buick Reatta included this feature. In 1990, the system was dropped because users found it “onerous and distracting” (Del-Colle, 2013). Adding to that there were some implications that the handling of the system seemed to have been too difficult for the users (Badal, 2008). Resistive touchscreens do not allow for gestures, but only for single push activations. Although there are serious drawbacks to resistive touchscreens, some features are superior to capacitive touchscreens, its successor technology: they can be operated with gloves, are less expensive and are not as dramatically prone to higher levels of humidity.

Surface capacitive touchscreens sense the change of induced capacity on the surface of the screen. When an electronic conductor moves into the electrical field, the field changes and thereby initiates activation. A human hand can function as an electronic conductor as touching the screen leads to a distortion of the electrical field. Because of these electric properties, wearing gloves often makes it impossible to use a capacitive touchscreen, not to speak of hand prostheses: as gloves mute or eliminate the conductive properties of a finger, the activation cannot take place, which occurs similarly with hand prostheses.

Projected capacitive touchscreens consist of a liquid-crystal display (LCD) layer, which is an actual screen, and one or more etched conductive indium tin oxide layers (ITO). These ITO form multiple horizontal and vertical electrode rows, which derive signals from a sensing chip. This chip can allocate the touched point on the surface. The position detection works due to distribution of the change in signals between the X and Y electrode rows (Philipp, 2013).

The triggering works for both sorts of capacitive touchscreens with capacitive devices only, not with e.g. a normal pen. After an activation the distortion can be traced back to its occurrence by sensors at the frame of the touchscreen (Sears, Plaisant, & Shneiderman, 1992). Surface capacitive touchscreens have the sensor located right at the surface of a device and these sensors are therefore prone to damage through dust, scratches or other contamination. The accuracy is not very high (Philipp, 2013). Another problem can occur due to the fact, that the induced field measures a longer activation of the screen than actually intended by the user, as activation lasts as long as the finger remains in the electric field. A longer sensing of the activation leads to a delayed deactivation, which delays a tactile feedback in turn, which may have been intended to be provided to the user. As the sensing is based on an induced field, faulty activation can happen quite easily, if the activation is not perpendicular but with a certain angle. Depending on the angle from which the intended activation takes place, nearby fields may be activated additionally or prior to the intended activation, which may result into faulty or accidental activation. Additionally, road vibrations during driving can also easily result in faulty activation due to unintended micro-movements of the user's hand.

In 2003, Honda introduced a touchscreen into its Acura TL (Honda, 2002). From 2009 on, nearly every bigger OEM brought capacitive touchscreens into one or more of his concept cars or models[6], in most cases into the center stack. In furthermost of the cases the touchscreen is for the interaction with a navigation system only. The reason for this lies in the technology of capacitive touchscreens. As mentioned before, capacitive touchscreens are more or less inaccurate in detecting the location of the finger. When it comes to driving, it is nearly impossible to move a finger accurately to one capacitive button because of road vibration (Ahmad et al., 2015). The consequence of this is, obviously, faulty activation and frustration of the user, with the possibility of drawing too much attention towards this secondary task and away from driving, which in turn might lead to severe safety impacts. Ford, which introduced a fully touch-based control panel in the center stack with the MyTouch system, decided to refrain from this strategy for future models

[6] all website visits at the 16.08.2012:
2009: Nissan (http://www.autosieger.de/article17248.html), Lexus (http://en.wikipedia.org/wiki/Remote_Touch),
Audi (http://telematicsnews.info/2009/12/01/audi-officially-launches-mmi-touch-handwriting-recognition/),
VW (http://www.auto.de/magazin/showArticle/article/44402/Cleveres-Geraet-Der-Touch-Adapter-Voice-von-Volkswagen-Zubehoerws/118129.html)
2010: Daimler (http://blog.mercedes-benz-passion.com/2010/02/f-800-style-neuartiges-bedien-und-anzeigesystem-cam-touch-pad/)
2011: GM (http://www.gm.com/content/gmcom/home/toolbar/search.html?num=30&q=touchscreen&start=0),
Toyota (http://www.golem.de/1110/87316.html),
Volvo (http://www.engadget.com/2011/09/13/volvo-unveils-concept-you-luxury-sedan-dripping-with-touchscree/)
2012: Hyundai(http://www.hyundai.de/pages/modelle/i30GD/innen.html), Kia (http://www.autogenau.de/news/modelle/der-neue-kia-ceed-debuetiert-in-genf.html),
Opel(http://www.opel.de/tools/suche.html?tp_gsa=search%3Fsite%3Dgmds%26client%3Dgmds_frontend%26q%3Dtouchscreen%26proxystylesheet%3Dgmds_frontend%26output%3Dxml_no_dtd%26filter%3D0%26sitesearch%3Dopel.de.prod-www.gm.plusline.net%26ie%3DUTF-8%26ip%3D82.98.123.150%26access%3Dp%26entqr%3D3%26oe%3DUTF-8%26ud%3D1%26sort%3Ddate%253AD%253AS%253Ad1), BMW (http://www.f30post.com/forums/showthread.php?t=716469), PSA (http://www.auto.de/magazin/showArticle/article/79900/Touchscreen-des-Peugeot-208-erhaelt-Design-Award)

after a tremendous amount of negative user feedback, which included complaints about the lack of tactile feedback from the system (Ramsey, 2013). Capacitive touchscreens in the steering wheel have so far been evaluated in university projects (e.g. Pfeiffer et al. (2010)), but are not implemented in series production so far.

It is clear that the proliferation of capacitive touch screens in consumer electronics is now, after a tumultuous start, generally well accepted, and has become almost synonymous with modern electronics. The majority of those capacitive touch screen implementations naturally requires the user to visually focus on the interaction task when operating the device, and in return the user is rewarded with immediate visual feedback. The interaction on a capacitive screen usually involves some sort of gesture, limited by the width of the screen. A bigger screen therefore allows for more precise sensing and is an essential need for capacitive touchscreens. When tactile feedback is provided by these systems, it is normally transferred through the whole device (Barua, 2013). As in the automotive context squeaks and rattles at unwanted areas are a huge issue, this kind of feedback transformation may also influence user acceptance and satisfaction negatively. The feedback given is normally perceivably delayed (Snell, 2008). There can be as much delay to the excitement of the tactile feedback that, in case of a single switch activation, the finger of the user is already removed, making the implementation of tactile feedback as a whole obsolete. This reduces a system which is intended to give feedback to one where the developers could as well save the effort of implementing it. If the finger has longer contact to the surface, as it is the case with gestures, tactile feedback makes more sense. These gesture interactions usually take longer than single switch activations and should generally be avoided while driving.

The importance of tactile feedback is not essential for effective operation of consumer electronic devices with capacitive touchscreens. As expressed earlier the importance of tactile, haptic and audible feedback increases substantially for systems where visual focus on the touchscreen should be minimized, such as those in an automotive context. It has to be ensured that the feedback reaches the user, which is due to the aforementioned reasons not always the case.

Electro active tactile feedback, when applied to consumer electronic devices, has typically been of the vibro-tactile style, using *Eccentric Rotating Masses* (ERM), and is typically felt primarily in the hand holding the device - not necessarily focused on the hand or finger interacting with the touchscreen *generic user interface* (GUI). Usage and effectiveness of such systems are varied, and none properly recreates the instant tactile feedback experienced when operating a typical rubber membrane or dome-cap-style mechanical switch. One of the reasons behind this is to be found in the inherent nature of the technologies being used and the way in which mechanical switches operate. For all the limitations mentioned above, it is clear that the imitation of a mechanical switch through capacitive buttons cannot succeed completely, no matter how carefully its design is performed.

Concluding on this it can be stated, that neither resistive nor capacitive touchscreens are optimal for the implementation in the automotive domain. Every single technical drawback does not only directly influence user quality perception negatively, but is also endangering and distracting for the user. Capacitive touchscreens need more precision in activation or deactivation, which could be implemented via active force sensing. On the automotive market the force sensing implementation is currently a solution which is too expensive for series production. Additionally, as long as these solutions are not able to provide accurate tactile feedback due to the capacitive sensing principle, they remain distracting. To combat this inaccuracy, a number of workarounds is existing, e.g. providing tactile feedback only on the release click, which in turn might be too late, but at least after a user corrected his movement. This shows that the issue is well known by the OEMs and that they try to solve the drawbacks of this technology. Resistive touchscreens should not be employed due to their visibility issues and the sluggish and delayed feedback they provide. Both resistive and capacitive touchscreens are limited to some sort of glass layer for the surface. To transfer this technology safely from the domain of consumer electronics to the automotive domain is overall very challenging for OEMs due to the reasons mentioned above. They are caught between the demand of their customers to implement and integrate touchscreen technology, which is not suitable for the automotive domain, and the standards and specifications of said domain to ensure safe driving.

#### *2.9.2.1 Design guidelines for touchscreens*

Generally, the DIN EN ISO 9241-920 regulates the design of tactile and haptic interaction. Main topics are input and output design, the encoding of information, context-specific encoding and other interaction guidelines. Key points of the standard are of course the design of the accuracy of the interaction, realistic experience, adaptation and temporal masking (which should be avoided) and the learnability of the interaction. The position where a tactile interaction takes place should be rather distal in order to be optimized for spatial resolution. This is the case when interacting with fingers or the hand, as this is quite far from the body center, and in line with the highest density of Pacini corpuscles. There are different information encoding possibilities: object shape, temporal pattern, vibration amplitude, vibration frequency, location, temperature, and thermal conductivity. Content specific encoding deals with text data, graphical data, and controls. The latter additionally covers force activation. Especially for avoiding accidental activation, force activation is recommended. The tactile frequency for feedback on touchscreens is recommended as frequencies between 10 Hz and 600 Hz, with the suggestion of using the area between 50 Hz and 250 Hz (DIN EN ISO 9241-920:2009). In order to avoid spatial masking of two stimuli, different frequencies for those, one below 80 Hz and the other above 100 Hz (DIN EN ISO 9241-920:2009), are recommended. The DIN EN ISO 16352: 2005 recommends the range between 100 Hz and 300 Hz, basing their recommendation on research conducted by Verrillo (1963). There are no recommendations for feedback delay times, as to date they are technically inherent and should be reduced as much as possible.

Tactile feedback in touchscreens can be classified into *inertial*, *piezo, surface, lateral, and electro-active polymer actuation* (Banter, 2010). Those actuators have a latency between 70 and 170 ms (Agawi, 2013). Typically ERMs or *Linear Resonating Actuators* (LRA's) are coupled to the screen or housing to create a tactile feedback upon the detection of an activation. The initiation of these tactile feedback events can additionally have a perceived lag, which may vary between 30-50 ms for ERM's and 20-30 ms for LRA's (Huotari, 2012). The feedback of mechanical switches is mostly more instantaneous, although for tactile feedback, a combination between the fastest actuator and the least delayed processing lies at 90 ms. The total time of 220 ms for the slowest combination can definitely not compete with mechanical switches in terms of instantaneous feedback.

In terms of just notable differences of amplitudes, recommendations range roughly between 0.4 to 6 dB, with lower recommendations between 0.4 – 3 dB (Craig, 1972; Fucci, Small, & Petrosino, 1982; Geldard, 1957; Globe & Hollins, 1993; Rinker, Craig, & Bernstein, 1998) and recommendations which cover the broader range from 0.5 – 6 dB (Craig, 1974; Gescheider, Bolanowski, Verrillo, Arpajian, & Ryna, 1990; Gescheider, Zwislocki, & Rasmussen, 1996).

These different JNDs depend on frequency and length of the signal, as well as with every vibrotactile perception, on the size of the stimulation area.

Another possibility for haptic feedback is force feedback. A multimodal feedback for virtual switches has been developed by Faeth (2012), where the displacement curves like depicted in Figure 7. *Force Time Diagram* is performed by force feedback. Force feedback can to this date only be provided in larger devices and cannot be integrated into the automotive context in a cost-effective manner for switches.

### 2.9.3 Virtual force sensitive switches

A *virtual switch* or *button* is defined by Faeth (2012) as a "switch in which the feedback generated for the user is not solely dependent on the movement of mechanical components" (p. 2). This definition includes touchscreens with or without tactile feedback, as a feedback on a touchscreen using only visual or auditory feedback counts as non-mechanical feedback as well. The term *button* is more frequently common in computer technology than in the automotive domain, where the term of reference is usually *switch*. When employing tactile feedback, different actuators can be used. For the determination of activation or deactivation, capacitive touchscreens employ the activation in their capacitive field. Few technologies on the market are capable of force-sensing, and the applicability of them in the automotive context is frequently not given.

The investigated 3D switches are a specific kind of virtual switches which are force sensitive and inherit the characteristics of a virtual switch. In addition the generation of auditory feedback, directly from the actuator device, can be timed in the way, that it accurately coincides with the tactile touch events, adding significant realism to the system. They are a force sensitive solution with linear output, which means that the difference between force input and sensor output has no deviation, in other words no hysteresis; and also zero creep. The switches have a small form factor, which is especially important for the reduction of designed space in the steering wheel. The vibrational and the auditory feedback is provided via a low frequency resonant acoustic vibration surface, which is fast-acting and allows for sharp feedback. They have complete freedom in terms of overlay material, opposed to resistive and capacitive touchscreens; even the thickness of the overlay can vary across the surface and the force sensing will still be accurate.

The successful transfer to the automotive context highly depends on the adaptation of auditory feedback to the particular environment and the reduction of the visual demand as far as possible. In contrast to that, the feedback of the activation and the deactivation of a function must be very precise. Tactile feedback can transport this information, but the specific design of the feedback and its vibration patterns is crucial. Another topic to be evaluated carefully is the actual switch behavior. Activation and deactivation force, snap ratio or, for scrolling functions, the time constant must be designed very carefully. The exact setting in terms of their usability while driving, with a focus on user acceptance and its measurement, is the scope of the present thesis.

## *2.10 Previous studies*

Several studies, previously conducted with the 3D switches, influenced the present thesis to some extend. They are briefly discussed in the following subchapters.

At first an expert evaluation was conducted. The material of the surface plays another important part, as this is completely variable with the present technology. A multimodal study with an earlier prototype investigated optimal combinations between auditory and tactile feedback, the auditory settings were investigated in a separate university project, and the first study concerning tactile feedback characteristics revealed the need for investigating more appropriate methods than existing, for the measurement of user acceptance for this new technology.

### 2.10.1 Expert evaluation

In User Centered Design expert evaluations take place at a very early development stage. The goal is usually to get first impressions on the design and the possible applications of the prototype from experts not being involved themselves in a current project. In this particular study, an additional goal was to get a first impression on possible settings for feedback and force settings, as well as on the general look and feel of the virtual switches.

#### *2.10.1.1 Method*

This expert evaluation, which can be described as a heuristic evaluation with expert subjects (Nielsen & Molich, 1990), took place with an earlier version of the current prototype. In total, five experts with a mixed knowledge about design, usability, marketing and electronics were interviewed, as this interdisciplinary approach is usually recommended (Nielsen, 1992). They were more or less associated with an OEM, and represented therefore their different design guidelines and styles.

At first, the prototype remained hidden under a cloth. The experts were asked some general questions about demographic data, their driving experience and car usage before the core interview started. A connection to the prototype was established via questions regarding multifunctional steering wheel familiarity and usage questions. In the next part the prototype remained hidden under the fabric still. The tactile feedback was not activated at this point and the expert only investigated the kinesthetic properties. The next part contained blind testing as well, with the additional activation of the tactile feedback in a fixed, standardized setting for every expert. Following this evaluation, the prototype was revealed and examined with the support of the visual modality. The subsequent and biggest part of the interview consisted of different threshold settings for activation and deactivation force, as well as tactile and acoustic feedback. The final part consisted of questions regarding possible functional applications and general value appeal questions. Each interview lasted between 1.5 and 2.5 hours.

##### *2.10.1.2 Results*

The interviews were transcribed and coded following the procedure proposed by qualitative content analysis (Mayring, 2000). The essential results of this study were the identified points for improvement: the feedback, acoustic as well as tactile, as they were not perceived as qualitatively high enough, and, as the degrees of freedom in terms of material and activation and deactivation force led to very different suggestions from the experts, further research in these areas with user testing. These findings led to the studies described, as well as partly to the development of the present dissertation. The full study is described in more detail in Diwischek, Essers, & Lisseman (2014).

### 2.10.2 Material design

This study was set up in order to investigate the interaction between different surface materials and vibrotactile feedback, as the degree of design freedom is highly restricted to the automotive applicability of the materials under investigation. For the switch technology itself, any material could be used for a surface.

##### *2.10.2.1 Method*

An earlier prototype of the present technology was used in this study. After a selection via a pretest, four automotive materials were evaluated as a switch surface, including leather, textile, silicon as a non-rigid automotive material and polycarbonate as a display-like surface (used in less expensive touchscreens). Participants first examined a hidden prototype without feedback and then with two different vibrotactile feedback stages. This procedure was repeated with the visual display of the prototype. A semantic differential after Wellings, Williams & Tennant (2009) was used to evaluate the different settings. Finger and room temperature were evaluated as co-variables before and after the experiment. 48 healthy subjects according to NHTSA age groups (Garrett, 2012) participated voluntarily in this study. The experiment lasted about 1.5 hours.

##### *2.10.2.2 Results*

Due to the addition of a valence scale, positive and and negative item poles could be investigated (Vogelpohl, 2014). Among other results, an interaction effect between the feedback intensity and the preference for a material could be observed. Harder surfaces were preferred with a weaker tactile feedback and softer materials with a stronger tactile feedback. This might be explained due to the fact that the feedback influenced the perception of the *hardness/softness* dimension, which is one of the most important dimension on a semantic differential scale focusing on materials (see e.g. Lederman & Klatzky, 2009; Okamoto, Nagano, & Yamada, 2013). Additionally some subjects reported difficulties in adapting the items of the semantic differential to the present prototype. The full study including detailed results can be looked up at Vogelpohl (2014).

For the present work it has to be kept in mind, that the material is an important influence factor for the feedback design and the feedback should be chosen accordingly to the material used as an overlay surface. Surface materials should be kept constant when evaluating different types of feedback or different types of switches in order to avoid interaction effects.

### 2.10.3 Single switch activation functionality

In one previous study using the same prototype as in the expert evaluation and the material study, the interaction of vibrotactile and auditory feedback was investigated in a driving simulator. This was conducted in order to better understand which setting combinations were more suitable for multimodal feedback presentation.

#### *2.10.3.1 Method*

40 healthy subjects participated voluntarily in the testing. With a pairwise comparison, error rate, time on task and additional driving parameters, nine different feedback intensity stages were compared while sitting and while driving. Finger and room temperature were measured before and after the experiment. This study only included single switch activation, which means that a push and hold function was not evaluated.

#### *2.10.3.2 Results*

Results showed that a medium acoustic feedback can be combined with a medium to high vibrotactile feedback. Extremes in combining both feedback modalities, such as both being very high or both being very low should be avoided, in terms of error rate as well as in terms of preference.

Although the results being not fully transferable due to technology changes, this study provides guidelines for further design on multimodal feedback for virtual switches. Especially the combination preferences influenced the feedback design of the final study of the present work. There were no differences in the preference rating between the sitting and the driving condition, implying that for preference ratings only, the driving task is not essential to provide the context. For safety-critical measures, the investigation while driving is crucial nevertheless. This is only applicable to a driving simulator and not to driving on a test track, as road vibrations will have a certain influence on the feedback perception. The full setup of the study, as well as the exact signal design, can be looked up at Last (2013).

The next studies discussed aim at dealing with detailed questions for the switch design.

### 2.10.4 Auditory settings

As the expert evaluation also pointed in the direction of improving the auditory feedback setting, a joint study with the Georgia Tech University was rolled out and supervised by employees of TK Holdings, Inc., Auburn Hills, and TAKATA AG, Berlin. The auditory setting used in the main experiment was specified in this previous cooperation study with the Integrated Acoustics Laboratory, School of Mechanical Engineering, of the Georgia Tech University.

#### *2.10.4.1 Method*

31 healthy subjects, students at the Georgia Tech University, participated voluntarily and evaluated thirteen signals, partly recorded sounds of automotive mechanical switches, and partly digital sounds. The experiment was divided into two parts: the evaluation of 16 psychoacoustic quality criteria on a five point Likert scale in terms of how well they described the sound presented (with the additional option not to judge via "I don't know"), and the mapping towards 17 common switch functions in the automotive domain. The psychoacoustic quality criteria were selected via a pretest with five students at Georgia Tech out of 26 adjectives used in different studies (Bech & Zacharov, 2007; Blattner, Sumikawa, & Greenberg, 1989; Brewster, Wright, & Edwards, 1993; Fastl, 2006b; Larsson, Opperud, Fredriksson, & Västfjäll, 2009; Suied et al., 2005; Zeitler, Ellermeier, & Fastl, 2004). Adjectives were excluded when frequently they received the reply "don't know", which indicated that these adjectives were not appropriate. The common switch functions were selected by TK Holdings, Inc. and TAKATA AG in terms of their frequency on steering wheels. A total of 13 sound samples were employed in this study.

A VX Pocket 440 soundcard was installed in the laptop over which the sample sounds and the questionnaires were presented. A pair of Sennheiser model HD 380 Pro headphones were used for sample sound presentation.

Several consent forms were first presented to the subjects and then they were taken through an audiogram in order to control for hearing loss. After that they were familiarized with the psychoacoustic quality criteria questionnaire, and filled in one questionnaire for each sound sample afterwards. Then they were familiarized with the different switch functions and matched each sample to one or more switch functions. The experiment lasted approximately one hour.

#### *2.10.4.2 Results*

The full results can be found in Edwards & Cunefare (2015). The result, that the most significant important items for the rating of the acoustics of switch sounds were sharp, high/low frequency, dynamic and fast, are of interest for this study. One sound sample performed extremely well on these dimensions and was included in the final study design of the present work. This item and its integration will be discussed further in chapter *Final Grid design.*

### 2.10.5 Frequency and waveform

This study is described in more detail as the previous ones, as its impact is higher for the present work; this applies in terms of vibrotactile feedback design as well as in terms of methodological questions arising from it, which as a final consequence led to the present work. The goal of this study was to investigate whether a modulation of a pure sine wave could contribute to user preference and which frequencies should be chosen from a range suggested by literature.

#### *2.10.5.1 Method*

16 healthy subjects, all employees at TK Holdings Inc., voluntarily participated in this study. They were recruited within four different age groups: 18-24, 25-39, 40-55 and 55+ years (age groups according to NHTSA recommendations (Garrett, 2012)), with four subjects in each age group. Gender was equally distributed over participants and age groups. Both native English speakers as well as German expatriates participated. They had at least three years of driving experience and no impairment in terms of sensory or motoric hand function. The presented stimuli consisted of four different frequencies and two different waveforms. The frequencies were at 105, 135, 175 and 230 Hertz, and calculated with a just-noticeable-difference (JND) of 30% (Mahns et al., 2006; Rothenberg et al., 1977) to ensure that a set of perceivable different stimuli was used. Additionally, pretests with two additional subjects were conducted to make sure the stimuli were differentiable. The two signals in this study present were

- a constant sine wave signal and
- a sine wave with a fall before the actual start of the sine wave and a dampening on the amplitude of the cycles after the initial peak, termed “fall-lunge-decay” (FLD).

Both waveforms were presented at the four frequencies described previously. The force threshold for activation was set at 4 N and at 2 N for deactivation. The area of contact was approximately 14 $mm^2$ in size. Finger and room temperature were measured before and after the experiment. To exclude interaction effects between sound from the exciter and the tactile perception, subjects were wearing headphones through which white noise was played during the evaluation parts of the experiment.

Subjects first conducted a full pairwise comparison between each of the frequency-waveform combinations and then a rating on a semantic differential, for each of the combinations separately.

Ties were allowed in the pairwise comparison when subjects could not determine a difference between the stimuli. The semantic differential used in this study consisted of adjectives developed by Wellings et al. (2009). The valence attribution of this instrument followed Vogelpohl (2014). Consequently, as the pairs in this semantic differential consist of adjectives with an inherent valence dimension now, it was possible to extract which setting is more likeable. Additionally to the existing semantic differential, an overall *good/bad* scale was added for enhanced valence anchoring. The order for both evaluation parts was randomized and it was ensured that no subject received the same order as any other subject. Subjects were instructed to rate as spontaneously as possible in the pairwise comparison as well as in the semantic differential, but were allowed to explore each setting as many times as they needed for basing their judgments upon. They always performed the pairwise comparison first and the semantic differential rating afterwards.

#### *2.10.5.2 Results*

As the pairwise comparison method consists of binomial decisions, a ranking can be obtained by calculating the frequency of the preferences of one stimulus against all other stimuli. For each subject, the consistency of the ratings had to be calculated in order to see if they performed an unacceptable amount of circular triads (being therefore self-contradictory) and had therefore to be excluded from further evaluation. This was done using the method proposed in Bortz, Liener, & Boehnke (2008), with the calculation of consistency coefficients for each subject. 69% of the sample answered consistently overall. Furthermore, the concordance of the sample had to be evaluated. The concordance describes the agreement of all subjects on the rating of the different stimuli and is independent of the consistency of the individual subjects (Bortz et al., 2008). The Akkordanzmaß (A) describes the concordance of a sample, but an additional $\chi^2$ test can be performed to test whether concordance is systematical or random. Despite a low A (0.04), the sampled judged concordantly with $\chi^2$ (34) = 55.64; $p < 0.05$. There was no effect on gender and only one age group (25-39 years) was not concordant in their judgments. However, it has to be kept in mind that the sample size was too small to investigate gender or age effects thoroughly. The ratings can be transformed from ordinal to metric scale by employing the Law of comparative judgment (Thurstone, 1927), when the assumptions of normal distribution and small variance of the observed judgments around the true judgment are met. Normal distribution is achieved via employing a z-transformation of the data and normalizing them to the lowest preference values. This transformed data was then investigated further calculating and ANOVA, which became significant with F (7/48) = 5.872, p < 0.001. A Tukey HSD post-hoc was performed with 95% family-wise confidence level to investigate for differences between the stimuli. Significant results (p < 0.05) are displayed in the Table 2. *Post-hoc comparisons using Tukey HSD*.

Several single characteristics of the different settings could be obtained through the rating of the semantic differential. In terms of quantitative differences, the positive valence scores and the negative valence scores were summed up for each setting and means were compared. The assumptions for an ANOVA were violated heavily in terms of normal distribution (Shapiro-Wilk with W = 0.94, p < 0.001; examined skewness and kurtosis revealed extreme values of 37.28 and 1395.95, respectively). A Levene's test for homogeneity of variances did not become significant with F (7/1400) = 0.99, p = 0.43. Consequently, a nonparametric method had to be chosen. A Kruskall-Wallis test was employed, which became significant with H (7) = 62.95, p < 0.001. Post-hoc tests were conducted using the Nemenyi-test with Chi-squared approximation in order to correct for ties. As the Kruskall-Wallis is rather an unconventional data processing procedure for the present data set, a backup two-Factorial ANOVA was calculated using the *ezANOVA* algorithm package in R. A main effect for *setting* became significant with F (7/105) = 2.52, p < 0.001. A follow-up ANOVA on the variable *setting* confirmed this result with F (7/1400) = 9.32, p < 0.001. As the normal distribution was violated very heavily, the more conservative approach of employing the Kruskall – Wallis was used for retrieving final results. It is noteworthy, however, that the non-corrected version of the post-hoc Nemenyi test leads to similar results as the post-hoc Bonferroni t-test, which might be due to the fact that both approaches are too liberal when taking the nonparametric distribution of the data set into account. The results of all three post-hoc evaluation procedures are displayed in the table below (cp. Table 3. *Post-hoc comparisons using three different methods*).

*Table 2. Post-hoc comparisons using Tukey HSD.*

| | Post-hoc comparisons (Tukey HSD) | | |
|---|---|---|---|
| **Stimuli** | ***Compared with*** | ***Mean difference*** | ***Significance*** |
| FLD 105 Hertz | Sine 105 | -1.38 | 0.0047075 |
| | Sine 175 | - 1.50 | 0.0016168 |
| | FLD 230 | -1.43 | 0.0030882 |
| | Sine 230 | -1.77 | 0.0001339 |
| FLD 135 Hertz | Sine 230 | -1.19 | 0.0230977 |

*Table 3. Post-hoc comparisons using three different methods.*

| Bonferroni corrected t-tests | | | | | | | |
|---|---|---|---|---|---|---|---|
| | FLD 230 | Sine 135 | Sine 105 | Sine 230 | FLD 175 | FLD 135 | FLD 105 |
| Sine 135 | 1 | - | - | - | - | - | - |
| Sine 105 | 1 | 0.36762 | - | - | - | - | - |
| Sine 230 | 1 | 0.55504 | 1 | - | - | - | - |
| FLD 175 | 0.17503 | 1 | **0.00057** | **0.00307** | - | - | - |
| FLD 135 | 0.54998 | 1 | **0.02853** | **0.01465** | 1 | - | - |
| FLD 105 | **7.40E-07** | **2.00E-05** | **1.20E-08** | **7.90E-07** | 0.19445 | **0.00579** | - |
| Sine 175 | 1 | 1 | 1 | 0.48163 | 0.64979 | 1 | **0.0004** |
| **Uncorrected Nemenyi** | | | | | | | |
| | FLD 230 | Sine 135 | Sine 105 | Sine 230 | FLD 175 | FLD 135 | FLD 105 |
| Sine 135 | 0.99123 | - | - | - | - | - | - |
| Sine 105 | 0.9387 | 0.47483 | - | - | - | - | - |
| Sine 230 | 0.51648 | 0.09918 | 0.99447 | - | - | - | - |
| FLD 175 | 0.35478 | 0.87478 | **0.0171** | **0.00089** | - | - | - |
| FLD 135 | 0.50625 | 0.95019 | **0.03592** | **0.00231** | 1 | - | - |
| FLD 105 | **7.70E-05** | **0.00299** | **1.30E-07** | **9.30E-10** | 0.2004 | 0.11696 | - |
| Sine 175 | 0.9993 | 0.99999 | 0.66412 | 0.19184 | 0.72821 | 0.85493 | **0.00097** |
| **Corrected Nemenyi** | | | | | | | |
| | FLD 230 | Sine 135 | Sine 105 | Sine 230 | FLD 175 | FLD 135 | FLD 105 |
| Sine 135 | 0.9985 | - | - | - | - | - | - |
| Sine 105 | 0.9873 | 0.7897 | - | - | - | - | - |
| Sine 230 | 0.8158 | 0.3787 | 0.9991 | - | - | - | - |
| FLD 175 | 0.7008 | 0.9703 | 0.1383 | **0.0199** | - | - | - |
| FLD 135 | 0.8096 | 0.99 | 0.2153 | **0.0382** | 1 | - | - |
| FLD 105 | **0.0034** | **0.0453** | **2.40E-05** | **4.00E-07** | 0.5404 | 0.4128 | - |
| Sine 175 | 0.9999 | 1 | 0.8927 | 0.5292 | 0.9199 | 0.9644 | **0.0211** |

Some subjects reported difficulties in providing a rating during the testing because the word pairs of the semantic differential were deemed unfitting for the present technology.

#### *2.10.5.3 Discussion*

In the pairwise comparison three different sine waves were significantly preferred over the least liked FLD (105 Hz). Additionally the most preferred of the sine waves – sine wave 230 Hz - was significantly preferred over the FLD 135 Hz. In the semantic differential rating all four sine waves were preferred over the least liked FLD (105 Hz). Additionally the most preferred sine wave (230 Hz) was significantly preferred over FLDs 135 Hz and 175 Hz. Only the FLD 230 Hz performed significantly better than the least liked FLD. Both instruments seem to imply concordantly, that a FLD is less preferred than a sine wave, especially at low frequencies. Higher frequencies seem to be preferred over lower ones, to the extent that the negative rating which the FLD received in general was compensated, but only at the highest frequency setting. The most preferred sine wave seems not to be in line with literature (Dabic et al., 2013), but as already discussed, actuator size and design do have a great impact on the preference of a signal. As this technology seems to behave differently, the frequency range up to 300 Hz should be investigated. The subjects having difficulties to rate on some scales on the semantic differential, was another difficulty which frequently occurred during testing. This might be due to the fact that the scale employed was one designed for mechanical, not for multimodal virtual steering wheel switches. In order to develop a scale and evaluate more switch characteristics, such as enhanced frequencies, amplitude of the signal activation and deactivation force, snap ratio and different functionalities, especially the push and hold function and its timing, as well as auditory characteristics, more studies need to be conducted and the method of measuring user acceptance should be improved. This is the main goal of the present thesis. The full results of this study can be read at Diwischek & Lisseman (2015).

# 3. Methods

The following subchapters deal with different methods of measuring concepts of usability while driving, starting with objective measures of efficiency and effectiveness, or performance. Subsequently, different methods of measuring user acceptance are discussed and conclusions are drawn for the method adaptation for virtual steering wheel switches.

## *3.1 Methods of performance measurement*

As described in the standards for usability (DIN EN ISO 9241-210: 2010) and suitability (DIN EN ISO 17287: 2003), measures of user performance show how good a user can fulfill a task (effectiveness). Efficiency in the driving context relates to the amount of physical and mental, including perceptional, resources the driver dedicates to a primary and to a secondary task. These two measures are operationalized in the conducted experiments as following:

Effectiveness (primary task):

- *Lane keeping quality,* as it describes the ability of the driver to maintain lateral control of the vehicle. Lane keeping quality is defined as the area of deviation from the lane center over a predefined course length.
- *Speed and its standard deviation*, as subjects are instructed always to maintain a certain speed; deviation from this would be an impact of the secondary task on the primary task in terms of longitudinal control the driver retains.
- *Steering wheel angle velocity and acceleration,* as this describes whether a sudden steering movement is employed in order to maintain lane keeping.

Effectiveness (secondary task):

- *Error rate,* relating to unintended errors (Reason, 1990). The distinction between *slips* (attentional errors, such as missing a traffic light because of distraction) and *lapses* (memory errors, such as forgetting to perform the secondary task) are not differentiated further because in a microinteraction task with little semantic content, this differentiation can hardly be performed. Errors in the present case contain activating the switch too little or too many times.

Efficiency (primary task):

- *Participant rating* on task difficulty via NASA-TLX (Hart & Staveland, 1988; Human Performance Group at NASA's Ames Research Center, 1986), a six-dimensional questionnaire on subjective task load. The used version will be the raw-TLX (Hart, 2006), which leaves out the weighting of the different workload scales. The scale will be employed in a reducted version with the scales from 0-20 opposed to scales from 0-100 in the evaluation, keeping the full range of response possibilities but evaluating single iterations, not in steps of five.

Efficiency (secondary task):

- *Time on task,* as this provides objective insight on the secondary task effort

### *3.2 Psychophysiological measurement methods*

Threshold detection is a very basic method to evaluate a subject's abilities to perceive tactile or auditory feedback, such as the aforementioned JNDs, absolute threshold (the absolute levels of perception), or confusion scaling (Susini, Lemaitre, & McAdams, 2011), which is basically a JND scaling with equidistant scales. These methods are necessary for determining the discriminable stimuli, but do not give any hints on subjective quality perception. Different stimuli need to be chosen corresponding according to requirements and specifications as well as user preference, so user studies among the recommended ranges need to be conducted to investigate the most preferred setting. JNDs aim at the rate at which a signal is differentiable and are therefore a necessary precondition, when setting up stimuli for an experiment. In the vibrotactile context, JNDs are highly dependent on the frequency and amplitude of a signal (Mahns et al., 2006; Rothenberg et al., 1977), which is in turn influenced by several factors, mentioned in the chapter *2.1.3.4 Confounding factors in tactile* perception. For each technology, these factors have to be taken into account when designing experiments. Pre-tests, especially with elderly participants, are crucial to generate a valid stimuli set.

### *3.3 Methods of subjective assessment*

There are different methods of subjective assessment, both in the tactile and in the auditory domain, which are used quite frequently. Those different methodologies are applied to extract the layers of perceptual dimensions. In the following part the methods *interview* with a focus on the usability context, *pairwise comparison, semantic differential, similarity estimation* and general *questionnaire* are discussed, as they are commonly used in tactile and auditory settings. Their advantages and disadvantages will be discussed in the following paragraphs. After this, the method *Repertory Grid* is introduced and discussed.

### 3.3.1 Interviews in the usability context

Interviews are a common procedure in psychological research. They can be conducted via phone, mail, face to face, or online. Questions and answers can be fixed (structured), fixed questions with open answers can exist (semi-structured), or both questions and answers can be open. With increasing flexibility the amount of data analysis increases rapidly. Open interviews can vary their flexibility as well, from fixed contents in guided interviews (*Leitfadeninterview*), fully open interviews, to narrative interviews, the latter allowing the interviewed subject to choose focus, pace and content of the interviews (Gläser & Laudel, 2006). In User Centered Design, expert interviews are a common practice in an early development stage. Expert evaluations are characterized inasmuch as the subjects have to have a high knowledge base in a certain field of interest. Another way of describing expert interviews in a technological domain is as heuristic evaluation with expert subjects (Nielsen & Molich, 1990). As a rule of thumb, three to five experts take part in an expert evaluation (Jeffries & Desuvire, 1992). Interdisciplinary knowledge is usually recommended (Nielsen, 1992). Apart from this user interviews are far less frequent, as the amount of users necessary for interviewing is higher than those of experts and lay subjects may encounter difficulties imagining something they are not aware of to the date of the interview.

Usually interviews are conducted in three phases: a warming up phase, guiding to the topic; the evaluation phase, which depends on the topic; and a closing phase. The evaluation phase starts with a low level of detail and increases the granularity depending on topics relevant for the interview partner (Schmidt, 2005). This leads to a high amount of qualitative data, which can be evaluated according to *Qualitative Inhaltsanalyse* (Mayring, 2000) or *Grounded Theory* (Dilger, 2000). Unfortunately, these heuristic expert evaluations often express only subjective opinions of the experts (Sarodnick & Brau, 2011). The probability of loosing sight of difficulties real end users might have with a system is therefore rather high. As a result this kind of interview should not be employed as a substitute for user testing, but rather as a filter method to screen a technology for obvious improvement possibilities in terms of usability. Other methods should be enclosed in order to complete the usability evaluation of a product.

### 3.3.2 Pairwise comparisons

The pairwise comparison has previously been used in preference testing for switch settings (Hoggan et al., 2008; Koskinen, Kaaresoja, & Laitinen, 2008). A set of different stimuli is chosen and each stimulus is tested against each other in terms of user preference. If a user cannot differentiate or is not sure about the judgment, a tie can be declared. Ties can either be ignored, divided equally between both stimuli or a decision is made by chance (David, 1959). When it comes to psychophysics stimuli, tactile and auditory memories are limited, so one advantage in applying this method lies in the less complex setup and that it involves little mental effort. Another one is that they can be put in an ordinal order (David, 1959), which can be transformed via the *law of comparative judgment* (Bortz & Döring, 2006) into metric data. The accordingly performed calculation of *consistency* and *concordance* can enhance the internal and external validity of the ratings (Bortz et al., 2008).

A few drawbacks apply. The pairwise comparison is based on the assumption that subjects base their decisions on one criterion (Bortz et al., 2008). Pure preference ratings do not imply the dimension on which the rating took place, which is especially critical when it comes to stimuli that are not one-dimensional. Although it can be detected via consistency calculation when subjects changed their decision criterion, and inconsistent subjects can be excluded from further analysis, this approach is debatable when actually multimodal settings are varied in a comparison. Depending on the masking and the dominance of the different modalities over each other in different settings, it may be well reasoned to change the decision criterion. Subjects which respond to these changes by changing their criterion accordingly may appear inconsistent and may be eliminated from the data set although they were giving far more accurate information. The pairwise comparison method does not inform about perceived characteristics, but only about preference within the stimuli set. The generality of the preference cannot be assumed this way. As the result is a preference order, there is no qualitative information available, and it remains unclear which aspects of the stimuli are preferred. Although the mental effort for the subjects is kept low, it also denies them the possibility to conduct deeper investigation of the stimuli. In spite of the reduction of mental effort, the pairwise comparison can be very time-consuming when conducted exhaustively depending on the number of stimuli.

### 3.3.3 Semantic Differential

Semantic differentials rely on the subjective ratings of materials on scales with opposing adjectives as their end-points (Osgood et al., 1957). Test subjects rate different objects on these scales, and the factors are determined as combinations of adjective words, that describe the materials. Okamoto, Nagano, & Yamada (2013) characterize the semantic differential based on their research as one of the most commonly used methods in tactile research. A problem with the evaluation of the subjective impression on a technological device with this method is that most questionnaires already provide the criteria with which the subjects should evaluate. These criteria may be inaccurate or unfitting in the subject's perspective and there has been evidence that this can lead to problematic data (Susini et al., 2011). One limitation of this method therefore lies on the strong reliance on an appropriate set of words. Additionally, when employing a semantic differential opposed to e.g. a pairwise comparison between different stimuli, an order or rank cannot be established, as each item is rated on its own. A possible countermeasure for that would be to integrate a valence dimension such as *good - bad* into the semantic differential to allow for comparability. In the original application of the semantic differential, Osgood et al. (1957) found three main dimensions on which the pairs loaded via factor analysis: evaluative (e.g. named *good – bad*), power (e.g. *strong - weak*) and arousal (e.g. *active – passive*). The evaluative factor seems to have the most powerful loading (Kamps, Marx, Mokken, & de Rijke, 2004). In a technological context the usage of a semantic differential is often focused on limited aspects of a technology. If semantic differentials are more tailored to their context, their generalizability and application in other domains usually suffers. For example, the questionnaire developed by Wellings et al. (2009) for mechanical switches, seems to be a valid instrument for the intended technology, but when used with the present, different technology, subjects had difficulties in estimating on these scales (cp. study described in chapter *Frequency and waveform*). Generally, a semantic differential leads to more information of the quality aspects than other methods (Fastl, 2006b). If this is intended, the applicability of the semantic differential used to the target technology has to be ensured.

### 3.3.4 Similarity estimation

The goal of this method is not to derive direct qualitative aspects of an object, but rather to understand which concepts are more closely related than others. For doing so, two stimuli are presented at the same time. Subjects then rate which concepts are similar or dissimilar. This can be conducted either on a questionnaire (Bortz & Döring, 2006) or with other distance methods, e.g. by allowing subjects to freely choose the distance between two post-its on a desk. This can only be conducted for a small number of stimuli, as the presentation of pairs has to include all possible pairings. This limitation can lead to less quality in the examination of the relation space between different stimuli, but is less time consuming than other methods. The similarity estimation has the advantage that it allows subjects to rate stimuli not apart from each other and on their own, but in their relationships to each other. Instead of a semantic differential, this gives more information about the context than of isolated quality criteria. On the other hand this method does not provide the opportunity of retaining an absolute rating of one setting, which may be desirable sometimes. The acuteness of the ratings is highly dependent on the subjects used in the study (Okamoto et al., 2013).

### 3.3.5 Questionnaires

Discussing the different characteristic options for a questionnaire would go beyond the scope of the present work; for a detailed overview see Bortz & Döring (2006). Some aspects which refer to the employment of the method in the automotive context are discussed in the following. Questionnaires are usually written replies to a certain topic. This method is cost- and time effective. Although it provides subjective insights to a topic, it is limited in terms of the fixed answering categories, which can be chosen by the conductor of the evaluation in regard to the topic (cp. Rohrmann, 1978; Vagias, 2006). Several questionnaires in the automotive domain exist for evaluating a signal. Most of them aim on the evaluation of system interaction, not on a microinteraction like a single switch activation. If questionnaires are more specified, their generalizability and application in other domains usually suffers, even more than in the case of the semantic differential, as with questionnaires, even the anchors influence the type of interpretation of the question categories. If the goal is to evaluate an overall system impression, with a large number of subjects, questionnaires are a valid method. If the focus is on specific qualities and aspects of a technology on a very granular basis, employing this method may lead to the loss of essential information, unless the questionnaire is a highly specified one.

### 3.3.6 Repertory Grid

The Repertory Grid technique (Kelly, 1955) is a method that aims at overcoming the gap between the researcher, who conducts a study on a topic, and the subject which is object of study by allowing the researcher closer insights into the view on a specific topic by accessing the view of a subject. The original domain of Repertory Grids was personal construct theory; however, its application has nowadays spread from this to organizational psychology, marketing, system design and product design. The Repertory Grid technique is a very profound method that aims at investigating an area between qualitative and quantitative research. As this is one of the goals of the method development in the present work, this method is discussed in more detail.

A Repertory Grid consists of a fixed set of elements of study (e.g. different kinds of cheese), bipolar constructs which subjects build their evaluation upon (e.g. "soft vs. hard", "intense vs. unobtrusive", etc.), and ratings, which are usually on a five- or seven point scale (Jankowicz, 2003). A topic is also often noted as being part of a Repertory Grid, which means that the subject has to be interviewed about something, that is actually part of his or her experience space, but this should be rather a matter of course in user studies. Typically, the Repertory Grid procedure is divided into five steps: topic definition, element choice, construct generation, rating and data evaluation.

Usually the topic is chosen by the researcher in charge. Two criteria apply here: the subjects need to be familiar with the topic and the broadness of the topic needs to be adequate (Fromm (1995), cited after Hemmecke (2012)). When elements are chosen, the number should be appropriate for the set as well as for the subjects. At least six different elements are usually required (Scheer & Catina, 1993). Too many elements may lead to confusion or fatigue of the subjects (Bell, 1990). The choice of elements crucially determines content and quality of the grid (Haritos, Gindidis, Doan & Bell, 2004; Neimeyer & Hagans, 2002; Bell, Vince & Costigan, 2002; cited after (Hemmecke, 2012)).

As all possible combinations cannot be used in one testing, the random selection of element groups is the usual approach (Scheer & Catina, 1993). It has to be kept in mind, especially when several factors are varied in one study, that the representativeness of the stimuli is ensured, which would advise a mixed approach between random selection and a balanced design. Elements should be homogenous (Easterby-Smith, 1980) and discrete (Stewart, Stewart, & Fonda, 1981), which means that each of them has to be a distinguishable object. These are important aspects when it comes to measuring perceptional aspects and constitutes the main challenges, when employing Repertory Grids in usability research.

In terms of sample size, 15-25 subjects are usually enough to cover all possibly producible constructs (Dunn, Cahill, Dukes, & Ginsberg, 1986). In addition, when the purpose of initial Repertory Grids is to construct other measurement techniques, such as semantic differentials, even smaller sample sizes can be used as well (Tan & Hunter, 2002).

There are different possibilities in eliciting constructs from subjects, depending on the number of elements used in one generating iteration. Firstly there is the possibility of presenting subjects with already present constructs made up by the experiment conductor (Tan & Hunter, 2002). This would resemble very much like a semantic differential rating and is usually not the purpose of a Repertory Grid. Secondly, by employing three elements and asking what the first two have in common compared to a third, *triads* are used; when the difference between two elements is evaluated only, the term *dyads* is applied. Dyads are commonly used when the chosen subjects require simplification or under time pressure (Rosenberger, 2015). Both dyads and triads refer to the minimum context form, meaning that the elements are pre-ordered. When all elements are present at the same time, the method is described as maximum context form. The latter is applied far less often and the results of employing it seem to be debatable (Schmitt & Altstötter-Gleich, 2010). A mixture of both supplied and elicited constructs can also be used (Easterby-Smith, 1980), e.g. when constructs for similar elements are known from previous studies.

When generating a grid via triads or dyads, the similarity pole is referred to as the *construct pole*, whereas the dissimilarity pole is referred to as the *contrast pole* (Schmitt & Altstötter-Gleich, 2010). There are two distinct methods for eliciting a grid: the *differentiation method* and the *opposition method* (Scheer & Catina, 1993). In the latter subjects generate the contrast pole by building up the semantic opposite of the construct pole; in the first one the contrast pole is elicited via the third object. Subjects are asked something like “what have objects 1 and 2 in common compared to object 3?” The goal of the opposition method is to lead to clear atonyms, as subjects have to refer to a verbal opposite. This leads away from the centering of the method around the present prototype and carries the danger of rather evaluation language capabilities than the prototype. The presentation of the elements for grid generation should be balanced, so that every element is present equally often (Scheer & Catina, 1993). Presupposed an appropriate number of elements present, it is not possible to present all possible combinations to a subject (the formula for all possible combinations equals $\binom{n}{3}$). It is recommended not to put subjects under time pressure, as some of the constructs build up only during the testing (Rosenberger, 2015). Analyzing the qualitative verbalized responses allows drawing conclusions to non-verbalized constructs, which are not the same as the verbalized responses, but their latent variable (Fransella, Bell, & Bannister, 2004). These non-verbalized constructs could be accessed e.g. via principal component analysis.

When linking constructs to elements, a rating procedure is often employed (Tan & Hunter, 2002). Both even and uneven numbers can be used as response categories and there are arguments for both solutions (Bortz & Döring, 2006). With even scales, subjects are forced to give an information about a tendency in one direction; uneven scales allow for a neutral response. The best retest-reliability values seem to be found when employing a five point scale (Bell, 1990) for a Repertory Grid rating.

There are different options for compiling the results of a Repertory Grid. Jankowicz (2003) provides three alternatives: describing single grids, analyzing relationships in single grids and analyzing multiple grids. The latter is the main focus of the present work. When analyzing the grids, they can be summed up and evaluated after the Procrustes-Method, which is a statistical shape analysis, or some form of adaptation of it (Werz, 2006). However, this method only leads to valuable results if the constructs are drawn from a fixed pool and not completely open to be generated by subjects. If the latter is the case, inductive compilation of the results can be applied to qualitative data (Mayring, 2000). This has the advantage that the coding of the material is driven by the context, but follows a standardized procedure to ensure the quality of the data. Usually the compilation takes place via a simple frequency count of the resulting constructs (Hunter, 1997). There is, however, a slight variance of this procedure present: Honey (1979) proposed to use different coders and statistically investigate the inter-rater-reliability. Another method to minimize information loss during the agglomeration of grid data is to include an *overall-construct,* which allows for drawing conclusions to the generated constructs by being set in relation to them at the end of each survey (Jankowicz, 2003). This method is aiming again at personality psychology and it can be assumed, that the emerging constructs as well as the overall constructs are much broader with personality psychology than in the present case of switch interaction.

The agglomerated data can then be evaluated further, investigating underlying factors. This is often conducted via exploratory factor analysis or principal component analysis (PCA) (Tan & Hunter, 2002). Employing PCA allows classifying the grid data in terms of single- or multidimensionality (Bell, 2004). Most Repertory Grids are sufficiently explained by two or three (Bell, 1990) or four (Riemann, 1983) factors, which usually results in 80% - 90% explained variance. All authors report that more factors can be developed as well, depending on the topic. An increasing number of factors points into the direction of a cognitively more complex topic (Bell, 2004).

### *3.4 Method selection and adaptation*

The Repertory Grid approach has a major flaw: it is hard to compare different grids, as each grid is made up by a subject, based on his impression of a topic. It is an aim of the proposed work to classify the constructs based on content analysis approach and derive an instrument which is formally a semantic differential, but constructed via users' impressions of virtual switches. It is also pursued that the first answer categories will remain free and individually to choose by each user. The result will therefore partly stay a grid. Because of the nature of tactile feedback, being stored only in working memory, and very hard to describe for most subjects, for almost every device it will make more sense to directly compare at least two elements, than only evaluating one on its own. This can be employed via building up new constructs for every new characteristic under investigation, additionally to the ones previously generated.

The aim of the employment of the Repertory Grid in the present work is to investigate whether this is an appropriate tool to investigate the acceptance of a virtual switch. Several adaptations of the method need to be conducted in order to meet this requirement.

In terms of elements, it is obvious that subjects need to interact with a virtual switch and evaluate it based on its characteristics. Several studies will therefore be employed with different aspects of the elements being varied. The total number of elements will vary for each study and for pre-studies, where grid generation is crucial, kept over six elements. In the chapter Perception thresholds, interactions between amplitude, vibration and frequency in subjective tactile signal evaluation support the assumption that subjects rather integrate the different aspects in their rating, which would support treating variations of different dimensions still as comparable elements.

It is not clear whether triads or dyads should be employed, as dyads may be appropriate for people who encounter perceptional difficulties. This will be varied across the pre-studies in order to see which method leads to more constructs per varied aspect. Dummy comparisons (with no objectifiable difference in contrast to e.g. different frequencies) will be used in one pre-study in order to investigate whether subjects draw opposite constructs from comparisons where there is no physical dimension to base their judgment upon. Element combinations will be drawn randomized, but the randomized orders will be balanced in order to make sure that all subjects are allowed to judge all items. As the focus of the present work is not on semantic opposites, the differentiation method seems to be more promising in the present context and will be employed. Both native English speakers as well as German speakers will be participating order to improve the transferability between both languages. Once generated constructs will be carried on to subsequent studies, as a) the topic remains the same and b) constant as well, in the way that if one setting performs better in the evaluation, its characteristics will be reused. This has the additional advantage that it allows for statistical evaluation of the established constructs. New elements will nevertheless be presented in every study in order to enlarge the construct pool. Additionally, two overall constructs will be added to the Repertory Grid technique: a valence scale consisting of the item *good – bad*, as this seems to be a promising approach from the chapter *Material design* and an ideal configuration at the end of the rating of the elements on the constructs, in order to see which configuration subjects prefer on the generated items. This ideal configuration can then be compared to the ratings of the different elements.

In terms of quality criteria, the discussion remains somewhat diffuse. A detailed manual will be generated for each study in order to improve *objectivity*, as this is a sufficient measure of process objectivity (Bortz & Döring, 2006).

As Fransella et al. (2004) state, the utility should be preferred over the *validity* of a Repertory Grid evaluation. This is a somewhat unsatisfying perspective when the goal is to make this data statistically quantifiable across grids. To enlarge ecological validity, it is aimed at conducting most studies in a driving simulator. Several studies have investigated the ecological validity of driving simulator studies (for an overview, s. Blana (1996), and the literature points in the direction that validity needs to be investigated task-specifically. Compared to testing on a desk, the ecological validity in a driving simulator increases, although one past study with the present technology did not find differences between the sitting and the driving condition. It would be optimal in terms of validity to evaluate in naturalistic driving studies, which goes beyond the scope of the present work. In order to investigate in the specific context, driving simulator data can be used to support *concurrent validity*, as well as error rate and time on task for the switches as a task during driving. For the first pre-study, an additional concurrent method, the pairwise comparison, is used to investigate this type of validity. In the main study, another concurrent method will be employed: the semating differential by Wellings et al. (2009), with the valence adaptation by Vogelpohl (2014) discussed previously.

*Reliability* is a necessary, but not sufficient precondition of validity: a method can be reliable, but not valid.This is not possible the other way around. A typical measure for reliability is the internal consistency or Cronbach's α, which will be investigated as soon as the data can be handled quantitatively. As Cronbach's α may not always be appropriate depending on the data, corresponding methods will be employed.

In terms of analysis, the inter-rater reliability cannot be checked due to economic reasons. It is feasible to identify potential interpretation gaps and cross-check with native speakers, as well as checking synonyms and antonyms in a dictionary.

## *3.5 Hardware*

### 3.5.1 3D Switches steering wheel

The switch sensors and actuators were mounted into a steering wheel for simulator integration; the switches were mounted in a small retainer with a polycarbon switch surface for interaction. The system which was evaluated in the two studies discussed in this paper has a tactile response time under 30 ms after the surface application of force. As human tactile perception averages at about 155 ms in experiments (Robinson (1934), cited after Kosinski (2012), this latency is sufficiently small and can be considered as instantaneous feedback. From a responsiveness perspective, such a system can deliver a tactile haptic response from the true application of force in 30 ms, which for human perception can be considered instantaneous. In the present study an advanced prototype was evaluated. According to Bjelland & Tangeland (2007), prototypes can be varied along five axis, namely:

- Physical ↔ Virtual
- Low-Fidelity ↔ High Fidelity
- Vertical ↔ Horizontal

- Exploratory ↔ Experimental
- Low user involvement ↔ High user involvement

This prototype would be a physical and high-fidelity prototype, as the subjects can directly interact with it and it is very close to a normal multifunctional steering wheel which people actually use in their car. It covers most features, but not to their full functionality, which makes it a horizontal prototype. Horizontal prototypes give an insight to the interface as a whole, but do not give any information about the underlying functionality (Nielsen, 2010). As the underlying functionalities are highly OEM-specific, testing for certain functionality would provide less generalizable results than testing for the overall settings. Because system specifications will be developed with it, it is an exploratory prototype and, as a consequence, it contains a rather high amount of user involvement.

### 3.5.2 Signal generation

Tactile signals were generated using an audio file modification software. This software allows for choosing among a range of different frequencies and choosing the addition of an audio file. Amplitudes were measured with different devices, and their setup and the acceleration measurements are discussed previously to each experiment. When signals were measured in $m/s^2$, the measured signals were transformed into the psychophysiological measurement of dB in reference to 1 micron using the formula

$$\frac{\mathrm{m/s^2}}{2\pi^2 * \upsilon^2 * 1000}$$

where $\upsilon$ is the frequency of the signal. This results in the displacement in mm. To convert this to dB in reference to 1 micron, the result has to be multiplied by 1000 and a logarithm with the factor 20 has to be applied to this product. The result is a displacement value in dB in reference to 1 micron, which can be compared with other psychometric vibrotactile studies, as well as held constant across different prototypes and studies.

The auditory signal of the technology was allocated at 2000 Hz, in order to avoid frequency confusion between auditory and vibrotactile signals.

### 3.5.3 Driving simulator

The simulator experiments were conducted in the driving simulator of TAKATA AG in Berlin.

*Figure 8. Driving simulator.*

This simulator has a fixed base. Its dashboard, seat and center stack was built up from a BMW 3 series cockpit. The transmission is automatic; therefore the simulator has neither a clutch nor a gear stick. Gas and brake pedal are of consumer electronics origin: they are add-ons for a Logitech G25 racing wheel for PC racing games. The physical mock-up is made of a ¼ part including driver's seat and full dashboard of a BMW 3 series (cp. Figure 8. *Driving simulator*).

Three projection screens provide the driver with a 117° horizontal and 47° vertical field of view, a resolution of 1280 x 720 pixels and a pixel density of 17.5 x 17.3 ppi. The image rate of the projection is 60 Hz. The time lag between driver input and simulator output varies between 110-136 milliseconds.

Main parts of the driving simulation are vehicle dynamics, the image generation for the driving scenarios, the simulation of other traffic participants, and the six-channel sound simulation. The vehicle dynamics are those of a BMW 5 Series. The communication interface with other hard- or software works via CAN, but can also be adjusted to TCP/IP. Scenarios can be designed individually and the data recording is very flexible but accurate. The software for running the driving simulation is SILAB from the Würzburger Institute for Traffic Sciences (Würzburger Institut für Verkehrswissenschaften GmbH).

### *3.6 Preliminary studies*

#### 3.6.1 First preliminary experiment: frequency and amplitude of the tactile feedback signal

The goal of this study was to broaden the frequency range of the signals investigated, and to analyze whether the preference differs in terms of the amplitude. In this testing a Repertory Grid was constructed with the subjects to define characteristics on which virtual switches could be evaluated.

##### *3.6.1.1 Method*

16 healthy subjects, all TK Holdings Inc. employees (15 at Pontiac, 1 at Auburn Hills), either native English speakers or German expatriates, participated voluntarily in this study. They were equally distributed in terms of gender and age groups, which were chosen correspondingly to NHTSA guidelines (Garrett, 2012). This study did not take place in a driving simulator, but in a lab setting. The hypotheses were

$H_1(1)$: There is a difference between the switch feedback and amplitude preferences for the different stimuli.

$H_0(1)$: There is no difference between the switch feedback and amplitude preferences for the different stimuli.

Three different frequencies were used in this experiment: 230 Hz, because this frequency emerged as the most preferred one in a previously conducted study; 300 Hz, because in the mentioned previous study, not the full range of possible frequencies was examined; and 175 Hz, because the adding of a higher frequency might change the preference rating of other stimuli. Two amplitudes were used in this experiment. Amplitudes were recorded using a custom USB driven high power frontend. The sensor was a Model M352C65 sensor by Platinum Stock Products. For 5 – 255 bit in amplitude, the vibrotactile acceleration was measured in steps of five. The peak-to-peak amplitude for each initial sine for each measurement point was used to determine the displacement as discussed in the chapter *Psychophysiological measurement of tactile perception.* Two amplitudes, high and low in intensity, were chosen according to JND calculation. A pretest with n = 3 made sure that the signals were differentiable.

The factorial design of the experiment is described in the following table (cp. Table 4. *Factorial design of the first preliminary experiment*).

The testing device was not mounted in a steering wheel, but lay flat on a desk in the aforementioned mount. White noise was played through noise-cancelling headphones during switch evaluation to make sure subjects tested the tactile feedback only.

*Table 4. Factorial design of the first preliminary experiment.*

| Frequency | **High feedback** | **Low feedback** |
|---|---|---|
| 175 Hz | $\bar{Y}_{11}$ | $\bar{Y}_{12}$ |
| 230 Hz | $\bar{Y}_{21}$ | $\bar{Y}_{22}$ |
| 300 Hz | $\bar{Y}_{31}$ | $\bar{Y}_{32}$ |

The subjects filled out a consent form at first and after that, room and finger temperature of the right hand were measured. The next step involved the pairwise comparison as discussed in chapter *Frequency and waveform*, and afterwards the Repertory Grid construction and rating. For generating Repertory Grids, two aspects were held constant (either amplitude or frequency level) for the first two elements, opposed to a third setting and were thus used to generate a Repertory Grid via triads. A variation between all factors was included to see whether there would be other constructs that subjects found which relied not on one single factor only. Each signal was then rated separately on the constructed Repertory Grid, which contained the addition of an overall *good – bad* scale for valence evaluation. At the end of all ratings, subjects were asked to imagine their ideal configuration on this grid and rate it accordingly. Only the single-push function was evaluated. The response anchors corresponded to those developed by Vagias (2006).

#### *3.6.1.2 Results of the first preliminary test*

##### 3.6.1.2.1 Pairwise Comparison

According to the aforementioned procedure, the consistency of the sample was calculated. 37% of the sample answered inconsistently, whereas 63% answered consistently. There were no age or gender effects in the consistency of the sample. In terms of concordance, the overall group answered not concordantly with an *Akkordanzmaß* A = -0.006683375 and a $\chi^2$ (39) = 17.11176041, $p = 0.4$. When further analyzing the data, it could be found that there were two distinct preference groups which accounted for the lack of concordance.

The first group, after z-transforming the values according to the *Law of comparative judgement*, preferred factor variation $\bar{Y}_{21}$ over factor variations $\bar{Y}_{12}$, $\bar{Y}_{22}$, and $\bar{Y}_{32}$, therefore disliking the lower amplitude. This could be further supported by statistical analysis. The data was normally distributed with the Shapiro-Wilk normality test W = 0.973, p = 0.5126, and the homogeneity of variances was given via Levene's test F (5/30) = 0.4518, p = 0.8086, $\omega^2 = 0.24$. Consequently an ANOVA was calculated and became significant with F (5/30) = 3.31, p = 0.0169. A Benjamini–Hochberg corrected post-hoc pairwise t-test revealed the following significant differences between aforementioned factor variations: $\bar{Y}_{21}$ to $\bar{Y}_{12}$ (p = 0.046), to $\bar{Y}_{22}$ (p = 0.043), and to $\bar{Y}_{32}$ (p = 0.043).

The second group, after z-transforming the values according to the *Law of comparative judgement*, preferred factor variation $\bar{Y}_{22}$ over $\bar{Y}_{11}$. This could be further supported by statistical analysis. The data was normally distributed with the Shapiro-Wilk normality test W = 0.9615, p = 0.2393, and the homogeneity of variances was given via Levene's test F (5/30) = 0.4518, p = 0.5436. Consequently, an ANOVA was calculated and became significant with F (5/30) = 3.52, p = 0.0127, $\omega^2 = 0.26$. A Benjamini–Hochberg corrected post-hoc pairwise t-test revealed in the following a significant difference between $\bar{Y}_{22}$ and $\bar{Y}_{11}$ (p = 0.0042).

###### 3.6.1.2.2 Repertory Grid Content Analysis

The content analysis of the Repertory Grids followed the procedure proposed in *Method selection and adaptation.* For the rating of the self-constructed Repertory Grid, all generated items were analyzed using a quality criterion of 25%. This procedure ensures that a minimum of consensus between the subjects on the evaluated items is present. A total of 40 items were generated, of which the eight remaining met this quality criterion. The emerging opposite pairs were *faint – predominant, smooth – rough, short – long, strong – weak, pleasant – clicky, slow – fast, hard – soft* and *mechanical – buzzy.*

As in the grid generation the triads were controlled for whether they were constructed on the frequency or the amplitude domain, two items were created in respect to frequency, three items to amplitude, and the remaining three were created with both. This is only a qualitative conclusion and cannot be support via statistical procedures.

#### 3.6.1.2.3 Repertory Grid Semantic Differential Analysis

The emerging items were treated as a semantic differential and evaluated via profile analysis. There were no significant differences of the different settings in terms of the good-bad scale: an ANOVA did not become significant with $F(5/90) = 0.492$, $p = 0.782$, $\omega^2 = 0.03$. The data structure is, taken as a sample, full of missing values due to the generation procedure; a statistical analysis does therefore only make limited sense in terms of scale evaluation.

The values were recoded and normalized according to the *Law of comparative judgment* in order to evaluate their absolute distance to the ideal scale. The factor variation $\bar{Y}_{21}$ seemed to be the closest one to the ideal in this calculation, whereas the factor variation $\bar{Y}_{22}$ was allocated rather in the middle and between other factor variations. Further data analysis was disregarded due to the data structure.

#### 3.6.1.2.4 Discussion

A set of 8 opposite pairs could be generated in this study. Some of these items match the ones proposed by Lederman & Klatzky (2009) or Wellings et al. (2009); others, such as *mechanical – buzzy*, are genuinely new and possibly due to the new characteristics of the present technology. For valence analysis and further confirmation, these items have to be evaluated in further studies. The addition of the *good – bad* scale revealed no additional results; this might be due to the fact that no extreme stimuli were chosen in the evaluation. The distance to ideal seems in contrast to be a promising approach in order to evaluate the valence of switch settings, but two caveats must be taken into account: first, this is subjective data and therefore prone to various unintended mistakes of the subjects, and second, subjects are usually not always good in determining their "ideal" configuration of something. Generally, the pairwise comparison and the Repertory Grid pointed in the direction of favouring a 230 Hz signal. The order between the different factor variations was different across both methods, but the extremes were congruent. The order of the factor variations was not statistically significant and could therefore be different by chance in the two different methods.

### 3.6.2 Second preliminary experiment: force settings of the skip and mode switches (driving simulator)

The purposes of the secondary preliminary experiment was to inspect the items generated in the previous study in terms of their underlying factors, to investigate on the interaction of tactile feedback and different force settings in terms of preference, effectiveness and effectivity, and to generate additional items via Repertory Grid for the questionnaire.

#### *3.6.2.1 Method*

In total 39 healthy subjects participated voluntarily in this study. They were equally distributed over the age groups recommended by NHTSA and equally distributed over gender, with the exception of the female age group 55+, where one recruited subject on the last day of the study did not show up. There was no possibility for additional recruiting as the driving simulator was already scheduled for another study the following day. The hypotheses were

$H_1$(1): There is a difference between the different switch feedbacks and force preferences.

$H_0$(1): There is no difference between the different switch feedbacks and force preferences.

$H_1$(2): There is a difference in performance between the switch feedback and force preferences.

$H_0$(2): There is no difference in performance between the switch feedback and force preferences.

The force settings for this study have been calculated in JND steps between 2.0 N and 6.5 N, as this is a range commonly used in automotive specifications. One recent experiment (Doerrer & Werthschuetzky, 2002) reported decreasing JND between 21.9 % (0.5 N) and 5.49 % (2.5 N). In Allin, Matsuoka, & Klatzky (2002) the JND was relatively constant at 10 % when increasing force from 2.25 N on, and another research team (Tan et al., 1992) reports a JND that lies between 5% and 10% for pinching motions between finger and thumb with a constant resisting force. This JND was found to be relatively stable over a range of different base force values between 2.5 N and 10 N. Therefore the JND of 10% was taken as a basis for JND calculation. As this would result in far too many stimuli for one study, four settings were selected. A pre-test with three subjects belonging to the 55+ age group was conducted in order to make sure that the different force values were distinguishable not only via JND calculation.

The feedback used was a pure 230 Hz sine wave, with low and high feedback amplitude, as described in the previous experiment. Amplitudes were recorded using a USB driven high power frontend by Head Acoustics, Inc., Octobox+. The sensor was a Model M352C65 sensor by Platinum Stock Products. For 5 – 255 bit in amplitude, the vibrotactile acceleration was measured in steps of five. For measurement setup, cp. Figure 9. *Acceleration measurement setup.* The peak-to-peak amplitude for each initial sine for each measurement point was used to determine the displacement as discussed in the chapter *Psychophysiological measurement of tactile perception.* Two feedbacks with a higher and lower value in dB in reference to 1µm were selected because of their conformity to the amplitudes of the stimuli in the previous study.

*Figure 9. Acceleration measurement setup.*

The factorial design of the experiment is described in the following table (Table 5. *Factorial design of the second preliminary experiment*).

*Table 5. Factorial design of the second preliminary experiment.*

| Forces | **High feedback** | **Low feedback** |
|---|---|---|
| Low | $\bar{Y}_{11}$ | $\bar{Y}_{12}$ |
| Medium low | $\bar{Y}_{21}$ | $\bar{Y}_{22}$ |
| Medium high | $\bar{Y}_{31}$ | $\bar{Y}_{32}$ |
| High | $\bar{Y}_{41}$ | $\bar{Y}_{42}$ |

As a test setting, the driving simulator described in the chapter *Driving Simulator* was used, but only with front projection in order to avoid noise issues which would have occurred with several projectors instead of one.

Subjects first filled out a consent form, then room and finger temperature of the right hand was measured. They drove a familiarization drive of seven minutes on a rural road with gentle curves. This was done in order to get a hold on subjects suffering from simulator sickness previously to the start of the real experiment, as simulator sickness usually occurs in the first minutes of a drive and does not improve or vanish over time (Lee, Younghak, & Jones, 1997). The last two minutes served as a baseline for the driving data in the data analysis. Subjects then filled out a NASA-TLX (Hart & Staveland, 1988) and then a demographic questionnaire (appendix: a) a *Second preliminary experiment Demographic questionnaire*), which included the valence ("pleasure") and the arousal scale of the SAM (Bradley & Lang, 1994), but not the dominance scale, as the validation of this scale is still debated. For Repertory Grid generation, two settings were held constant on force level opposed to one other setting and were thus used to generate a Repertory Grid via triads; a variation between all factors was included to see whether there would be other constructs that the subjects found, which relied not on one single factor only. This Repertory Grid was generated both for the single - as well as the push and hold - function and can be found in appendix a) b *Repertory Grid*.

Then subjects rated the different stimuli on a five point scale, on the newly generated item pairs and on the ones that were generated in the previous study. The response anchors corresponded to those developed by Vagias (2006). The subjects were driving when they performed the evaluation. At the end of the rating of the different switch configurations, subjects rated on the questionnaire how their ideal configuration would look like. After this they conducted a performance task: For the skip function, forward and backward skip were used; subjects were instructed to press and hold the switch for five activations forwards and for two activations backward. The different activations were fed back to the subjects via the tactile feedback. For the mode function, subjects were instructed to activate and deactivate the switch four times as fast as possible. This was due to the fact that for subjects with smaller hands, re-gripping might be necessary and therefore the performance of the push-and-hold switch could not be transferred to the other. After the performance task the subjects filled out the NASA-TLX (Hart & Staveland, 1988) to rate the difficulty of the driving task in the performance condition.

The same procedure was conducted first either for single activation (mode switch) or push-and-hold (skip), and then for the other function modality. The order of the functions was balanced for the subjects and the order of the stimuli randomized and balanced across the different subjects. Subjects were instructed throughout the course to drive with 70 km/h per hour and to prioritize to the driving task when a secondary task was to be conducted.

#### *3.6.2.2 Results of the second preliminary test*

##### 3.6.2.2.1 Switch performance

In order to investigate whether there was a difference between the switch settings in terms of switch performance, efficiency and effectiveness of the switches were measured. The operationalization was conducted through time on task (efficiency) and error rate (effectiveness).

In terms of error rate a binomial test with 10% accepted error was conducted. For the calculation of the mode error rate 27 of 320 data points were missing, in the skip data set 88 of 320 data points. The missing values were not substituted for this calculation. For the skip backward function no setting passed the test criterion, which means that the task itself might have been too difficult to perform. For the skip forward and the mode function, the factor iterations $\bar{Y}_{41}$ (mode: $p = 0.035$, skip: $p < 0.001$), $\bar{Y}_{42}$ (mode: $p = 0.001$, skip: $p < 0.001$), and $\bar{Y}_{32}$ (mode: $p = 0.003$, skip: $p = 0.021$) did not pass the criterion. All other factors iterations were acceptable in terms of error rate.

In terms of time on task of the switch performance, there were no significant differences between the different switch settings.

For the mode function, the data was not distributed normally (Shapiro-Wilk-Test with $W = 0.41$, $p < 0.001$), but the variances were homogenous (Levene-Test with $F(266/7) = 0.96$, $p= 0.46$). As a result to this, a Friedman's ANOVA, the non-parametric equivalent of an oneway ANOVA, was conducted, and did not become statistically significant with $\chi^2 (7) = 10.64$, $p = 0.1552$.

For the skip function, the data was not distributed normally (Shapiro-Wilk-Test with $W = 0.79$, $p < 0.001$), although the variances were homogenous (Levene-Test with $F(296/7) = 0.61$, $p = 0.75$). Correspondingly, a Friedman's ANOVA was conducted as well, and did become statistically significant with $\chi^2 (8) = 32.71$, $p < 0.001$, $\omega^2 = 0.01$. A post-hoc Nemenyi test revealed no significant differences between a distinct pair of settings, and can be found in appendix a) d *Post-hoc Nemenyi results for the skip function*.

##### 3.6.2.2.2 Driving performance

In order to investigate whether there was a difference between the switch settings in terms of driving performance, efficiency and effectiveness were measured. Whereas effectiveness was measured via the driving parameters mean velocity, standard deviation of mean velocity, lane keeping quality, lane position, steering wheel velocity and steering wheel angle, efficiency was measured via the NASA-TLX between the baseline drive and the performance task.

To analyze the differences between the driving performance tasks, an oneway-ANOVA was conducted. The preconditions for an ANOVA were met with Shapiro-Wilk-Tests between the conditions being

- Baseline– Mode: $W = 0.96$, $p = 0.1841$
- Baseline– Skip: $W = 0.98$, $p = 0.6033$
- Mode – Skip: $W = 0.95$, $p = 0.08637$

and the Levene Test for variance homogeneity being $F\ (2/108) = 2.01$, $p = 0.1393$, and Mauchly's test for sphericity being $W=0.86$, $p = 0.07$. The ANOVA became statistically significant with $F\ (2/72) = 28.11$, $p < 0.001$, $\eta^2 = 0.15$. A post-hoc bonferroni corrected t-Test revealed differences between all conditions (cp. Figure 10. *NASA-TLX workload for the second preliminary test score across driving conditions*).

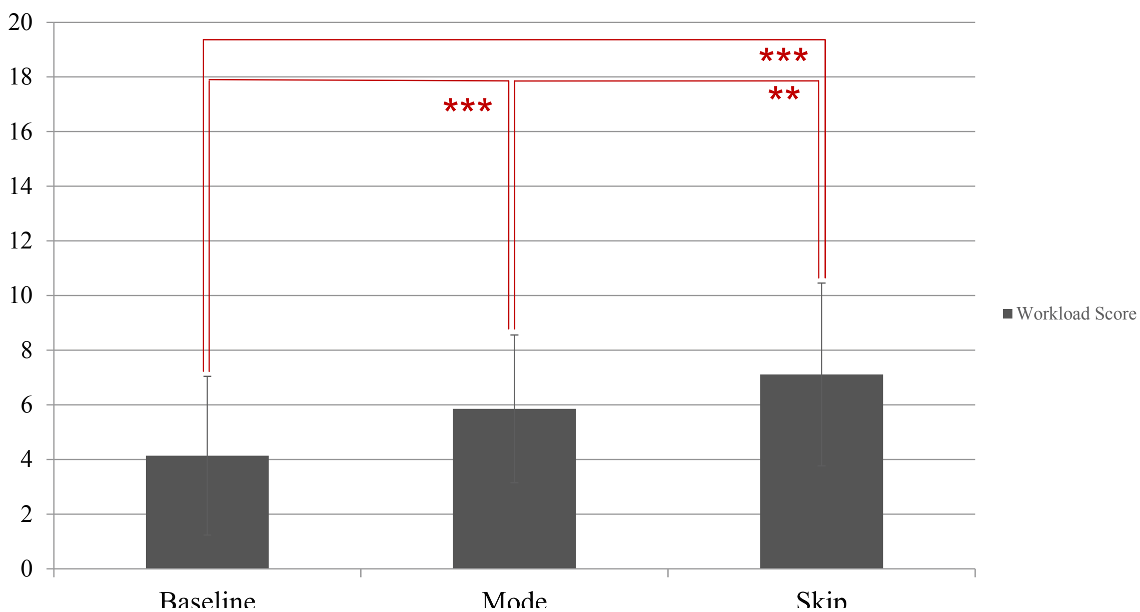



*Figure 10. NASA-TLX workload for the second preliminary test score across driving conditions.*

There were no relevant differences in terms of driving performance between the switches. Although there was one significant difference in terms of speed, both values deviated from the instructed speed of 70 km/h, one positively, one negatively, and were therefore discarded for the evaluation. Detailed results can be found in appendix a) c *Driving Performance Results.*

#### 3.6.2.2.3 Subjective preference data

In order to analyze the preferred switch configurations, the distance to the ideal configuration and the results of the good – bad – scale at the end of each of the switch setting questionnaire were measured.

For the ideal configuration comparison in the mode function, the sample was not normally distributed (Shapiro-Wilk-Test with $W = 0.94$, $p < 0.001$) and the homogeneity of variances was given (Levene's Test with $F(8/72) = 1.90$, $p = 0.0728$). A Friedman's ANOVA was therefore conducted and became significant with $\chi^2(7) = 53.42$, $p < 0.001$, $\omega^2 = 0.32$. A post-hoc Nemenyi test revealed highly significant differences to the ideal configuration for settings $\bar{Y}_{21}$ ($p = 0.05$), $\bar{Y}_{31}$ ($p = 0.00481$), $\bar{Y}_{41}$ ($p < 0.001$), and $\bar{Y}_{42}$ ($p = 0.00144$). An ANOVA was conducted as a backup and became highly significant with $F(8/72) = 5.75$, $p < 0.001$. A posthoc bonferroni corrected t-test showed significant distances to the ideal configuration for settings $\bar{Y}_{21}$ ($p = 0.027$), $\bar{Y}_{31}$ ($p = 0.00401$), $\bar{Y}_{41}$ ($p < 0.001$), and $\bar{Y}_{42}$ ($p = 0.002$).

For the ideal configuration comparison in the skip function, the sample was not distributed normally (Shapiro-Wilk-Test with 0.95, $p = 0.001921$) and the homogeneity of variances was not given (Levene's Test with $F(8/72) = 2.02$, $p = 0.0554$). A Friedman's ANOVA was therefore conducted and became significant with $\chi^2(7) = 50.66$, $p < 0.001$, $\omega^2 = 0.29$. A post-hoc Nemenyi test revealed highly significant differences to the ideal configuration for settings $\bar{Y}_{21}$ ($p = 0.00567$), $\bar{Y}_{31}$ ($p = 0.00408$), $\bar{Y}_{41}$ ($p < 0.001$), and $\bar{Y}_{42}$ ($p = 0.00120$). An ANOVA was conducted as a backup and became highly significant with $F(8/72) = 5.04$, $p < 0.001$. A posthoc bonferroni corrected t-test showed significant distances to the ideal configuration for settings $\bar{Y}_{21}$ ($p = 0.0075$), $\bar{Y}_{31}$ ($p = 0.0043$), $\bar{Y}_{41}$ ($p < 0.001$), and $\bar{Y}_{42}$ ($p = 0.0026$).

In terms of the differences on the good – bad – scale the mode data was not distributed normally (Shapiro-Wilk-Test with $W = 0.90$, $p < 0.001$) and homogeneity of variances was given (Levene's Test with $F(7/304) = 1.63$, $p = 0.1267$). A Friedman's ANOVA was therefore conducted and became significant with $\chi^2$ (7) = 53.79, p < 0.001, $\omega^2$ = 0.13. A post-hoc Nemenyi test revealed highly significant differences between the settings $\bar{Y}_{11,}\bar{Y}_{12}$ and settings $\bar{Y}_{41,}\bar{Y}_{42}$; a highly significant difference between setting $\bar{Y}_{12}$ and setting $\bar{Y}_{31}$ and a significant difference between setting $\bar{Y}_{22}$ and setting $\bar{Y}_{42}$. This was backed up by an ANOVA which was conducted and became significant with F (7/304) = 7.47, p < 0.001. A post-hoc Tukey HSD revealed highly significant differences between settings $\bar{Y}_{11,}\bar{Y}_{12}$ and settings $\bar{Y}_{41,}\bar{Y}_{42}$; a highly significant difference between setting $\bar{Y}_{12}$ and setting $\bar{Y}_{31}$ and significant differences between setting $\bar{Y}_{22}$ and settings $\bar{Y}_{41}$ and $\bar{Y}_{42}$.

The skip data was not distributed normally (Shapiro-Wilk-Test with W = 0.90, p < 0.001) and homogeneity of variances was given (Levene's Test with F (7/304) = 0. 28, p = 0.9614). A Friedman's ANOVA was therefore conducted and became significant with $\chi^2$ (7) = 35.15, p < 0.001, $\omega^2$ = 0.07. A post-hoc Nemenyi test revealed highly significant differences between settings $\bar{Y}_{11,}$ $\bar{Y}_{12}$ and setting $\bar{Y}_{41,}$ and a significant difference between setting $\bar{Y}_{22}$ and setting $\bar{Y}_{41}$. An ANOVA was conducted as a backup and became significant with F (7/304) = 4.46, p < 0.001. A post-hoc Tukey HSD revealed highly significant differences between settings $\bar{Y}_{11,}\bar{Y}_{12}$ and setting $\bar{Y}_{41}$; a significant difference between setting $\bar{Y}_{12}$ and setting $\bar{Y}_{31}$, and a significant difference between setting $\bar{Y}_{22}$ and setting $\bar{Y}_{41}$. The full post-hoc test data can be viewed in appendix a) e *Post-hoc Nemenyi and Tukey HSDs for good/bad rating*.

##### 3.6.2.2.4 Questionnaire evaluation

In the Repertory Grid generation, only one new item emerged from the present study: *responsive vs. sluggish* (*leichtgängig-schwergängig* in German). The questionnaire data was then evaluated in two forms: an item polarity analysis and a principal component analysis (PCA).

In terms of item polarity, each item was correlated to the good/bad scale at the end of the questionnaire. This was conducted for each setting to avoid the influence of the different switch configurations on the overall correlations. For both mode and skip functionality, the correlation data was not distributed normally with Shapiro-Wilk Test W = 0.88, p < 0.001. Because of that, a Spearman correlation was conducted for each setting with the good – bad – scale as a nonparametric correlation calculation. A summary of the results can be found in the table below (Table 6. *Correlations with good/bad scale*).

*Table 6. Correlations with good/bad scale.*

| | | Correlations with good/bad scale per setting | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Mode** | | $\bar{Y}_{11}$ | $\bar{Y}_{21}$ | $\bar{Y}_{31}$ | $\bar{Y}_{41}$ | $\bar{Y}_{12}$ | $\bar{Y}_{22}$ | $\bar{Y}_{32}$ | $\bar{Y}_{42}$ | **Correlation** |
| Faint | Predominant | 0.12 | -0.20 | 0.00 | 0.10 | -0.33 | -0.22 | 0.26 | 0.00 | None |
| Smooth | Rough | 0.56 | 0.15 | 0.38 | 0.38 | -0.03 | 0.62 | 0.38 | 0.53 | Positive |
| Short | Long | 0.24 | -0.09 | 0.34 | 0.36 | 0.12 | 0.18 | 0.08 | 0.24 | Positive |
| Strong | Weak | 0.04 | -0.08 | 0.17 | -0.23 | 0.21 | 0.14 | -0.19 | -0.23 | None |
| Pleasant | Clicky | 0.75 | 0.62 | 0.63 | 0.63 | 0.63 | 0.51 | 0.43 | 0.50 | Positive |
| Slow | Fast | -0.27 | -0.22 | -0.06 | -0.17 | -0.07 | -0.05 | -0.12 | -0.16 | None |
| Hard | Soft | -0.16 | -0.17 | -0.11 | -0.29 | -0.44 | -0.09 | -0.51 | -0.19 | Negative |
| Mechanical | Buzzy | 0.07 | 0.00 | -0.07 | -0.29 | -0.03 | 0.08 | -0.15 | -0.21 | Negative |
| | | **Correlations with good/bad scale per setting** | | | | | | | | |
| **Skip** | | $\bar{Y}_{11}$ | $\bar{Y}_{21}$ | $\bar{Y}_{31}$ | $\bar{Y}_{41}$ | $\bar{Y}_{12}$ | $\bar{Y}_{22}$ | $\bar{Y}_{32}$ | $\bar{Y}_{42}$ | |
| Faint | Predominant | 0.11 | 0.22 | 0.56 | 0.03 | -0.05 | -0.20 | 0.15 | 0.28 | Positive |
| Smooth | Rough | 0.49 | 0.35 | 0.21 | 0.40 | 0.48 | 0.30 | 0.25 | 0.38 | Positive |
| Short | Long | 0.64 | 0.54 | 0.47 | 0.03 | 0.32 | 0.07 | 0.49 | 0.42 | Positive |
| Strong | Weak | -0.11 | -0.34 | -0.41 | -0.19 | 0.04 | -0.04 | -0.22 | -0.30 | Negative |
| Pleasant | Clicky | 0.75 | 0.68 | 0.43 | 0.64 | 0.57 | 0.78 | 0.72 | 0.62 | Positive |
| Slow | Fast | -0.23 | -0.24 | -0.31 | -0.26 | -0.23 | -0.20 | -0.32 | -0.18 | Negative |
| Hard | Soft | -0.48 | -0.74 | -0.47 | -0.42 | -0.55 | -0.56 | -0.70 | -0.57 | Negative |
| Mechanical | Buzzy | -0.14 | -0.13 | -0.17 | -0.17 | -0.42 | -0.12 | -0.03 | -0.06 | Negative |

Generally, it can be observed that the correlations for the mode functionality are less distinct than the ones for the skip functionality. None of the correlation trends contradict each other, though.

To investigate the underlying factors, a principal component analysis was run. The nine items were rotated via varimax. The Bartlett test became significant with $\chi^2$ (36) = 710.23, $p < 0.001$ (mode) and $\chi^2$ (36) = 700.23, $p < 0.001$ (skip)

The sampling adequacy was ensured with a Kaiser-Meyer-Olkin criterion (KMO) = .7 (mode)/ .77 (skip). As the number of cases was well over 300, the communality values of over .5 for each item (mode and skip) are sufficient (Field, Miles, & Field, 2012; cp. Table 7. *Factor loadings of principal component analysis* for exact values).

Scree plots showed the possibilities of either extracting one factor, which would then be switch characteristics, or four, which would then lead to different underlying factors (cp. Figure 11. *Mode Scree Plot* and Figure 12. *Skip Scree Plot*).

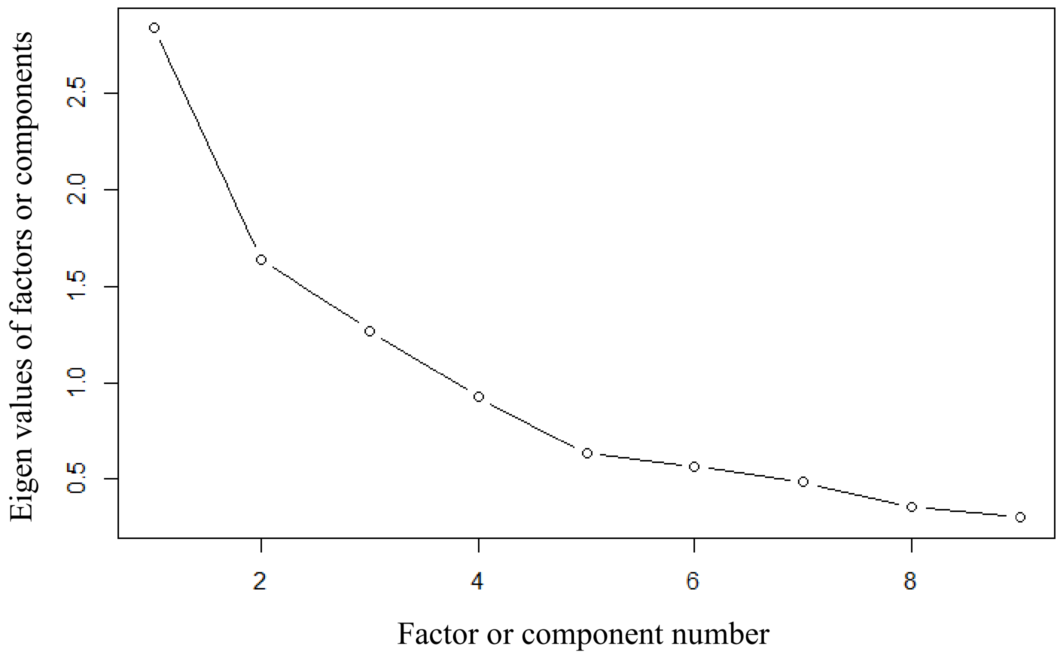



*Figure 11. Mode Scree Plot.*

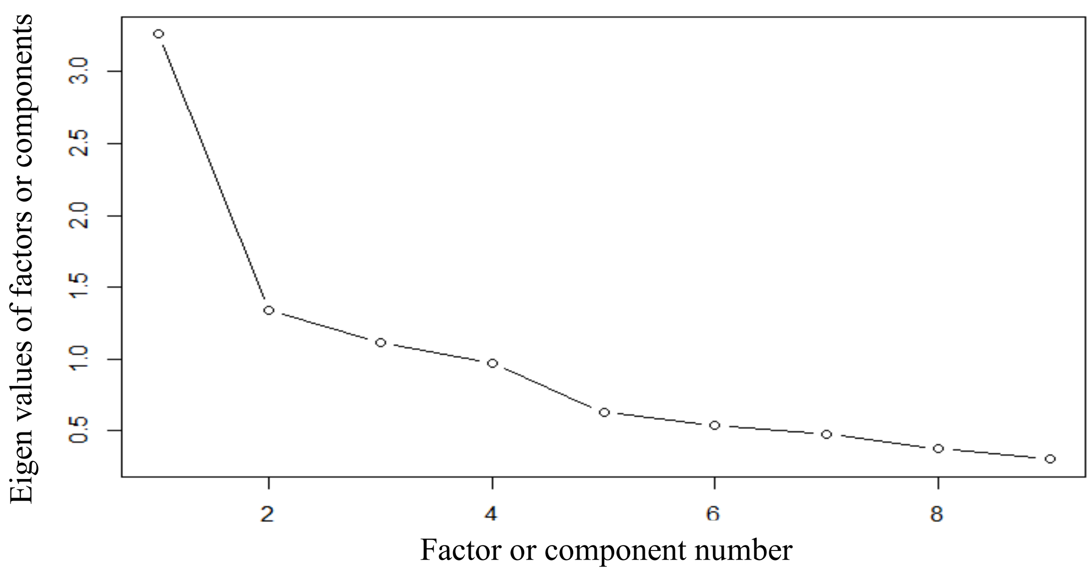



*Figure 12. Skip Scree Plot.*

The fit based upon off diagonal values was 0.91 for mode and .93 for skip. This indicates that four factors are more likely to be an adequate decision when taking the sample size into account (cp. Field et al., 2012). The first four factors had eigenvalues above .92/.97 (mode/skip) and explained 74% of the variance in both cases. Additionally given the closeness of the eigenvalues to Kaiser's Criterion (= 1), although this is not always a sufficient measure (Bell, 2004), four factors were obtained for further analysis. An analysis of residuals showed that the proportion of the total number of residuals is at 44% (mode)/ 36% (skip), which is within the acceptable limit (cp. Field et al., 2012). The root-mean-square residual is at 0.084 (mode)/ and 0.082 (skip), which might indicate that more factors could be extracted in addition.

The obtained factor loadings are summarized in the table below (Table 7. *Factor loadings of principal component analysis*).

*Table 7. Factor loadings of principal component analysis.*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | **Factor Loadings** | |
| **Mode** | | item | **RC1** | **RC2** | **RC3** | **RC4** | h2 | u2 |
| Clicky | Pleasant | 5 | 0.84 | | | | 0.76 | 0.24 |
| Bad | Good | 9 | 0.81 | | | | 0.69 | 0.31 |
| Rough | Smooth | 2 | 0.74 | | | | 0.61 | 0.39 |
| Strong | Weak | 4 | | 0.89 | | | 0.81 | 0.19 |
| Predominant | Faint | 1 | | -0.83 | | | 0.71 | 0.29 |
| Hard | Soft | 7 | 0.43 | 0.62 | | 0.33 | 0.68 | 0.32 |
| Fast | Slow | 6 | | | -0.88 | | 0.8 | 0.2 |
| Long | Short | 3 | | | 0.79 | | 0.73 | 0.27 |
| Mechanical | Buzzy | 8 | | | | 0.94 | 0.88 | 0.12 |

| **Skip** | | item | **RC1** | **RC3** | **RC2** | **RC4** | h2 | u2 |
|---|---|---|---|---|---|---|---|---|
| Clicky | Pleasant | 5 | 0.85 | | | | 0.75 | 0.249 |
| Bad | Good | 9 | 0.8 | 0.31 | | | 0.75 | 0.254 |
| Rough | Smooth | 2 | 0.73 | | | | 0.6 | 0.402 |
| Hard | Soft | 7 | 0.66 | | 0.33 | | 0.71 | 0.295 |
| Fast | Slow | 6 | | -0.87 | | | 0.78 | 0.222 |
| Long | Short | 3 | | 0.78 | | | 0.7 | 0.303 |
| Predominant | Faint | 1 | | | -0.87 | | 0.76 | 0.236 |
| Strong | Weak | 4 | | | 0.82 | | 0.73 | 0.27 |
| Mechanical | Buzzy | 8 | | | | 0.95 | 0.92 | 0.085 |

The variance explained by this model is at 74% (cp. appendix a) f *Principal component analysis* . View the same source for the full analysis for both mode and skip function).

#### 3.6.2.2.5 Summary and conclusion

Throughout the data, there were no contradictions between the mode and the skip functionality, which shows that for skip and mode function, performance and preference can be treated similarly.

For future studies, an adapted questionnaire with item polarity will be implemented. In total four items need to be re-poled according to their correlation results. In terms of settings, as $\bar{Y}_{41,}$ $\bar{Y}_{42,}$ and $\bar{Y}_{32}$ did not pass the error rate criterion, they should be discarded as they might be safety-critical. Settings $\bar{Y}_{21}$, $\bar{Y}_{31}$, $\bar{Y}_{41}$, and $\bar{Y}_{42}$ differed significantly from the ideal configuration and should therefore be discarded as well. On the good – bad – scale, settings $\bar{Y}_{11,}$ $\bar{Y}_{12,}$ and $\bar{Y}_{22}$ were significantly better than several other settings. As setting $\bar{Y}_{12}$ has the most significant differences, it is used furtherly in the subsequent studies. It is noteworthy that this setting is allocated on the lower end of the automotive force scale for mechanical switches; its higher preference performance might be due to the fact that it incorporates some aspects of capacitive switches, which are already familiar to subjects. This is a hint that the *mere exposure effect* (Zajonc, 1968) may have an influence, which is the effect that subjects tend to prefer things they are used to over novel stimuli.

For the principal component analysis, four factors were obtained. As the first factors include items such as *good/bad*, *clicky/pleasant* and *rough/smooth*, it is determined as *valence*. The second factor includes the items *strong/weak*, *predominant/faint*, and *hard/soft*, which is allocated with the meaning *potency* characteristic of a switch configuration. The third category includes *dynamics* items such as *slow/fast* and *long/short*, which is why they were labeled accordingly. This corresponds to the three dimensions of a semantic differential after Osgood et al. (1957). The last dimension describes how *mechanical* or *buzzy*, in other words how comparable the configuration feels to a mechanical or a virtual switch, and is therefore labeled *virtuality*. The presented dummy comparison did not lead to new items in the qualitative analysis, as not enough subjects (minimum 25%) did come up with a new construct for this comparison.

### 3.6.3 Third preliminary experiment: Time interlock and force delta

The purpose of the third experiment was to validate the factor structure of the previous experiment. Additionally, two components of switch signal design were investigated in terms of usability (preference, effectiveness and effectivity): the time between pulses (tbp) for the skip function, in other words: how much time should pass between each push of the push and hold, and the snap ratio for the mode function, or, to to put it like this, how the relationship between activation and deactivation force should be set.

#### *3.6.3.1 Method*

In total, 32 healthy subjects participated voluntarily in this study. They were equally distributed over the age groups recommended by NHTSA, but not equally distributed over gender, as the subjects had to be recruited within the company this time for economic reasons. The hypotheses are

> $H_1(1)$: There is a difference between the time between pulses preferences for different settings.
>
> $H_0(1)$: There is no difference between the time between pulses preferences for different settings.
>
> $H_1(2)$: There is a difference in performance between the time between pulses preferences for different settings.
>
> $H_0(2)$: There is no difference in performance between the time between pulses preferences for different settings.
>
> $H_1(3)$: There is a difference between the snap ratio preferences for different settings.
>
> $H_0(3)$: There is no difference between the snap ratio preferences for different settings.
>
> $H_1(4)$: There is a difference in performance between the snap ratio preferences for different settings.
>
> $H_0(4)$: There is no difference in performance between the snap ratio preferences for different settings.

The feedback used was a pure 230 Hz sine wave, with low feedback amplitude, as this was the most successful setting in terms of usability in the previous experiment. As this experiment took place on an advanced prototype, the amplitudes had to be measured again, as they were not expected to be congruent. Amplitudes were recorded using a USB driven high power frontend by Head Acoustics, Inc., Octobox+. The sensor was a Model M352C65 sensor by Platinum Stock Products. For 5 – 255 bit in amplitude, the vibrotactile acceleration was measured in steps of five. The peak-to-peak amplitude for each initial sine for each measurement point was used to determine the displacement as discussed in the chapter *Psychophysiological measurement of tactile* *perception.* A feedback with a lower value in dB in reference to 1µm could be selected because of its congruence to the amplitude of the low feedback stimuli in the previous study.

The factorial design of the experiment is described in the following table (Table 8. *Factorial design of the third preliminary experiment*).

*Table 8. Factorial design of the third preliminary experiment.*

| Snap Ratio | **Mode** |
|---|---|
| Low | $\bar{Y}_{11}$ |
| Medium | $\bar{Y}_{21}$ |
| High | $\bar{Y}_{31}$ |

| tbp | **Skip** |
|---|---|
| Slow | $\bar{Y}_{12}$ |
| Medium | $\bar{Y}_{22}$ |
| Fast | $\bar{Y}_{32}$ |

As a test setting, the driving simulator described in the chapter *Driving Simulator* was used, but only with front projection in order to avoid noise issues which would have occurred with several projectors instead of one.

Subjects first filled out a consent form, then room and finger temperature of the right hand were measured. They drove a familiarization drive of seven minutes on a rural road with gentle curves. This was done in order to get a hold on subjects suffering from simulator sickness previously to the start of the real experiment. The last two minutes served as a baseline for the driving data in the data analysis in line with the previous study. Subjects then filled out a NASA-TLX (Hart & Staveland, 1988) and then a demographic questionnaire (appendix a) a *Second preliminary experiment Demographic questionnaire*), which included the valence ("pleasure") and the arousal scale of the SAM (Bradley & Lang, 1994), but not the dominance scale, as the validation of this scale is still debated. For Repertory Grid generation, dyads were used in this testing. This Repertory Grid was generated both for the mode, as well as for the skip-switch function and can be found in appendix b) a *Repertory Grid Design*.

Then, subjects rated the different stimuli on a five point scale, on the newly generated item pairs and on the ones that were generated in the previous study. The response anchors corresponded to those developed by Vagias (2006). The subjects were driving when they performed the evaluation and the experiment conductor took down the rating for them. At the end of the rating of the different switch configurations, subjects rated on the questionnaire what their ideal configuration would look like. After this they conducted a performance task: For the skip function, forward skip was used, the subjects were instructed to press and hold the switch for five activations forwards. The different activations were fed back to the subjects via the tactile feedback. For the mode function, the subjects were instructed to activate and deactivate the switch four times as fast as possible. After the performance task, the subjects filled out the NASA-TLX (Hart & Staveland, 1988) to rate the difficulty of the driving task in the performance condition.

The same procedure was conducted first either for mode or skip, and then for the other function modality. The order of the functions was balanced for the subjects and the order of the stimuli randomized and balanced across the different subjects. The subjects were instructed throughout the course to drive with 70 km/h per hour and to prioritize to the driving task when a secondary task was to be conducted.

#### *3.6.3.2 Results of the third preliminary test*

##### 3.6.3.2.1 Switch performance

In order to investigate whether there was a difference between the switch settings in terms of performance, efficiency and effectiveness of the switches were measured. The operationalization was effected through time on task (efficiency) and error rate (effectiveness).

In terms of the snap ratio, 8 out of 96 data points were missing and were not replaced. The previously used criterion for the binomial test with 10% error rate and 95% confidence interval did not lead to any discrimination between the different variations. As a consequence the error rate was calculated with 15% error rate and a 95% confidence interval and as a backup with 10% error rate and a confidence interval of 99%. Both adjustments of the error rate led to significantly more errors for variations $\bar{Y}_{21}$ ($p = 0.0152$) and $\bar{Y}_{31}$ ($p = 0.01139$).

There were 9 out of 96 data points missing for the time between pulses (tbp) condition. As the previous criterion for the binomial test with 10% error rate and 95% confidence interval did not lead to any discrimination between the different variations again, both the adjusted error rate calculation with a binomial test with 15% error rate and confidence interval of 95% as well as a binomial test with 10% error rate and confidence interval of 99.9% led to the same result. Factor iterations $\bar{Y}_{12}$ and $\bar{Y}_{32}$ had significant (both $p < 0.001$) more errors than allowed by the criteria.

The time on task was calculated for both switch configurations separately. As there were missing data points present in the switch performance data set, missing values were replaced by employing the expectation maximization (EM) method (Dempster et al., 1977) using the *Amelia* package in R. For the snap ratio condition a total of 11 out of 96 data points were replaced. As the assumption of normal distribution (Shapiro Wilk test, $W = 0.92$, $p < 0.001$) and homogeneity of variances (Levene's test with $F(2/93) = 2.90$, $p = 0.05995$) were both violated, a Friedman's ANOVA was calculated with $\chi^2(2) = 4.84$, $p = 0.08898$, $\omega^2 = 0.02$.

For calculating the time on task for the skip condition, the duration of each variation had to be taken out through pre-processing. Additionally, the number of faulty activations had to be taken into account and data was normalized over this. A total of 15 out of 96 data points were replaced using the EM method. The data was not distributed normally, as a Shapiro-Wilk became significant with $W = 0.27$, $p < 0.001$. The homogeneity of variance was given via Levene's test with $F(2/84) = 0.58$, $p = 0.562$. As a result, a Friedman's ANOVA was calculated and did not become significant with $\chi^2(2) = 1.28$, $p = 0.5271$, $\omega^2 = 0.001$.

##### 3.6.3.2.2 Driving performance

In order to investigate whether there was a difference between the switch settings in terms of driving performance, efficiency and effectiveness were measured. Whereas effectiveness was operationalized via the driving parameters mean velocity, standard deviation of mean velocity, lane keeping quality, lane position, steering wheel angle velocity and steering wheel angle acceleration, efficiency was operationalized via the NASA-TLX between the baseline drive and the performance task.

For the NASA-TLX, the preconditions for calculating an ANOVA were violated with a Shapiro-Wilk-Test with $W = 0.92$, $p < 0.001$. A Levene's test with $F(2/62) = 1.61$, $p = 0.2056$ did not become significant. A Friedman's ANOVA became significant with $\chi^2(2) = 24.02$, $p < 0.001$, $\omega^2 = 0.087$. Post-hoc Nemenyi tests revealed significant differences between Baseline and each condition, but not between the conditions themselves (see Figure 13. *NASA-TLX results of the third preliminary study*).

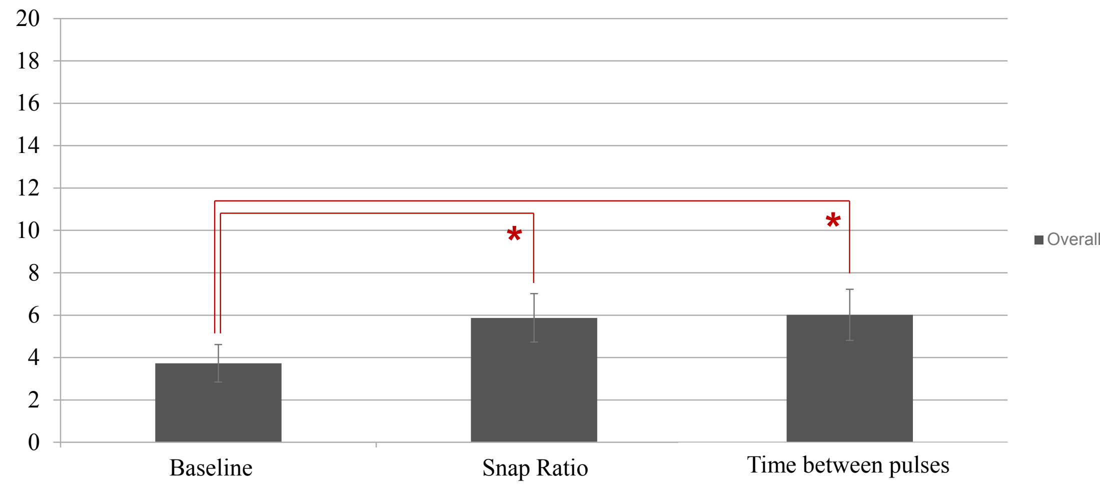



*Figure 13. NASA-TLX results of the third preliminary study.*

Significant results for the effectiveness of the driving data are reported below.

In the snap ratio (mode) rating condition, the data was not distributed normally (Shapiro-Wilk with W = 0.94, $p < 0.001$), so a Friedman's ANOVA was calculated for the lane keeping quality and became significant with $\chi^2$ (2) = 6.06, $p = 0.04826$, $\omega^2 = 0.002$. A Friedman-Nemenyi-Posthoc-Test revealed a significant difference (p = 0.046) between factor variations $\bar{Y}_{11}$ and $\bar{Y}_{21}$, with the latter having a more desired value (a lower deviation from the lane keeping quality, which would at an optimum equal zero).

In the time between pulses rating condition, the data was not distributed normally (Shapiro-Wilk with W = 0.43, $p < 0.001$), so a Friedman's ANOVA was calculated for the steering wheel angle velocity and became significant with $\chi^2$ (2) = 15.30, $p < 0.001$, $\omega^2 = 0.16$. A Friedman-Nemenyi-Posthoc-Test revealed a significant difference (p = 0.0093) between factor variations $\bar{Y}_{12}$ and $\bar{Y}_{22}$, with the latter having a more desired value (a lower steering wheel angle velocity).

Another Friedman's ANOVA was calculated for the steering wheel angle acceleration in the time between pulses rating condition, as the data was not distributed normally, either (Shapiro-Wilk with W = 0.40, $p < 0.001$) and became significant with $\chi^2$ (2) = 12.5, $p = 0.00193$, $\omega^2 = 0.03$. A Friedman-Nemenyi-Posthoc-Test revealed a significant difference (p = 0.033) between factor variations $\bar{Y}_{12}$ and $\bar{Y}_{32}$, with the latter having a more desired value (a lower steering wheel angle acceleration). All driving data can be found in appendix b) b *Driving performance results*.

In the time between pulses performance condition, the data was not distributed normally (Shapiro-Wilk with W = 0.86, $p < 0.001$), so a Friedman's ANOVA was calculated for the lane keeping quality and became significant with $\chi^2 (2) = 6.06$, $p = 0.04826$, $\omega^2 = 0.003$. A Friedman-Nemenyi-Posthoc-Test revealed a significant difference ($p = 0.046$) between factor variations $\bar{Y}_{12}$ and $\bar{Y}_{22}$, with the latter having a more desired value (higher lane keeping quality with a lower deviation from the lane keeping quality, which would at an optimum equal zero).

Note that except for the last condition, the differences did not occur in the performance condition and the validity of these results is therefore diminished, especially when taking the small effect size ($\omega^2$) into account.

#### 3.6.3.2.3 Subjective preference data

In order to analyze the preferred switch configurations, the results of the good – bad – scale at the end of each of the switch settings and the distance to the ideal configuration were measured.

The following paragraph reports the results of the good – bad – scale. For the snap ratio calculation, the assumption of normality was violated (Shapiro-Wilk Test with W = 0.85, $p < 0.001$), but the assumption of homogeneity of variances was met (Levene's Test with $F (2/93) = 0.78$, $p= 0.461$). An ANOVA was calculated and did not become significant with $F (2/93) = 0.20$, $p = 0.823$. In terms of the time between pulses, the assumption of normality was violated as well (Shapiro-Wilk Test with W = 0.88, $p < 0.001$) and the Levene's Test with $F (2/93) = 2.78$, $p = 0.06712$ did not become significant and the assumption of homogeneity of variance was therefore met. An ANOVA became nearly significant with $F (2/93) = 2.97$, $p = 0.0564$. A post-hoc calculation with Bonferroni corrected t-tests showed the tendency of a significant difference between low and medium with $p = 0.052$.

In terms of distance to the ideal configuration, the assumption of normality was violated (Shapiro – Wilk with W = 0.88, $p < 0.001$), but the homogeneity of variances was given (Levene's Test with $F (3/1276) = 0.52$, $p= 0.6706$). Therefore an ANOVA was calculated and did become significant with $F (3/1276) = 3.15$, $p = 0.0242$, $\omega^2 = 0.005$. A post-hoc Tukey HSD revealed significant differences to the ideal configuration for $\bar{Y}_{21}$ ($p = 0.0467267$) and $\bar{Y}_{31}$ ($p = 0.0429102$). The results are displayed in the following (Figure 14. *Snap ratio distance to ideal configuration*).

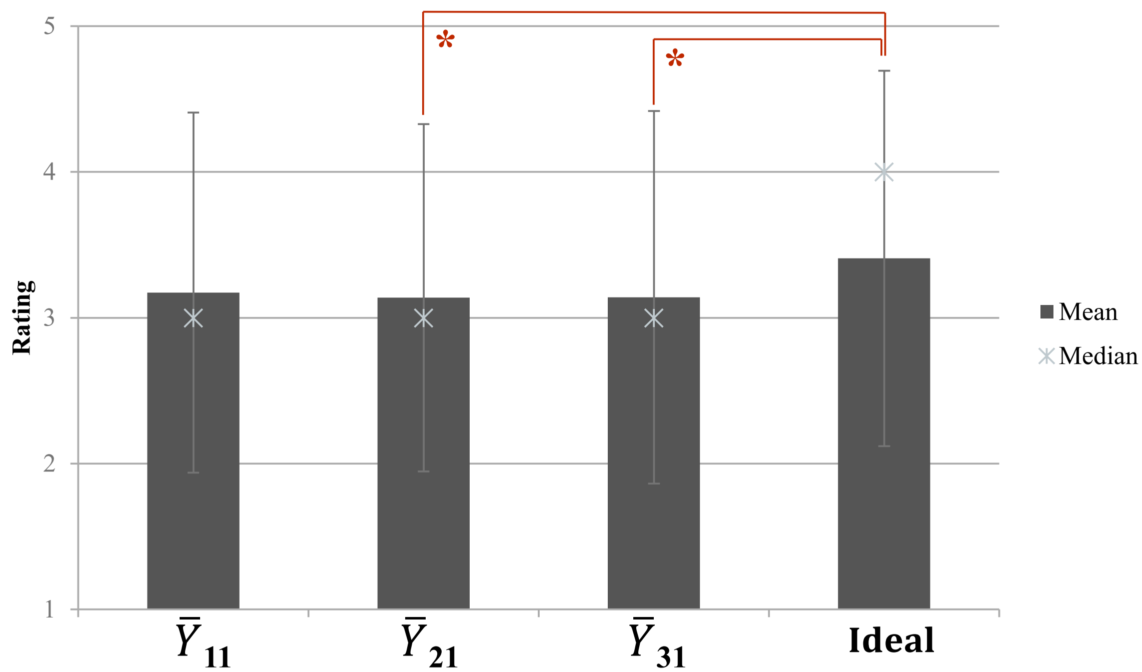



*Figure 14. Snap ratio distance to ideal configuration.*

In the time between pulses rating, the assumption of normality was violated (Shapiro – Wilk Test with $W = 0.90$, $p < 0.001$), but the homogeneity of variances was given (Levene's Test with $F(3/1275) = 0.91$, $p = 0.4368$). A subsequent ANOVA did become significant with $F(3/1275) = 5.13$, $p = 0.00157$, $\omega^2 = 0.01$. A post-hoc (Tukey HSD) test showed significant differences between the ideal configuration and $\bar{Y}_{32,}$ as well as between $\bar{Y}_{32}$ and $\bar{Y}_{12}$, the latter having the closest mean to the ideal configuration. They are displayed in the following (Figure 15. *Time between pulses distance to ideal configuration*).

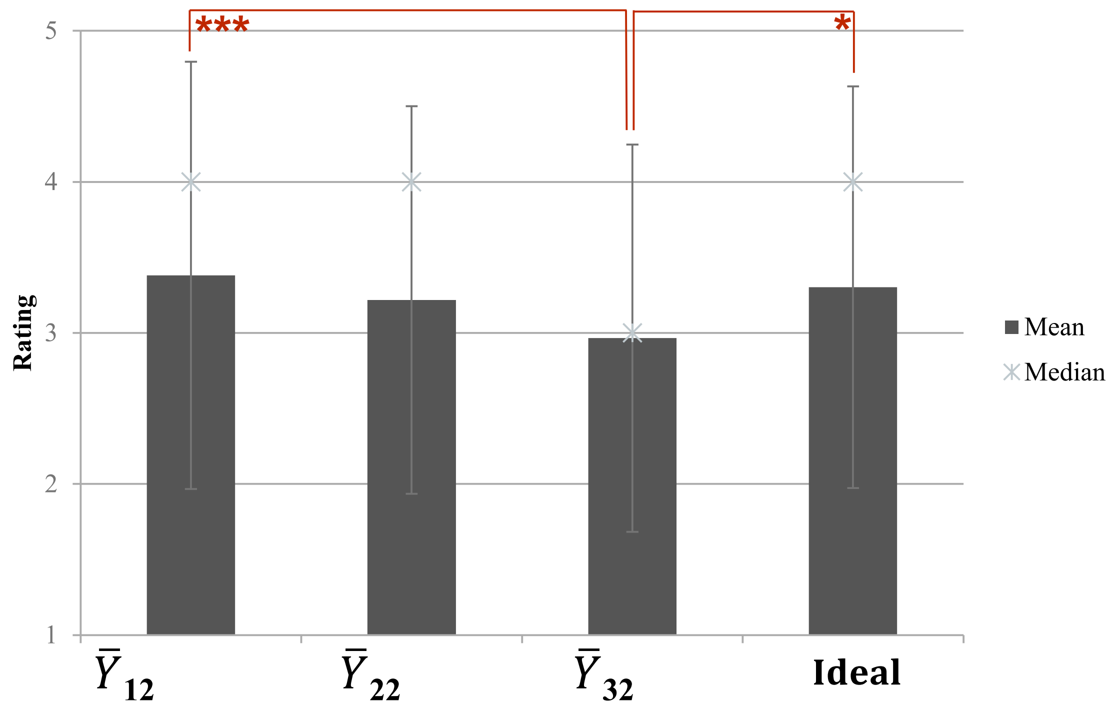



*Figure 15. Time between pulses distance to ideal configuration.*

The full preference data for both configurations can be found in appendix b) c *Preference data*.

3.6.3.2.4 Questionnaire evaluation

No new items emerged for the snap ratio as well as the time between pulses testing.

3.6.3.2.5 Principal component analysis

For the principal component analysis, both data sets from mode and skip functionality were combined, as they tended to point in similar directions in the last study. The ten items were rotated via varimax, as calculations of oblique rotation showed little intercorrelation between the different factors (cp. a) d *Principal component analysis*). The Bartlett test became significant with $\chi^2$ (45) = 472.75, $p < 0.001$.The sampling adequacy was ensured with a Kaiser-Meyer-Olkin criterion (KMO) = .64, which is slightly lower than in the previous study. The scree plot showed ambiguous results in the present data set (cp. Figure 16. *Scree plot of third preliminary study*).

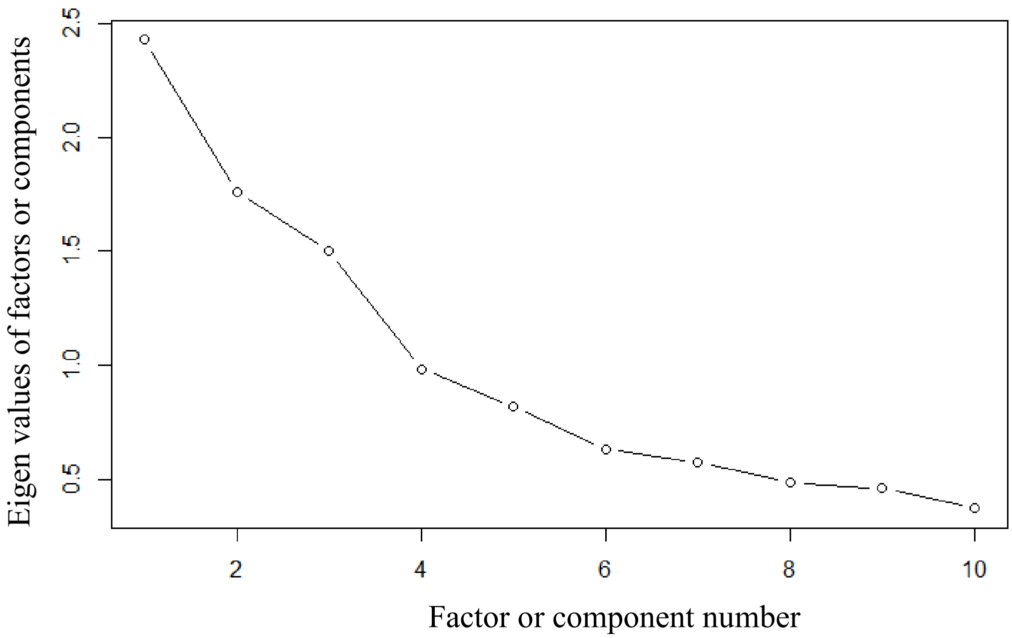



*Figure 16. Scree plot of third preliminary study.*

It could be argued for the extraction of either three or five factors. The Kaiser's criterion (cp. appendix a) d *Principal component analysis*) is in favor of three factors, but some of the communalities are below the acceptable value of .6 (Klopp, 2015), where the small sample size has to be taken into account. The fit of the model is also rather low with .8, but this might be due to the small sample size as well. Whereas increasing the factor number to four does still lead to communality values below .6 and the fit of the model increases slightly to .81, increasing the factor number to five meets the communality criterion, but the overall fit of the model remains with .86 under the criterion of .9. As the ratio of factors to items should at least be 1:3 ( Wagner, Drescher, & Reichel, 2010), the increase of factor numbers should be occured with caution only. As this points rather in the direction of taking a smaller fit into account, five factors were obtained for final analysis. The factors are summarized in the table below (Table 9. *Factor analysis results of the third preliminary study*).

*Table 9. Factor analysis results of the third preliminary study.*

| item | | | **RC1** | **RC3** | **RC2** | **RC5** | **RC4** | h2 | u2 |
|---|---|---|---|---|---|---|---|---|---|
| Weak | Strong | 5 | .81 | | | | | .66 | .34 |
| Faint | Predominant | 2 | .76 | | | | | .71 | .29 |
| Soft | Hard | 8 | .67 | | | .33 | | .63 | .37 |
| Fast | Slow | 7 | | .82 | | | | .68 | .32 |
| Short | Long | 4 | | .75 | | .34 | | .74 | .26 |
| Responsive | Sluggish | 1 | .35 | .68 | | | | .69 | .31 |
| Good | Bad | 10 | | | .89 | | | .82 | .18 |
| Pleasant | Clicky | 6 | | | .81 | | | .76 | .24 |
| Smooth | Rough | 3 | | | | .88 | | .84 | .16 |
| Mechanical | Buzzy | 9 | | | | | .97 | .94 | .06 |

It should be noted that the order of the items differs from the previous study, as it was adapted to the valence ratings. The explained variance of this model with five factors is at 75%, which is at a medium level (cp. (Field et al., 2012)).

Another possibility to analyze multidimensionality is to investigate the Cronbach's α (Cronbach, 1951), which is usually a measure for reliability in terms of internal consistency. The author recommends to measure reliability for each scale separately. If there would be only one factor, the reliability of the overall scale should be higher than for each subscale. This is not the case for the present data set (s. appendix b) e *Cronbach's α*, where the overall α is 0.63, compared to the three first subscales with α = .64, .66, and 0.67. It is not possible to calculate α for factors with a single item.

The first factor contains the same items as in the previous study; the second factor lost the dimension s*mooth – rough*, which is now loading on a single factor (RC5). The third factor gained the new dimension r*esponsive – sluggish* and remained otherwise the same. The remaining factor RC4 remained also the same. Full results of the PCA can be found in appendix b) d *Principal component analysis*.

### *3.6.3.3 Discussion*

The error rate results of the snap ratio condition point in the direction of using $\bar{Y}_{11}$ in further studies, as the two other factor variations did not pass the criterion. It should be noted that the differences between the snap ratios was perceivable, but very small, which might account for the adjustment need of the criterion. There were some missing values, but to an acceptable extend. Missing values were not replaced, but omitted from the overall calculation. In the time between pulses condition, the error rate points in the direction of employing $\bar{Y}_{22}$, as this was the only factor variation passing the criterion as well.

It should be noted that the driving data is ambiguous and cannot be interpreted clearly, especially as most of the results were derived not in the performance, but in the rating condition, with a rather small effect size. Nevertheless the result points in the time between pulses performance condition in the direction of disregarding setting $\bar{Y}_{12}$, as it performs significantly worse than other settings. This is supported by the time between pulses rating condition, where this factor variation performs always worse than one of the other conditions. The snap ratio rating points in the direction of a significant better performance of $\bar{Y}_{21}$ over $\bar{Y}_{11}$, but not in a performance task, so it is questionable whether this is a reliable result.

The preference data seems to support the error rate results. For the snap ratio, $\bar{Y}_{11}$ had no significant differences to the ideal configuration, whereas the other two factor variations did have them. As subjective and objective data are in line, it is recommended to use a low deactivation force in relation to the activation force for virtual switches. This result might be due to the fact that the perceived deepness of the switch is higher when the release point is at a very low force. The interaction becomes more precise and richer when the switch is perceived as not being flat and shallow, as well as the accuracy is eased by this.

There were little results from the *good – bad* – scale in this testing, which might be due to the fact the differences between the conditions were very small and the sample was also reduced due to economic reasons. Whereas there were no results for the snap ratio rating, the time between pulses showed a non-significant tendency in the direction of letting go of $\bar{Y}_{12}$.

The time between pulses distance to the ideal configuration comparison points in the direction of disregarding a long time between pulses, as subjects significantly did not prefer setting $\bar{Y}_{32}$. This is backed up by the error rate results and complemented further in disregarding $\bar{Y}_{12}$, as this factor variation did not pass the error criterion. It seems that the lowest time between pulses is too fast for accurate interaction, whereas the highest is too slow to be perceived as qualitatively valuable. In conclusion, factor variation $\bar{Y}_{22}$ is recommended for virtual switch configuration with a push-and-hold function.

For the principal component analysis it can be stated that at large, the underlying factors of the questionnaire remained the same, although the quality criteria decreased a little bit, which might be due to the decreased number of cases compared to the previous study. There were no new items added to the instrument in this iteration of the testing. Either the items named by the subjects were already a part of the questionnaire or the quality criterion of 25% of subjects naming an item was not passed. This might imply that the number of items is already sufficient for describing the present topic, or that the dyad method is not sufficient for this kind of technology. To exclude the latter possibility, the final study will again employ a triad.

It should be noted that the number of subjects were relatively small and the addition of more subjects might have led to new items as well as improved quality criteria for the principal component analysis as well as for Cronbach's α.

## *3.7 Methodology of the main experiment*

In the main experiment, the goal was to evaluate three possible technologies for steering wheel switches: a contemporary mechanical switch, an automotive-applicable capacitive switch and the aforementioned force-sensitive virtual switch technology.

### 3.7.1 Participants

40 healthy subjects participated voluntarily in this study. They were equally distributed over the age groups recommended by NHTSA (Garrett, 2012) and also equally balanced over gender. 20% of the subjects were recruited internally, with no connection to the project. The remaining 80% were recruited from the TAKATA AG subject data base. Exclusion critera were auditory or tactile perception disabilities and thickly rimmed glasses, as they were problematic for the eyetracking system: the shadow of the rim tended to obscure the pupil, thus enhancing the difficulty of pupil detection. 39 subjects were right-handed and one subject both-handed.

### 3.7.2 Hypotheses

As the main aspects of usability, efficiency, effectiveness and acceptance, are to be investigated, the first two hypotheses deal with user preference and performance:

$H_1(1)$: There is a difference between the switches in terms of driving performance.

$H_0(1)$: There is no difference between the switches in terms of driving performance.

$H_1(2)$: There is a difference in performance between the switches.

$H_0(2)$: There is no difference in performance between the switches.

The distraction potential between the different switch types are evaluated additionally. For this purpose an eyetracking system will be employed and the third hypothesis is accordingly:

$H_1(3)$: There is a difference in distraction between the switches.

$H_0(3)$: There is no difference in distraction between the switches.

As in previous studies, the preference of the different technologies is evaluated. The fourth hypothesis is formulated consequently:

$H_1(4)$: There is a difference in preference between the switches.

$H_0(4)$: There is no difference in preference between the switches.

And, the final evaluation for the acceptance measurement of the developed multimodal virtual switch scale (MVSS) has to take place. Therefore, the last hypothesis is:

$H_1(5)$: The MVSS is a valid measurement of acceptance.

$H_0(5)$: The MVSS is not a valid measurement of acceptance.

As the goal of the main study was to compare three technologies for steering wheel switches, the factorial design was rather simple. It is displayed in the following table (Table 10. *Factorial Design for the main experiment*).

### 3.7.3 Stimuli preparation

*Table 10. Factorial Design for the main experiment.*

| Switch Technology | **Factor** |
|---|---|
| Mechanical | $\bar{Y}_{11}$ |
| Capacitive | $\bar{Y}_{21}$ |
| 3D Switch | $\bar{Y}_{31}$ |

For the creation of a multimodal signal for the virtual switch, the auditory sample retrieved from the study *Auditory settings* was edited using *Audacity*[7], a freeware program which allows for sound editing. The auditory signal had to be sampled down, formatted accordingly and adapted in terms of signal length for the present evaluation. This signal was employed both for the virtual force-sensitive and the capacitive switch.

From the past study, a 230 Hz signal with lower amplitude is combined with the audio signal. The activation force is at 2N, with a snap ratio of 70%, as this was evaluated as an optimal setting for virtual force – sensitive switches. For the interaction, a single switch acitivation is investigated, as neither capacitive nor mechanical are fully capable of a true push-and-hold function.

### 3.7.4 Steering wheels

Three steering wheels are compared in this study: a series production mechanical switch, a capacitive switch, and a force-sensitive virtual switch. All steering wheels are modified 2014 Jeep Grand Cherokee SRT steering wheels. The mechanical switches are used as present, but communication to the simulator was established. For the capacitive switch, an additional sound was played through the driving simulation software, which was the same sound as with the force – sensitive virtual (3D) switch. As it is usual with smartphones, the vibrations are transferred via the hand which holds the smartphone and not with the finger interacting with the display, there was no vibrotactile feedback implemented for the capacitive switch. This is implemented due to increasing external validity, as the interaction with a capacitive screen in a car usually gives no or too delayed vibrotactile feedback for human sensor capabilities. All three steering wheels have to be exchanged during testing.

### 3.7.5 Eyetracking equipment

The eyetracking employed in this study is a binocular Dikablis professional system by ERGONEERS. This system tracks with a frequency of 60 Hz and has an accuracy of 0.05° visual

[7] http://audacityteam.org/

angle for pupil tracking and a 0.1° - 0.3° visual angle for glance direction accuracy. The camera resolution for the eye camera is 648 x 488 pixels, whereas the scene camera is full HD (1920 x 1080 pixels) with an aperture angle of 90°. For head movement is compensated via QR markers in the scene, preferable with a physical proximity to the areas of interest (AOI). The angle of the markers in the scene camera allow for a compensation of these movements, where otherwise a head tracker would be required.

#### 3.7.6 Final Grid design

The final grid for the main experiment consisted of 13 items which were generated in respect to vibrotactile characteristics of a virtual switch. In a previous study (*Auditory settings*) conducted with the virtual switch technology, only four items emerged for characterizing a preferred switch sound: *Sharp, high frequency, dynamic* and *fast* (Edwards & Cunefare, 2015). As these items were not generated on a semantic differential, there were no opposite poles for the items present. Ideally, these items should have been validated with another Repertory Grid study; unfortunately, time and resource constraints prohibited this approach. So for integrating the items into the present method, the opposite poles have to be constructed. The easiest way to do this would be using the word "not" before each named item on the questionnaire (as a psychoacoustic application for the word *sharp* with the opposite *not sharp* see Bech & Zacharov (2007); Schulte-Fortkamp, Quehl, Mellert, & Remmers (2001)). This approach has the advantage that for words for which opposites are hard to construct, a negative pole can be constructed, which is accurate yet not manipulating. On the other hand the item is not really bipolar in its original sense, but unidimensional with varying intensity. This would lead to inconsistency within the scale, something that should be avoided in user testing.

As the item *fast* with the opposite of *slow* is already part of the questionnaire, it would not make sense to list this item twice, so it can be disregarded as an addition. For the item *high frequency*, the opposite pole *low frequency* was listed on the psychoacoustic questionnaire as a separate items which seemed to have been functioned as a true opposite; in other words: when subjects rated a sound *very much* on the *high frequency* scale, they rated it *not at all* on the *low frequency* scale and vice versa (Edwards & Cunefare, 2015). This opposite pair can also be found at Schulte-Fortkamp et al. (2001).

*Sharp*, on the other hand, had no statistically determinable opposite on this particular questionnaire; in psychoacoustic semantic differentials, a frequently used opposite for this item is *dull* (see e.g. Hashimoto & Hatano, 2001; Kuwano & Namba, 2001; Zeitler & Hellbrück, 2001; Zeitler, Ellermeier, & Fastl, 2004). From a strictly musical notation perspective though, where a *sharp* accidental does the opposite of a *flat* accidental: the first one lifts the sound and the other lowers it for a half-tone, the opposite chosen in this approach should be *flat*. That criterion derives from a different development than psychoacoustics and although there are many attempts of explaining musical preference with psychoacoustic measures (e.g. Zeller, 2009), the direct translation between musical notations and psychoacoustic measurements remains unsolved. Taking this into account for the final grid generation, the opposite pair *sharp – dull* is chosen.

For the *dynamic* item, the construction of an opposite tone is not trivial. Although some references do include a dynamic dimension (e.g. an activity level like in Larsson, Opperud, Fredriksson, & Västfjäll (2009)), the exact wording and consequently a matching opposite is not present. *Calm* is often present in psychoacoustic semantic differentials (Fastl, 2006b; Hashimoto & Hatano, 2001; Kuwano & Namba, 2001; Larsson et al., 2009; Schulte-Fortkamp et al., 2001; Zeitler et al., 2004), but each opposite implies very different meanings from *dynamic*: *shrill* (Hashimoto & Hatano, 2001; Kuwano & Namba, 2001), *agitating* (Fastl, 2006b) or *shaking* (Schulte-Fortkamp et al., 2001), each of the items representing the context-specific pairs of each evaluation, but not directly transferable to virtual switch sounds. There is one pair in Fastl (2006b) that rather seems to capture the *dynamic* dimension: *busy – tranquil*. This dimension has the advantage of inheriting dynamic parameters without the strong valence attributions occurring for the ones mentioned above. A brief cross-check with native English speakers allowed for using *tranquil* as the opposite pole of *dynamic*.

The order of the questionnaire had also to be adapted. Whereas in previous studies the order of the elements was optimized for data evaluation purposes, it can lead to reply tendencies when the items have a common positive or negative pole side (Field et al., 2012). This was adapted for the final study by varying the pole for half of the items. For the analysis, these items were recoded and the recoding noted on each variable (s. appendix c) b *Grid design for main experiment*).

A small pre-study with eight subjects evaluated possible additional grid pairs for the three different prototypes via triads. All subjects were employees of TAKATA AG, but without any connection to the project. Three females and five males were tested. As already discussed in the chapter *Method selection and adaptation*, this number should be sufficient for generating grid pairs. The arrangement of the triad was varied and balanced for the subjects. All possible combinations were tested. Surprisingly, only one new item emerged for the final grid, and only in the condition where mechanical and 3D switches were paired opposed to the capacitive switch: *feedback* vs. *no feedback*. The complete results of the pretest can be viewed in appendix c) a *Results of the pretest*. As an original planned ninth subject would not have led to a new category when applying the 25% criterion, only eighth subjects were questioned.

The final grid can be viewed for reference in appendix c) 0 *Grid design for main experiment.*

### 3.7.7 Procedure

Subjects first filled out a consent form, then room and finger temperature of the right hand were measured and the eyetracking system was adapted to the subjects. They first drove a familiarization drive which contained straight highway driving, two times overtaking on the highway, straight driving on a rural road and gentle curves on a rural road. Subjects were instructed throughout the course to drive with 130 km/h on the highway and 70 km/h on the rural road, respectively and to prioritize the driving task higher when a secondary task was to be conducted. This course was also employed for all other tasks. Then they filled out the NASA-TLX (Hart & Staveland, 1988) and then a demographic questionnaire (appendix a) a Second preliminary experiment Demographic questionnaire), which included the valence (“pleasure”) and the arousal scale of the SAM (Bradley & Lang, 1994). After this familiarization phase, the first part of the main experiment began. Subjects drove the course again, this time with a performance task of pressing a certain button four times as fast as possible for each maneuver. The maneuvers were driving on a two-lane highway, overtaking on the two-lane highway, driving straight on a one-lane rural road and driving gentle curves on a rural road, thus slightly increasing the primary task difficulty at each measurement point. After each maneuver the subjects were asked to fill out the final generated Repertory Grid and a corresponding questionnaire, the aforementioned instrument by Wellings et al. (2009), which is included in appendix c) c *Switch Perception Questionnaire* (adapted after Wellings et al., 2009). After finishing the whole course, subjects filled out the NASA-TLX once more and were asked to wait outside of the simulator while the steering wheels were changed. This was repeated for the other two conditions. The order of the steering wheels was randomized and balanced across subjects. After completing all three courses, the ideal switch configuration of each subject was evaluated via another Repertory Grid questionnaire. After finishing the experiment finger and room temperature were measured for a second time. Subjects were paid and bid farewell.

# 4. Results

No significant results emerged in terms of finger or room temperature, or the SAM. The measured co-variables therefore seemed not to have influenced the results.

## *4.1 Switch performance*

In line with the previous studies the switch performance was measured in terms of efficiency and effectiveness, operationalized through time on task (efficiency) and error rate (effectiveness).

The error rate was calculated with 10% allowed error and a 95% confidence interval, both for the switch location (pressing at the correct point) and the correct number of activations. For the location, 22 data points were missing and were not replaced. Only the capacitive switch passed the error criterion with $p = 0.001006$ (2.61% error rate), whereas the mechanical switch with 9.5% ($p = 1$) error rate and the 3D switch with 6.08% ($p = 0.13$) error rate did not pass the criterion for switch location.

For correct number of activations, all three switches had significantly less success than in 90% of the cases: $p < 0.001$ (19.7% error rate) for mechanical switches, $p < 0.001$ (36.60%) for capacitive, and $p < 0.001$ (41.84%) for 3D switches. None of the switches passed the error rate criterion in this case.

The time on task data was not distributed normally with $W = 0.7764$, $p < 0.001$ (Shapiro – Wilk – Test), and homogeneity of variances was not given with $F (2/465) = 5.61$, $p = 0.003933$. In order to investigate if there were any effects in the data set, a robust two-factorial ANOVA was calculated at first and afterwards, non-parametric tests were applied. A main effect for switch type became significant with $F (2/76) = 62.38$, $p < 0.001$, $\eta^2 = 0.08$, and an interaction effect between switch type and measurement point (driving maneuver) with $F (6/228) = 2.42$, $p = 0.012$, $\eta^2 = 0.01$.

The code implemented by Mangiafico (2015) was used as a robust post-hoc procedure, with Holm correction applied. The posthoc test for the effect on switch type illustrated that the significant differences were either between capacitive and mechanical ($p < 0.001$) or 3D and mechanical switches ($p < 0.001$), with mechanical switches having the shortest mean time on task; 3D and capacitive switches had very similar mean times on task. To investigate the interaction effect further, an interaction diagram was plotted (cp. Figure 17. *Interaction effect for time on task*).

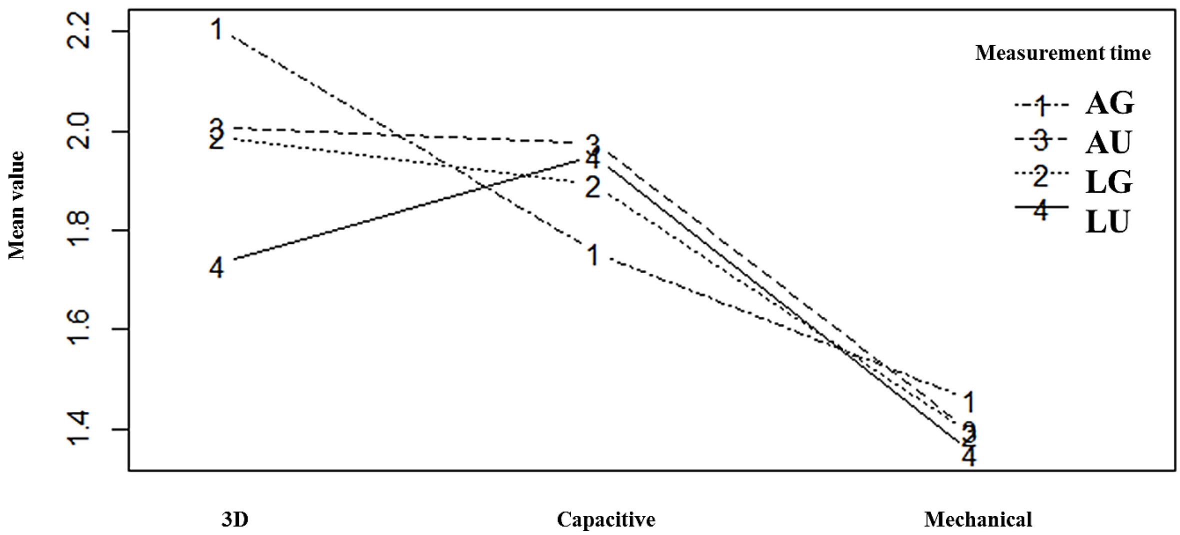



*Figure 17. Interaction effect for time on task. AG = Straight driving on a highway, AU = overtaking on a highway, LG = Straight driving on a rural road, LU = light curves on a rural road.*

Whereas the time on task for the mechanical switch remain similar over all four measurement points, the capacitive time on task increases and stagnates at a higher level for the last two measurement points, compared to the first one. The time on task for the capacitive switches were therefore reversely influenced by the measurement point. When the task became more difficult, the time on task increased. The exactly opposite effect can be found for the 3D switches: with increasing familiarization with the switches, the time on task decreases, having the fastest time on task in the most difficult situation. The interaction effect therefore points in the direction of a learning effect.

### *4.2 Driving performance*

In order to investigate whether there was a difference between the switch settings in terms of driving performance, efficiency and effectiveness were measured. Whereas effectiveness was operationalized via the driving parameters mean velocity, standard deviation of mean velocity, lane keeping quality, lane position, steering wheel angle velocity and steering wheel angle acceleration, efficiency was operationalized via the NASA-TLX between the baseline drive and the performance task.

For the NASA-TLX the assumption of normality was violated (Shapiro – Wilk – Test with $W = 0.93$, $p < 0.001$), but the assumption of homogeneity of variances was met (Levene's Test with $F (3/160) = 0.6$, $p = 0.62$). A robust two-factorial ANOVA for the analysis of workload distribution between the switches did not became significant with $F (3/120) = 1.29$, $p = 0.28$.

The data of the standard deviation of speed was not distributed normally with $W = 0.44$, $p < 0.001$. A robust two-factorial ANOVA was calculated which had a significant effect on driving condition, even with a violation of Mauchly's Sphericity and a resultant Greenhouse-Geisser correction, of $F (3/96) = 153.41$, $p < 0.001$, $\eta^2 = 0.23$. A post-hoc Nemenyi test revealed the follwing significant differences between the driving conditions (cp. Figure 18. *Standard deviation of speed*).

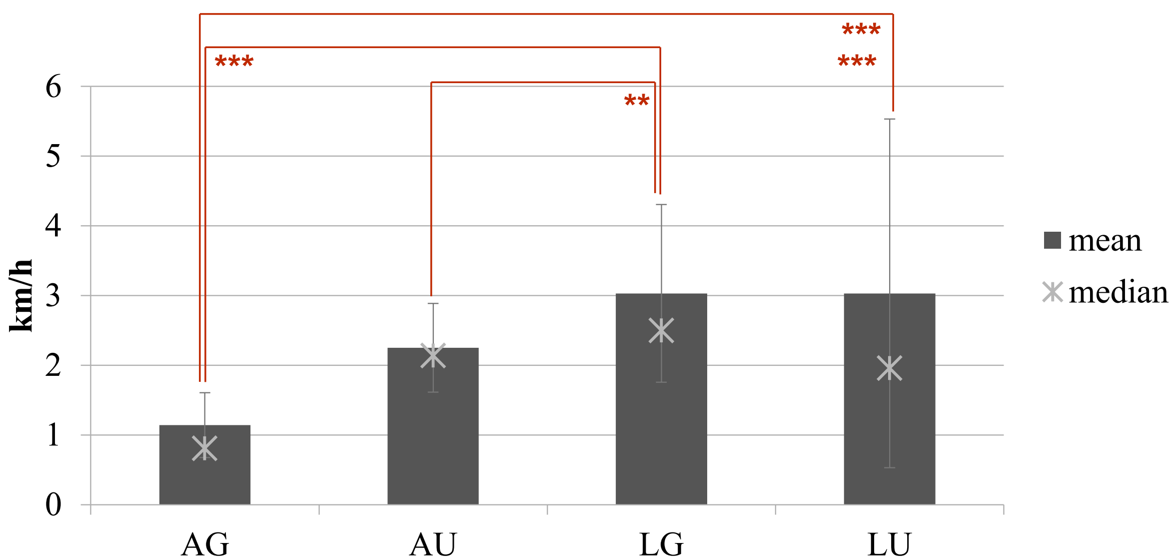



*Figure 18. Standard deviation of speed. AG = Straight driving on a highway, AU = overtaking on a highway, LG = Straight driving on a rural road, LU = light curves on a rural road.*

This implies that the subjects had more difficulties in keeping the speed on the instructed level from one course to the other, emphasizing that the courses had a different level of difficulty. There was also an interaction effect between switch type and task significant after sphericity correction in this variable ($F (9/288) = 4.72$, $p = 0.01$, $\eta^2 = 0.008$). As the main effect on switch type did not become significant, this was not further investigated.

In terms of lane keeping quality, the data was not distributed normally with W = 0.83, $p < 0.001$. A robust two-factorial ANOVA was calculated in lack of a non-parametric procedure and had a significant effect on task with $F(3/96) = 424.00$, $p < 0.001$, $\eta^2 = 0.52$. A robust post-hoc test revealed significant differences between all driving conditions. The results can be viewed in the following table (Table 11. *Lane keeping quality differences for driving condition*).

*Table 11. Lane keeping quality differences for driving condition.*

| | Statistic | p-value | p adjusted |
|---|---|---|---|
| AG - AU | -8.006 | 0 | 0 |
| AG - LG | -1.135 | 0 | 0 |
| AG - LU | -5.137 | 0 | 0 |
| AU - LG | 6.871 | 0 | 0 |
| AU - LU | 2.869 | 0 | 0 |
| LG - LU | -4.002 | 0 | 0 |

For the steering wheel velocity, the data was not distributed normally with Shapiro – Wilk W = 0.8977, $p < 0.001$. A robust two-factorial ANOVA was therefore calculated and had a significant effect on driving condition with $F(3/117) = 2302.60$, $p < 0.001$, $\eta^2 = 0.76$. A robust post-hoc test revealed significant differences between almost all driving conditions and can be found in the following table (Table 12. *Steering wheel velocity differences for driving condition).*

*Table 12. Steering wheel velocity differences for driving condition.*

| Comparison | Statistic | p value | p adjusted |
|---|---|---|---|
| AG - AU | -5.575 | 0 | 0 |
| AG - LG | -5.775 | 0 | 0 |
| AG - LU | -21.99 | 0 | 0 |
| AU - LG | -0.1999 | 0.884 | 0.884 |
| AU - LU | -16.41 | 0 | 0 |
| LG - LU | -16.21 | 0 | 0 |

For the steering wheel acceleration, the data was not distributed normally with Shapiro – Wilk W = 0.94, $p < 0.001$. A robust two-factorial ANOVA was therefore calculated and had a significant effect on driving condition with $F(3/117) = 1116.61$, $p < 0.001$, $\eta^2 = 0.65$. A robust post-hoc test revealed significant differences between almost all driving conditions and can be found in the following table (cp. Table 13. *Steering wheel acceleration differences for driving condition*).

In order to investigate the impact of the steering wheel jerks, their velocity was investigated by employing a combined function of lane keeping quality and steering wheel angle velocity. When a subject deviates heavily from a lane, it is investigated how fast the steering movement occurs, thereby providing insight on the intensity of corrective steering movements. The basis for the calculation is the individual baseline of each subject (the familiarization drive), and 20 cm from this baseline was estimated as a critical deviation.

*Table 13. Steering wheel acceleration differences for driving condition.*

| Comparison | Statistic | p value | p adjusted |
|---|---|---|---|
| AG - AU | - 36.96 | 0 | 0 |
| AG - LG | - 47.49 | 0 | 0 |
| AG - LU | -152 | 0 | 0 |
| AU - LG | -10.53 | 0.056 | 0.056 |
| AU - LU | -115 | 0 | 0 |
| LG - LU | -104.5 | 0 | 0 |

The data of the absolute number of deviations is not distributed normally ($W = 0.8862$, $p < 0.001$). The assumption of sphericity is violated as well for the main effect on measurement time ($W = 0.56$, $p < 0.001$). 13 missing values out of 480 are replaced using the EM Algorithm. A robust two-factorial ANOVA was calculated because of the assumption violations and a significant effect on measurement time occurred with $F(3/117) = 477.13$, $p = 0.001$, $\eta^2 = 0.3$. A robust post - hoc test showed significant differences between almost all driving maneuvers (cp. Table 14. *Significant effects on driving maneuver in numbers of critical lane deviations*). P values were adjusted using the Holm correction.

There was no significant effect on switch type with $F(2/78) = 2.27$, $p = 0.329$.

*Table 14. Significant effects on driving maneuver in numbers of critical lane deviations.*

| Comparison | Statistic | p value | p adjusted |
|---|---|---|---|
| AG - AU | -2.85 | 0 | 0 |
| AG - LG | -1.03 | 0.004 | 0.008 |
| AG - LU | -3.40 | 0 | 0 |
| AU - LG | 1.82 | 0 | 0 |
| AU - LU | -0.55 | 0.132 | 0.132 |
| LG - LU | -2.37 | 0 | 0 |

| Comparison | Statistic | p value | p adjusted |
|---|---|---|---|
| AG - AU | 0.30 | 0.74 | 0.74 |
| AG - LG | 0.86 | 0.244 | 0.72 |
| AG - LU | -2.25 | 0.04 | 0.16 |
| AU - LG | 0.56 | 0.48 | 0.74 |
| AU - LU | -2.55 | 0.02 | 0.1 |
| LG - LU | -3.10 | 0 | 0 |

In terms of steering wheel jerk impact, the maximal deviation velocity data is not normally distributed with $W = 0.9803$, $p < 0.001$. The assumption of sphericity was violated for the main effect on measurement time ($W = 0.71$, $p = 0.021$) and for the interaction effect between measurement time and switch type ($W = 0.34$, $p = 0.004$). 160 out of 480 values are missing and are replaced with the EM Algorithm. A robust two-factorial ANOVA is calculated and a significant occurs for each main effect, but not for the interaction between both effects. The main effect on measurement time is significant with $F(2/80) = 15.6174$, $p = 0.001$, $\eta^2 = 0.01$. The robust Holm-corrected post-hoc test results are displayed in the table below (cp. Table 15. *Significant post-hoc results for the main effect on driving maneuver*).

The main effect on switch type does not become significant with $F(2/76) = 2.2168$, $p = 0.334$, $\eta^2 = 0.009$. If the sample size is increased via generating additional cases while keeping the mean and median stable, the differences between capacitive and both mechanical and 3D switch become significant (see Figure 19. *Differences per switch type for steering wheel jerks*).

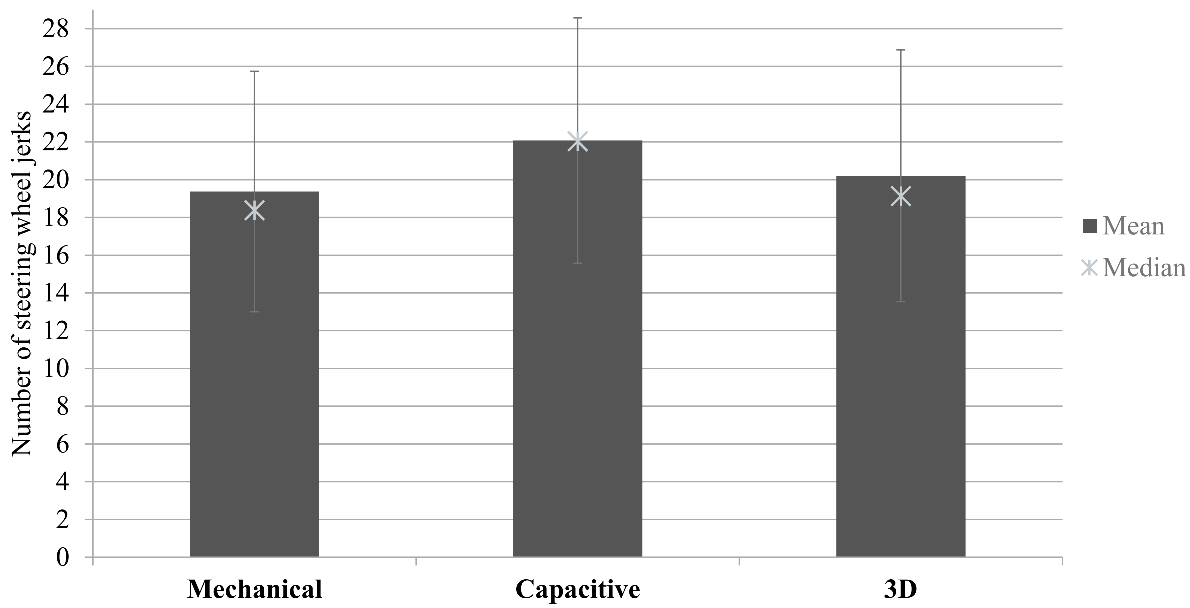



*Figure 19. Differences per switch type for steering wheel jerks.*

### *4.3 Subjective preference data*

As with the previous experiments, the subjective preferences were analyzed in terms of the good – bad – scale and the distance to the ideal configuration. Additionally, the results of the reference instrument are reported.

The following paragraph reports the results of the good – bad – scale. The assumption of normality was violated (Shapiro-Wilk $W = 0.97$, $p = 0.00476$), but the assumption of homogeneity of variances was met (Levene‘s Test with $F(2/120) = 0.24$ $p = 0.7876$). A Friedman’s ANOVA was calculated and did become significant with $\chi^2$ (2) = 22.31, $p < 0.001$, $\omega^2 = 0.15$. A posthoc Nemenyi Test revealed significant differences between mechanical and both capacitive ($p < 0.001$) and 3D switch ($p < 0.001$). For the case of only taking the last measurement point into account, the posthoc test did not become significant.

In terms of distance to ideal, the assumption of normality was violated (Shapiro-Wilk W = 0.97, $p < 0.001$), and the homogeneity of variances was not given (Levene's Test with F (3/2292) = 39.96, $p < 0.001$). A robust two-factorial ANOVA was therefore calculated. Two main effects and an interaction effect between the two factors emerged significant after sphericity corrections. The main effects were the measurement scale (F (13/520) = 274.13, $p < 0.001$; $\eta^2 = 0.054$) and the switch configuration (F (3/120) = 121.15, $p < 0.001$; $\eta^2 = 0.12$) and the interaction between scale and switch configuration (F (39/1560) = 566.38, $p < 0.001$; $\eta^2 = 0.17$). Both main effects were analyzed further with a robust post-hoc Hommel corrected t-test. For the switch configurations, highly significant differences to the ideal configuration for both capacitive ($p < 0.001$) and 3D switch ($p < 0.001$) emerged, as well as highly significant differences to the mechanical configuration for both capacitive ($p < 0.001$) and 3D switch ($p < 0.001$). The differences between the different measurement scales revealed highly significant differences between various configurations which can be found in appendix c) e *Posthoc tests for the main effects*, as well as the other aforementioned calculations.

To investigate the interaction effects further, a profile was extracted from the data (s. Figure 20. *Profile of the distance to ideal*).

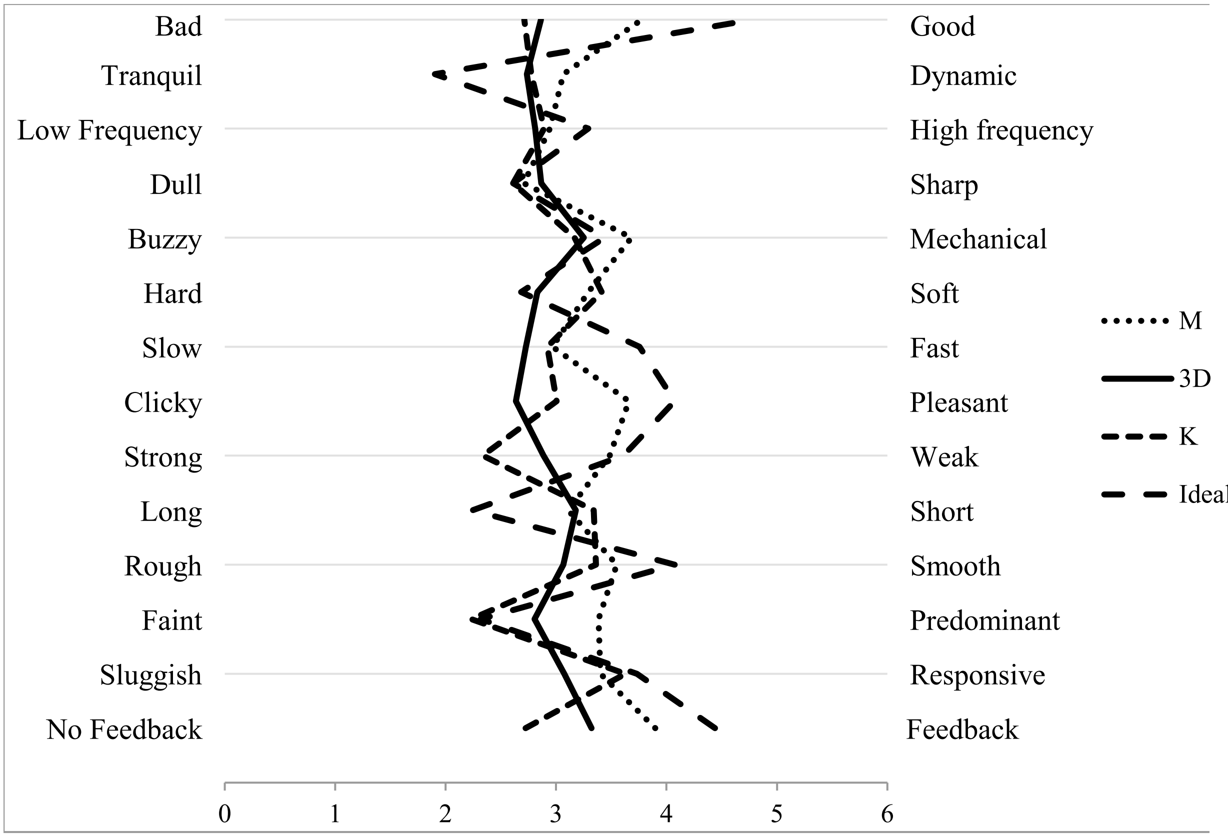



*Figure 20. Profile of the distance to ideal.*

As the profile depicts, the interpretability of the main effects cannot be assumed because the course of the lines are drastically varying (*disordinal* interaction). A possible way to further analyze the results would be, although to be conducted with caution, building up a sum score of the measurement scales, where the significant differences between the switch configurations and the ideal scale occur. With this procedure random effects can be prevented from the analysis. This sum score of the significant differences to the ideal led to the following results (cp. Figure 21. *Sum Scores of the significant differences to the ideal configuration*):

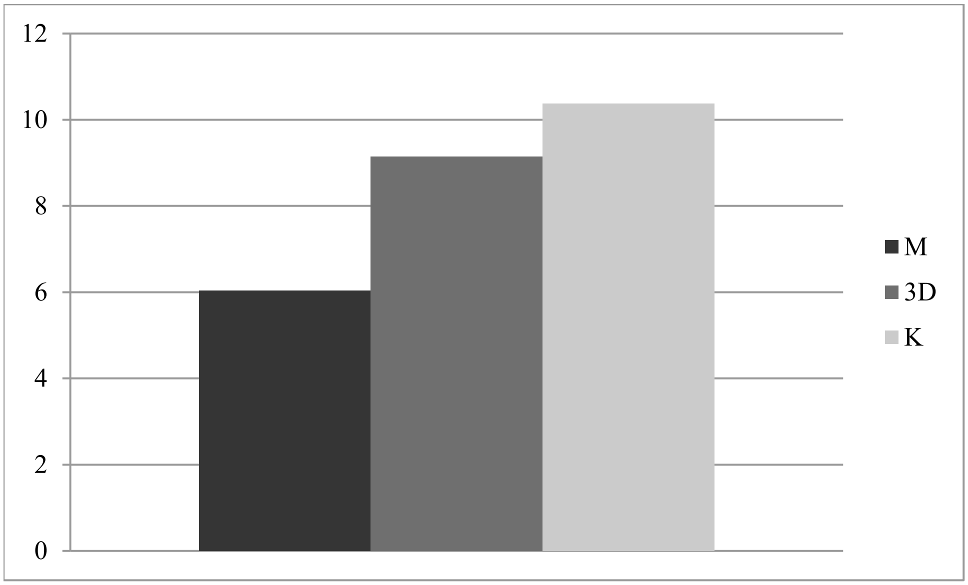



*Figure 21. Sum Scores of the significant differences to the ideal configuration. M = Mechanical, 3D = 3D switch, K = Capacitive. A smaller value is better on this scale.*

The interpretation of this sum score is, however, difficult and the data basis for the calculation of any statistic method too unreliable to be justified. The sum score graphic can therefore only be interpreted as a trend towards mechanical being rated better than 3D and capacitive switch, and the 3D switch being rated better as the capacitive switch.

The parallel test instrument data was not distributed normally with Shapiro-Wilk $W = 0.99$, $p < 0.001$, but homogeneity of variances was given (Levene's test with $F (2/1227) = 2.66$, $p = 0.07049$. A sphericity-corrected two-way ANOVA became significant with a main effect on measurement scale ($F (9/360) = 5.03$, $p < 0.001$, $\eta^2 = 0.04$) and an interaction effect between switch configuration and scale ($F (18/720) = 3.94$, $p < 0.001$, $\eta^2 = 0.05$). No effect emerged for the different switches.

### *4.4 Principal Component Analysis*

As the data of the questionnaire was not distributed normally (cp. chapter *Subjective preference data*), the correlation matrix was calculated with a Spearman correlation. Fourteen items were rotated via varimax for the principal component analysis in a first step, in conformity with the previous studies. The Bartlett test became significant with $\chi^2$ (91) = 2502.19, $p < 0.001$. The sampling adequacy was ensured with a Kaiser-Meyer-Olkin criterion (KMO) = .83, which showed a clear improvement compared to the previous studies. The scree plot showed ambiguous results in the present data set in terms of factor determination (cp. Figure 22. *Scree plot of main experiment*).

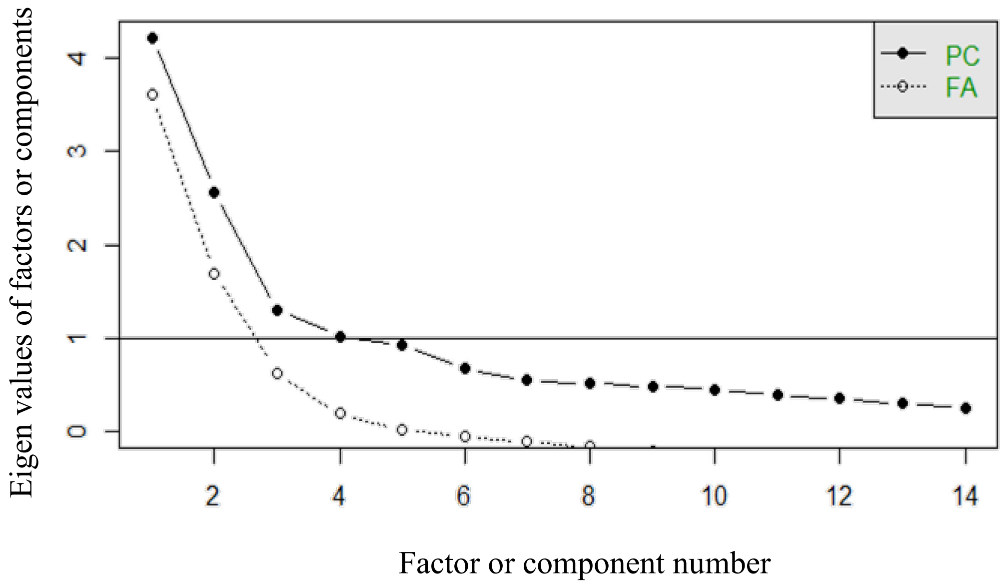



*Figure 22. Scree plot of main experiment. PC stands for PCA, whereas additionally a scree plot for an eventual factorial analysis (FA) was plotted, but can be disregarded.*

Applying different and more detailed measures for factor determination led to ambiguous results. The VSS argued for two factors, whereas Velicer MAP pointed in the direction of three factors. The empirical BIC and Kaiser's Criterion were in favour of a four-factor structure, whereas the sample size adjusted BIC pointed in the direction of five factors (cp. appendix c) f *Factor determination*). As an underestimation of the factors by VSS was very likely, the PCA was conducted for three, four and five factors in order to determine the best factor solution. The full three analyses can be viewed in appendix c) g *Principal component analysis with five, four and three factors.*

The model fit of all three analyses is very equal with three factors at 0.95, four factors at 0.94, five factors at = 0.96, and in an acceptable range. The communalities for three or four factors are not always > .6, which should not be the case. Additionally, with three factors, there is no dimension on which the item *mechanical – buzzy* is loading; this item is simply ignored. The residual sum of squares, a measure for the fit of the data to the model, is also not in favour for either the three- or the four factor solution. On the other hand, the five factor solution with a varimax rotation has two single-item factors, which is a situation that should be avoided for a reliability analysis of a questionnaire.

To investigate the overall fit of the four factor solution, the method via employing Cronbach's α described in chapter *Third preliminary experiment: Principal component analysis* was applied. Subscales were calculated for four factors, and it was intended to compare each of their Cronbach's α to an overall Cronbach's α of the whole scale (s. appendix c) h *Cronbach's α of the main experiment* for full data). As the Guttman's Lambda 6 value is constantly higher than the Cronbach's α, this in another indicator that the scale is not univariate (Revelle, 2011), and points at the inappropriateness for the Cronbach's α as a measure of reliability. To apply a more adequate measure, McDonald's ω was calculated following the procedure recommended by Zinbarg, Revelle, Yovel, & Li (2005). The full calculation is enclosed in appendix c) i McDonald's omega. It is noteworthy that this analysis implies that the scale *mechanical – buzzy* is not related to other scores (cp. Figure 23. *McDonald's omega determined factor structure*), which does not change with enhancing the number of factors. Additionally, if a univariate option would be chosen, the first two scores would not contribute to the scale, although providing an important part of the first factor F1.

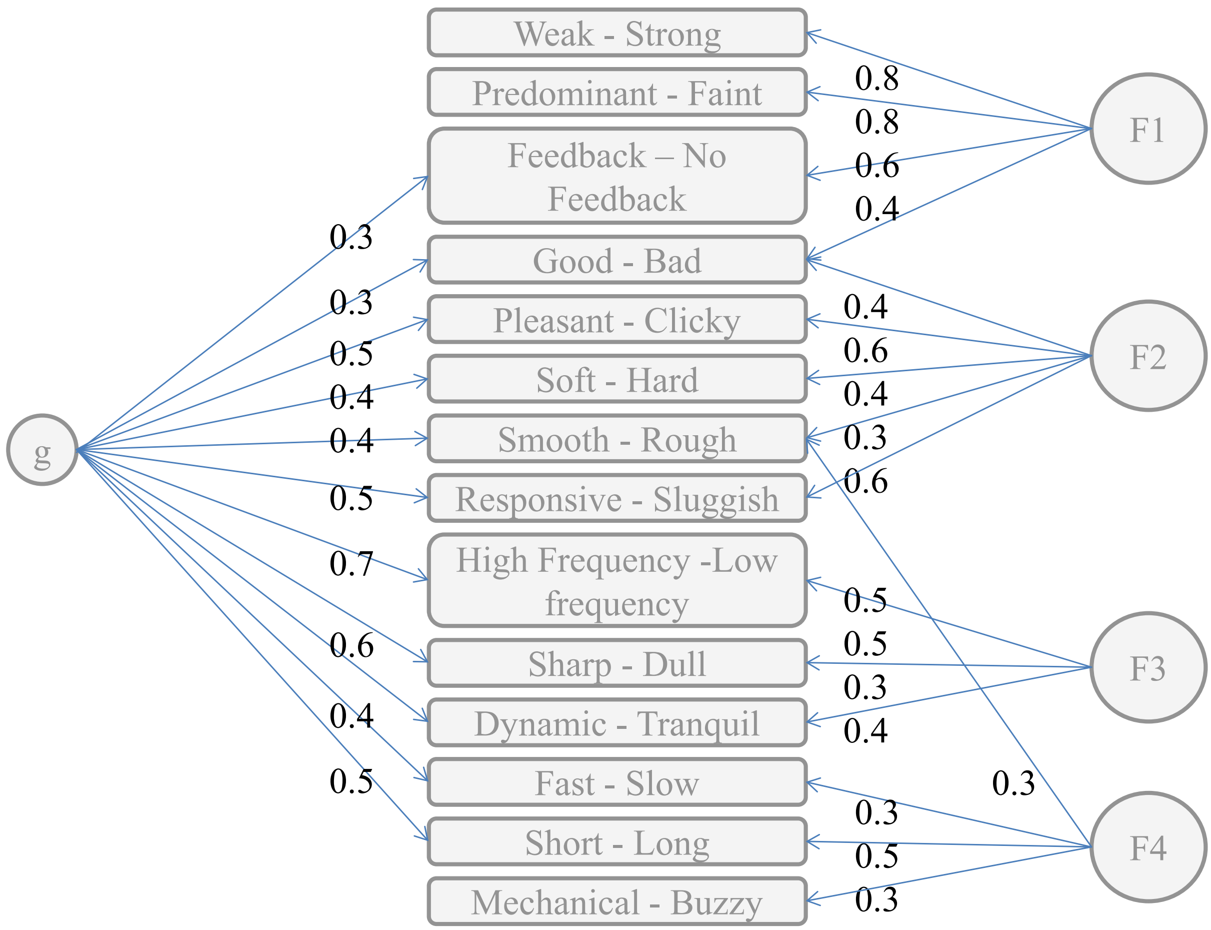



*Figure 23. McDonald's omega determined factor structure.*

To investigate the factor structure of the present data set further, a PCA with five underlying factors was calculated. As the independency of the different factors cannot be postulated given the structure in Figure 23. *McDonald's omega determined factor structure*, an oblique rotation was chosen, *promax* suggested by Thurstone (1927, cited after Abdi (2003)), with five underlying factors. The following factor structure emerged from this approach (cp. Table 16. *Factorial structure of final grid*):

*Table 16. Factorial structure of final grid.*

| | PC2 | PC3 | PC1 | PC5 | PC4 | h2 | u2 |
|---|---|---|---|---|---|---|---|
| Weak - Strong | 0.87 | | | | | 0.72 | 0.28 |
| Predominant - Faint | 0.87 | | | | | 0.75 | 0.25 |
| Feedback – No Feedback | 0.75 | | | | | 0.63 | 0.37 |
| Good - Bad | 0.52 | 0.51 | | | | 0.77 | 0.23 |
| Pleasant - Clicky | -0.42 | 0.85 | | | | 0.71 | 0.29 |
| Soft - Hard | | 0.83 | | | | 0.73 | 0.27 |
| Smooth - Rough | | 0.61 | | 0.31 | | 0.55 | 0.45 |
| Responsive - Sluggish | | 0.57 | | 0.34 | | 0.67 | 0.33 |
| High Frequency - Low frequency | | | 0.97 | -0.37 | | 0.74 | 0.26 |
| Sharp - Dull | | | 0.75 | | | 0.64 | 0.36 |
| Dynamic - Tranquil | | | 0.55 | | | 0.71 | 0.29 |
| Fast - Slow | | | | 0.97 | | 0.77 | 0.23 |
| Short - Long | | | 0.46 | 0.54 | | 0.66 | 0.34 |
| Mechanical - Buzzy | | | | | 1 | 0.98 | 0.02 |

The factor structure corresponds to the one proposed by the calculation of McDonald's ω, with *mechanical – buzzy* having an own score. The overall contribution of this score is very small (0.02), which also corresponds to McDonald's ω. The full analysis is enclosed in appendix c) j *Promax rotation with five factors.*

The number of residuals and their root mean square were calculated in order to investigate the fit further. An analysis of residuals shows that the proportion of the total number of residuals is at 34%, which is well within the acceptable limit (cp. Field et al., 2012). The root mean square residual is at 0.059, which is likewise within the acceptable limit (cp. Field et al., 2012).

The *good – bad* score is nearly equally loading on factor 2 or factor 3, which makes the interpretation and the allocation of this item particularly difficult. And although the correlation with the *pleasant – clicky* scale is the highest (cp. appendix c) k *Interitem Correlations*), the loading on the factor 2 is just slightly higher than the loading on factor 3, which includes the *pleasant – clicky* scale. When investigating the relationship between factor 3 and factor 2, the correlation between them is only at .28, which implies that the valence of a switch is rated on both factors with a different focus. The scale which accounts mostly for the auditory items, PC1, contains *High Frequency – Low frequency*, *Sharp – Dull*, and *Dynamic – Tranquil*. As these aspects are semantically related to the to the concept of the energetic content of a switch signal, the term *energy* is chosen for this dimension, although it could be argued that as they are all auditory, they might be labeled auditory as well. PC3, its items being part of the *dynamics* scale in the previous study, contains the items *Pleasant – Clicky*, a valence aspect, *Soft – Hard, Smooth – Rough,* and *Responsive – Sluggish*. As all these items describe how fluid or seamless an interaction takes place to a certain extent, the label for this scale is *fluidity*. PC5, previously related to the *dynamics* scale as well, contains the items *Fast – Slow* and *Short – Long*, clearly related to *velocity* and therefore labeled accordingly. The *potency* scale (PC2) and the *virtuality* scale (PC4) remained the same compared to the previous study. The final factor structure is as follows (cp. Figure 24. *Final Factor Structure*):

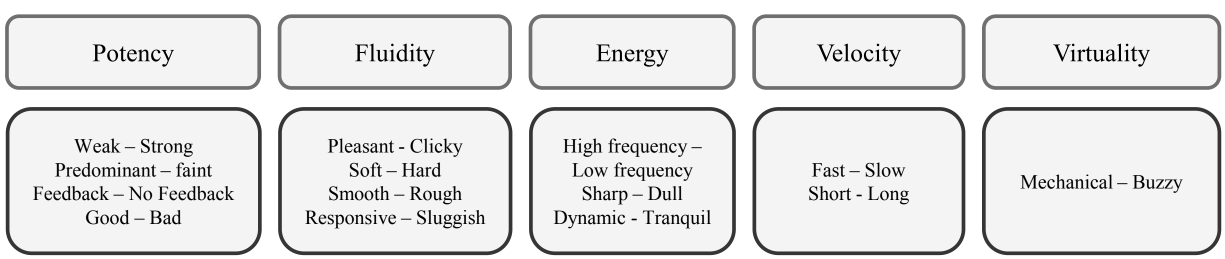



*Figure 24. Final Factor Structure.*

### *4.5 Eyetracking data*

The distraction during the different tasks was defined as percent dwell time on two AOIs, being the road AOI and the steering wheel AOI. Both AOIs were evaluated separately.

Two subjects had to be excluded from the data entirely because the eyetracking equipment did not work properly. With three subjects data was missing for one or two conditions. To keep these subjects in the analysis, missing data was replaced using the EM-algorithm which was discussed earlier. The percentage of single missing data points was exactly the same for the road and steering wheel AOI (1.04%). This data was replaced using the EM-algorithm as well.

For the AOI road, the data was not distributed normally (investigated via Shapiro-Wilk with $W = 0.77$, $p < 0.001$, but the homogeneity of variances was givenen via Levene's test with $F\ (3/\ 756) = 1.03$, $p = 0.3764$. Many outliers

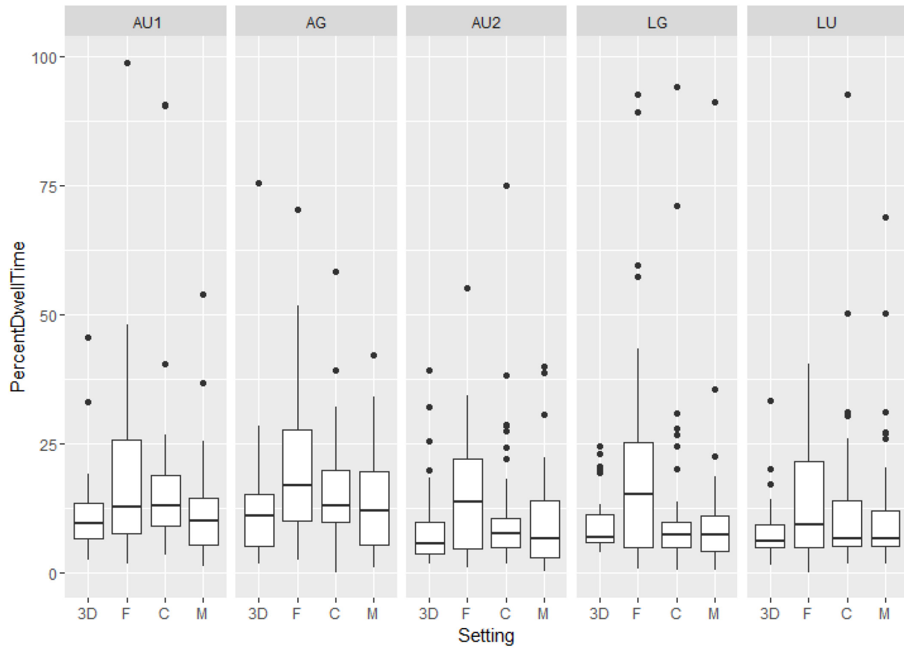



*Figure 25. Boxplots for steering wheel AOI. Driving situations from left to right: First overtaking situation on a highway, straight driving on a highway, second overtaking situation on a highway, rural road straight driving, and rural road slalom. Settings: 3D = 3D switch, F = Familiarization, C = Capacitive, M = Mechanical.*

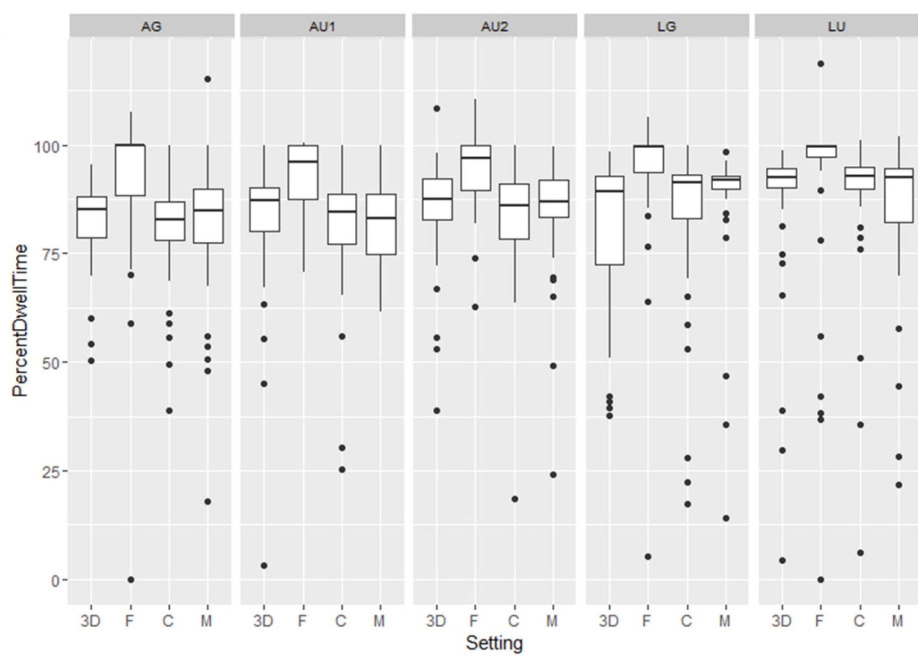



*Figure 26. Boxplots for road AOI. Legend equivalent to Figure 25.*

are present in the data set; as different subjects may have different glance behavior indeed, these remained in the data set.

To investigate differences between the gaze behaviors, a robust two-way ANOVA was calculated. A significant effect could be found for switch type ($F$ (3/ 756) = 185.81, $p < 0.001$; $\eta^2 = 0.06$) and the driving condition ($F$ (3/ 756) = 64.10, $p < 0.001$, $\eta^2 = 0.01$). The interaction between both variables did not become significant with $F$ (3/ 756) = 20.12, $p = 0.095$.

A robust Holm-corrected post-hoc procedure was employed. Significant differences are displayed in the graphic below (cp. Figure 27. *Post-hoc results for road AOI*).

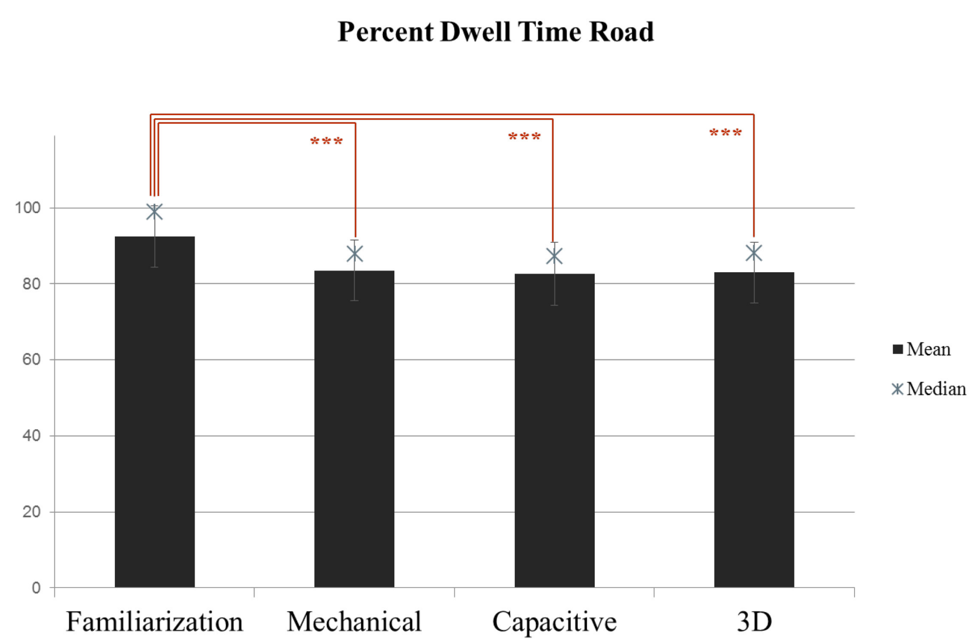



*Figure 27. Post-hoc results for road AOI.*

For the AOI Steering wheel, neither the precondition of normal distribution (Shapiro-Wilk with $W = 0.71$, $p < 0.001$) nor the homogeneity of variances (Levene's test with $F$ (3/756) = 8.66, $p < 0.001$) was given. The outliers remained in the present data set as well. To investigate differences between the gaze behaviors, a robust two-way ANOVA was calculated. A significant effect could be found for switch type ($F$ (3/111) = 557.50, $p < 0.001$; $\eta^2 = 0.038$) and the driving condition (F

(4/148) = 33.78, $p < 0.001$, $\eta^2 = 0.017852623$). The interaction between both variables did become significant with F (12/444) = 43.27, $p < 0.001$, $\eta^2 = 0.007$).

The significant mean and median differences are displayed in the graphic below (cp. Figure 28. *Median and mean of percent dwell time on steering wheel AOI*).

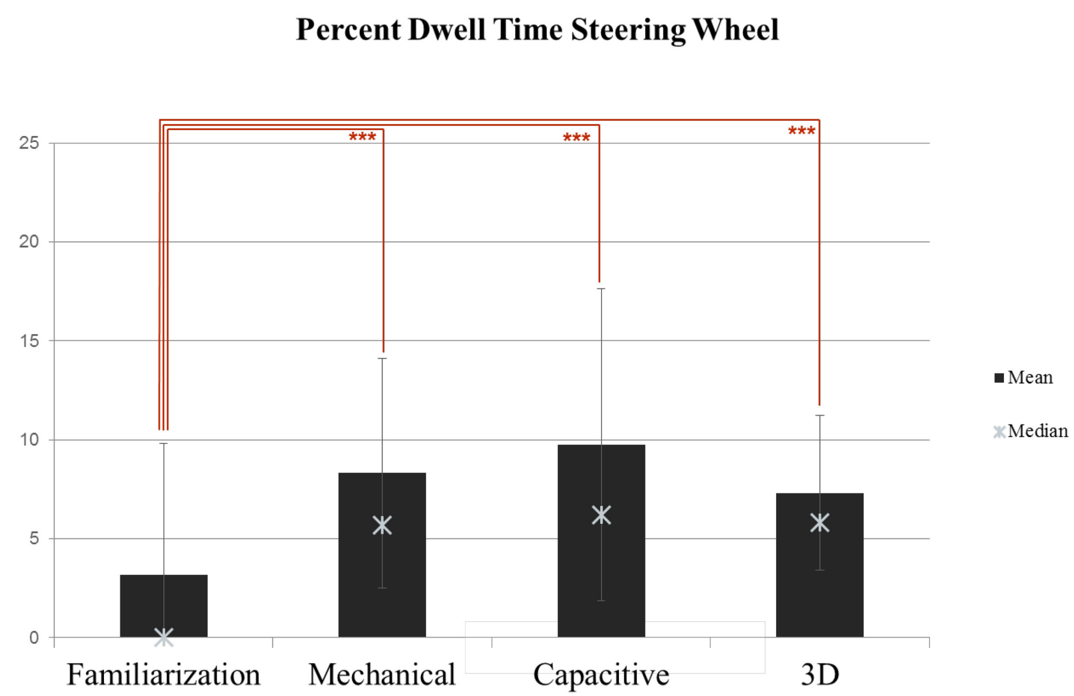



*Figure 28. Median and mean of percent dwell time on steering wheel AOI.*

Other measures investigated include the mean glance duration, the maximum glance duration and the number of glances over 2 s. None of these measures were distributed normally (mean glance duration W = 0.44, maximum glance duration W = 0.46, number of glances over 2s W = 0.50; for all DVs, $p < 0.001$). The homogeneity in the data set was only given for the number of glances over 2s with F (3/736) = 1.72, p= 0.162. For both mean and maximum glance duration, this precondition was violated with F (3/741) = 3.12, p= 0.025, and F (3/736) = 11.117, $p < 0.001$, respectively.

Both maximum glance duration and number of glances over 2 seconds did become significant in a robust ANOVA with F (3/108) = 8.9426, p = 0.034 and F (3/108) = 13.8235, p = 0.005, respectively, but did not lead to significant results in the post-hoc tests. This is most likely due to the fact that the robust ANOVA algorithm still calculates with a mean value. In contrast to that, the median was used for calculation of the post-hoc tests, a more conservative approach. Fisher's exact test comparing the number of glances over 2 seconds for the different switch types led not to significant differences with p = 0.4888 likewise.

The mean glance duration did reveal a significant effect between certain conditions. A robust ANOVA became significant with $F(3/108) = 18.3354$, $p = 0.001$, $\eta^2 = 0.08$. The post-hoc results are displayed in the figure below (Figure 29. *Mean glance duration on the steering wheel per switch type*).

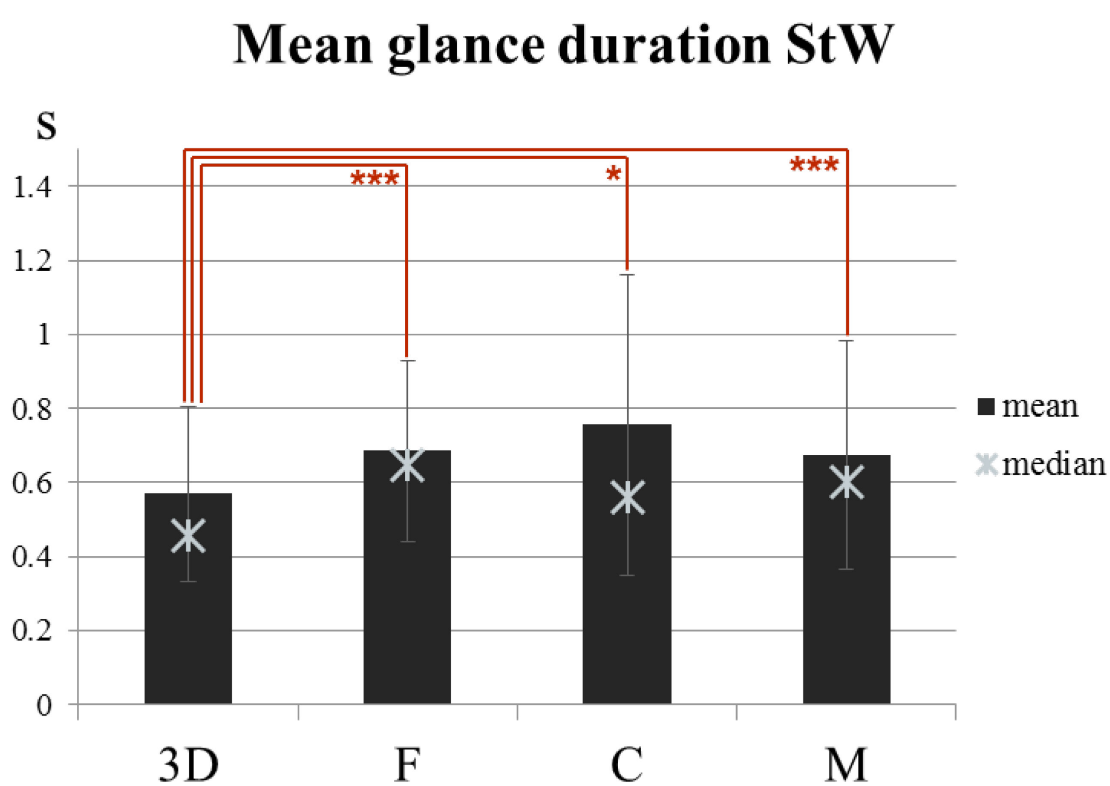



*Figure 29. Mean glance duration on the steering wheel per switch type. 3D = 3D switch, F = Familiarization, C = Capacitive, M = Mechanical.*

## 5. Discussion

The present work aims at generating a method for the measurement of subjective preference in the area of automotive multimodal feedback. This instrument is designed to fill the gap between quantitative and qualitative research by establishing items with users in a controlled comparison, allowing qualitative data to be added for each new study and for its evaluation and interpretation in a quantitative way. The evaluation of a 3D switch compared to mechanical and capacitive ones is its second goal, both with subjective preference and workload data as well as objective data in terms of eyetracking, driving, and performance data. Each part or the results is discussed in the following.

In terms of method generation, the fifth hypothesis (*The MVSS is a valid measurement of acceptance*) is to be investigated. The sample can be established as adequate (KMO – criterion) and the sphericity assumption is met (Bartlett's test). The principal component analysis could therefore be used to investigate underlying factors. As the data was usually not distributed normally, Spearman correlations are employed for the component analysis. The underlying factor structure was investigated very conservatively, especially as some of the results of the third preliminary experiment were ambiguous, as they are based on a very small subject size. Through a supplementary analysis with McDonald's ω, the factor size of the principal component analysis was verified. Five factors are therefore chosen in the analysis. The retrieved factors are labeled *potency, fluidity, energy, velocity* and *virtuality.*

Regarding reliability, the stability of the factors is ambiguous. Compared to the model developed in the previous studies, the dimensions changed to a certain degree, which is influencing the stability aspect of reliability. The *dynamics* dimension seemed to be split now between two different aspects of it, *fluidity* and *velocity,* which are correlated between .4 and .5 with each other. The addition of the *energy* scale is due to the addition of an auditory component and its items through the psychoacoustic study. *Potency* and *virtuality* remained stable over the studies.

Concerning internal consistency, communalities are above .6, which is seen as a proper value in terms of quality of the data set (Field et al., 2012). The general fit of the chosen model is at .96, which meets the quality criterion that the model fit should be > .9. The reliability according to McDonald's ω equals 0.87, which is in the upper range of good reliability. Internal consistency can therefore be assumed. This is also supported by a lower BIC value of the factor structure compared to a general model and a lower root mean square of residuals of the factor structure compared to the one of a general model. As the test is rather short and has only a few items, the odds of obtaining this level of reliability are rather low. The number of subjects might also be an issue, as in the third preliminary study with the fewest subjects, the quality criteria diminished somewhat. It is therefore recommended to use a higher number of subjects than 30, if possible, when applying this instrument.

In conclusion the reliability of the instrument can be assumed, although the addition of items or a possible exclusion of one scale might influence the overall reliability of the instrument.

A more difficult point is the discussion on validity. Content validity was aimed at through the development of the questionnaire data with users in direct interaction with the product. In terms of concurrent validity, the results of the MVSS and the scale developed by Wellings, T. et al. (2009) do not point in a similar direction. It can therefore not be assumed that they measure the same thing. As a result the scale cannot be used to determine concurrent validity of the MVSS. Ecological validity is intended to be maximized through a driving simulator environment. By selecting subjects from each age group, randomizing the presentation of the stimuli and testing drivers only, it was aimed at optimizing the external validity.

The scale has so far only been used in a German and an English speaking setting. The latter part has only been used in the first iteration and regarding the auditory items with native English speakers. Although there has been a cross-check with native English speakers for the remaining items, this can only be regarded as preliminary validated. The items do not apply on other cultural contexts so far.

For generating new items the dyad method did not bring up new items, although Rosenberger (2015) does state this method as being easier for the subjects. A possible explanation could be that in order to generate items for psychophysiological applications, more examples are needed by the subjects to imagine an adequate item pair. It is therefore recommended that triads are used for additional item generation, if the instrument is enhanced in future studies. The final instrument can be found in appendix c) d*) Final* Instrument, both in a German and an English version.

For the switch comparison, driving performance, switch performance, preference and distraction were evaluated in order to allow for an evaluation which is as holistic as possible. The first hypothesis *(There is a difference between the switches in terms of driving performance)* has to be evaluated carefully. The subjective workload as a measure of efficiency shows no differences between the different driving maneuvers or the different switches. This might be due to the fact that subjects could have had problems with evaluating their own performance. The objective driving data does show two effects: firstly, the driving data does show significant differences for the measurement points throughout the driven course. This implies that the difficulty of the different maneuvers was varied as intended. Secondly, the capacitive switch shows higher steering wheel jerks than mechanical or 3D switch, but not to an extend where these differences are significant. From the subjective data on driving performance, the null hypothesis cannot be rejected, this would also not be possible from the objective data. As the subjective data is prone to several biases (e.g. recall bias), and the differences surface for various driving parameters, the null hypothesis cannot be rejected for the time being, but this must be handled with caution, as the steering wheel jerk differences do become significant with a larger, bootstrapped sample.

The switch performance data revealed somewhat contradictory results, which makes the evaluation of the second hypothesis (*There is a difference in performance between the switches*) rather difficult. None of the switches did pass the error criterion in the number of correct activations; for location, only the capacitive switch passed the error criterion, and in the time on task condition, the mechanical switch performed significantly better than both capacitive and 3D switch, with both showing a different behavior over time. The capacitive switch may have been biased in terms of sensing area, as an electric field is induced, which might be an explanation for its superior performance in terms of location. The time on task for capacitive increased over measurement point and therefore driving time, whereas the the 3D switch decreased in the same order of maneuvers. The first trend might be interpreted as a trend towards worse performance of capacitive switches under increasing driving difficulty. A similar trend was found in Ahmad et al. (2015), although the decrease surfaced due to deteriorating road surface characteristics, not driving maneuvers. The second observed trend can be interpreted differently. Mechanical switches, though not always steering wheel switches, are used many times a day by every average person. Capacitive surfaces become more and more present to an extent, that this is also a device subjects are familiar with. 3D switches are, on the other hand, a completely novel interaction form. The time on task data reflects that with more practice, and even with increasing primary task difficulty, subjects become faster in the interaction with them. Although there are differences in performance between the switches, they are therefore contradictory. A sound conclusion can only be drawn inasmuch as that mechanical switches are faster than both capacitive and 3D, so the null hypothesis cannot be rejected for all measures of switch performance.

In terms of the third hypothesis (*There is a difference in distraction between the switches*), only significant differences between the baseline and all three switch types can be found for the percent dwell time on both steering wheel and road. Differences in glance time only emerge for the mean glance duration, not for the maximum glance duration. As the mean times are all below 2 seconds, the significant difference in mean glance duration is existent, but its external power is not very large. The number of glances over 2 seconds did not reveal any significant differences between the switches, not in terms of mean values, nor in terms of frequency analysis. The eyetracking behaviour might change when a more complex task is employed with the different kind of switches. Until this is furtherly investigated, the null hypothesis cannot be rejected, with the exception of the mean duration of glances, where the 3D switches outperform mechanical and capacitive switches significantly.

The preference results are less ambiguous. The results of the scale developed in this work differ from the results obtained in the instrument by (Wellings, T. et al., 2009), which implies that they measure something differently. As the method generation of items on the instrument by (Wellings, T. et al., 2009) is expert-based and not based on subjects directly interacting with the switches and the technology is different, this result is not highly unlikely. The good – bad results point in the direction of subjective preference towards the technology the subjects are most used to: mechanical switches. Both capacitive and 3D switches perform equally and significantly less. In terms of distance to the ideal, the differences which emerge significantly to the ideal vary: mechanical switches have the least, 3D switches are somewhere in the middle, and capacitive switches have the most. It is statistically not possible to determine whether these number of significant differences to the ideal are significant, as more than one row would be needed for an exact Fisher test. As only the count of statistical differences is taken into account for this score, it can be argued that this score is valid because it is already based on a significant data set. In lack of another investigation method, the fourth hypothesis, *There is a difference in preference between the switches,* is kept and the null hypothesis is rejected.

A few drawbacks apply to these results. The effect sizes of the reported results are comparatively small to normal statistical analysis. Small effects can, depending on the context, nevertheless have strong impacts (Rosenthal, 1993). In a safety-critical environment even consequences from small effects should be applied, as long as they constitute an improvement to the status quo.

The testing took place in a static driving simulator and the applicability of them is limited, as noise and road vibration do have an impact on the interaction with the switches. Two of the main benefits of 3D switches, accurate location and force measurement, develop their full potential on real roads only. The accidental operations of capacitive switches occur due to road surfaces, especially roughness (Ahmad et al., 2015), and the capacitive feedback based on auditory signals is masked more by environmental noise on a road than in a driving simulator. Instead of using a screen, the normal capacitive surface, a normal switch surface was used for better comparability in addition. A screen has a higher distraction potential than a switch overlay. Capacitive switches gave feedback only via sound, whereas mechanical switches gave feedback via sound and movement and 3D switches via sound and vibrotactile feedback. This difference occurred because of the external validity of the study. In reality most automotive applications do not inherit tactile feedback, and if they do, the feedback is delayed remarkably. It is therefore possible that the results were biased because on the unimodal feedback of the capacitive switch. Furthermore, the sensing principle of the capacitive switch biases the results in its favor; the sensing area of the 3D switch was exactly the 14 $mm^2$ field on the switch, whereas the electric field, which was induced by the capacitive switch, was most likely larger than that. As the second preliminary study has shown, subjects prefer less force application for the 3D switches, even less force than contemporary mechanical switches. This might be explained due to the increasing presence of capacitive interaction surfaces, which may change the way interaction in a car will take place in the near future. A bias towards the positive evaluation and the positive performance of the capacitive switch can therefore not be ruled out.

Concludingly, the instrument developed over the different studies has shown that it covers the most important aspects of tactile and auditory feedback for virtual steering wheel switches. It remains open to variations if other technologies or use cases are under investigation. The recommended procedure for adding new scales is employing the Repertory Grid technique using triads. For the technology under investigation, a benefit over capacitive switches could be identified, although the experimental setup was biased in favor of capacitive switches to a certain degree. Further work should include the application of the instrument to other domains than the automotive context and the investigation of 3D switches on a real road, maybe even with added functionality.

## Literature

1-Consciousness-Sense-Touch-Anatomy-Receptors. (2012, February 5). Retrieved July 8, 2013, from http://www.outline-of-knowledge.info/OKD/1/Consciousness/Sense/Touch/ Anatomy/Receptors/1-Consciousness-Sense-Touch-Anatomy-Receptors.html

Abdi, H. (2003). Factor rotations in factor analyses. *Encyclopedia for Research Methods for the Social Sciences. Sage: Thousand Oaks, CA*, 792–795.

Agawi (2013). TouchMarks II: Touchscreen Latencies in Flagship Tablets. Retrieved from http://appglimpse.com/blog/touchmarks-ii-touchscreen-latencies-in-flagship-tablets/

Ahmad, B. I., Langdon, P. M., Godsill, S. J., Hardy, R., Skrypchuk, L., & Donkor, R. (2015). Touchscreen usability and input performance in vehicles under different road conditions: an evaluative study. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 47–54). ACM.

Allin, S., Matsuoka, Y. & Klatzky, R. (2002). Measuring Just Noticeable Differences For Haptic Force Feedback: Implications for Rehabilitation. In *Proceedings of the 10th Symp. On Haptic Interfaces For Virtual Envir. & Teleoperator Systs* (p. 4). Orlando, FL, USA.

Allport, D., Antonis, B., & Reynolds, P. (1972). On the division of attention: A disproof of the single channel hypothesis. *Quarterly Journal of Experimental Psychology*, *24*(2), 225–235. Retrieved from 10.1080/00335557243000102

Altinsoy, M. E. (2008). The Effect of Auditory Cues on the Audiotactile Roughness Perception: Modulation Frequency and Sound Pressure Level. In A. Pirhonen & S. Brewster (Eds.), *Haptic and Audio Interaction Design* (Vol. 5270). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from http://www.springerlink.com/index/10.1007/978-3-540-87883-4

Altinsoy, M. E., & Merchel, S. (2009). Audiotactile Feedback Design for Touch Screens. In M. E. Altinsoy, U. Jekosch, & S. Brewster (Eds.), *Haptic and Audio Interaction Design* (Vol. 5763). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from http://www.springerlink.com/index/10.1007/978-3-642-04076-4

Anguelov, N. (2008). *Haptische und akustische Kenngrößen zur Objektivierung und Optimierung der Wertanmutung von Schaltern und Bedienfeldern für den Kfz-Innenraum*. Value Appeal. Technische Universität Dresden

Badal, J. (2008, June 23). When Design Goes Bad Lessons from some spectacular product flops. *Wall Street Journal*. Retrieved from http://online.wsj.com/news/articles/SB121390461372989357

Bähr, M. & Frotscher, M. (2009). *Neurologisch-topische Diagnostik: Anatomie - Funktion - Klinik*. Stuttgart: Georg Thieme Verlag. Retrieved from http://books.google.de/books?id=Bk2uYfpXwsUC&printsec=frontcover&dq=Neurologische+Anatomie&hl=de&sa=X&ei=xe5oUozwKcup0AXxioDoCg&ved=0CD8Q6AEwAA#v=onepage&q&f=false

Banter, B. (2010). Touch Screens and Touch Surfaces are Enriched by Haptic Force-Feedback. *Information Display*.

Barua, A. (2013, June). *Dynamic Haptic Feedback in Comparing Spatial Information* (Masterarbeit). University of Tampere.

Bech, S., & Zacharov, N. (2007). *Perceptual audio evaluation-Theory, method and application*. John Wiley & Sons.

Bell, R. C. (1990). Analytic issues in the use of repertory grid technique. *Advances in Personal Construct Psychology*, *1*, 25–48.

Bell, R. C., Vince, J. & Costigan, J. (2002). Which vary more in repertory grid data: constructs or elements? *Journal of Constructivist Psychology*, 15, 305-314.

Bell, R. C. (2004). When is my grid cognitively complex and when is it simple? Some approaches to deciding. *Personal Construct Theory & Practice*, *1*, 28–32.

Bhise, V. D. (2012). *Ergonomics in the Automotive Design Process*. Boca Raton: CRC Press. Retrieved from http://books.google.de/books?id=XgVHut1fVJQC&printsec=frontcover&hl=de#v=onepage&q&f=false

Birbaumer, N., & Schmidt, R. (1996). *Biologische Psychologie*. Berlin: Springer.

Bjelland, H. V., & Tangeland, K. (2007). User-Centered Design Proposals for Prototyping Haptic User Interfaces. In *HAID 2007* (pp. 110–120). Berlin Heidelberg: Springer.

Blana, E. (1996). *Driving simulator validation studies: a review* (No. ITS working paper 480) (p. 74). University of Leeds.

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. *Human–Computer Interaction*, *4*(1), 11–44.

Bolanowski, S. J., Gescheider, G., Verrillo, R., & Checkosky, C. (1988). Four channels mediate the mechanical aspects of touch. *Journal of the Acoustical Society of America*, *84*(5).

Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler*. Berlin Heidelberg: Springer.

Bortz, J., Liener, G. A., & Boehnke. (2008). *Verteilungsfreie Methoden in der Biostatistik*. Heidelberg: Springer Medizin.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Jorunal of Behavioral Therapy and Experimental Psychiatry*, *25*(1), 49–59.

Brammer, A. J., & Piercy, J. E. (1998). Rationale for measuring vibrotactile perception at the fingertips as proposed for standardization in ISO CD 13091-1. In R. Lundström, L. Burström, T. Nilsson, & S. Bylund (Eds.). Presented at the Eighth International Conference on Hand-Arm Vibration, Umeå, Sweden.

Brewster, S. A., Wright, P. C., & Edwards, A. D. (1993). An evaluation of earcons for use in auditory human-computer interfaces. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems* (pp. 222–227). ACM.

Brewster, S. A. & Brown, L. A. (2004). Tactons: Structured Tactile Messages for Non-Visual Information Display. In *5th Australasian User Interface Conference* (p. 9). Dunedin.

Brickman, B. J., Hettinger, L. J., & Haas, M. W. (2000). Multisensory Interface Design for Complex Task Domains: Replacing Information Overload With Meaning in Tactical Crew Stations. *The International Journal of Aviation Psychology*, *10*(3), 273–290. http://doi.org/10.1207/S15327108IJAP1003_04

Bubb, H. (2003). Fahrerassistenz - primaer ein Beitrag zum Komfort oder fuer die Sicherheit? / Driver assistance - firstly a contribution to primary safety or rather to comfort? In *Der Fahrer im 21. Jahrhundert. Anforderungen, Anwendungen, Aspekte fuer Mensch-Maschine-Systeme. Tagung der VDI-Gesellschaft Fahrzeug- und Verkehrstechnik* (Vol. 1768, pp. 25–44). Braunschweig: VDI-Verlag.

Burdea, G. (1996). *Force and touch Feedback for Virtual Reality*. New York: Wiley.

Cao, Y., Van Der Sluis, F., Theune, M., & Nijholt, A. (2010). Evaluating informative auditory and tactile cues for in-vehicle information systems. In *Proceedings of the 2nd International*

*Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 102–109). Retrieved from http://dl.acm.org/citation.cfm?id=1969791

Carpenter, P., Just, M., & Rayner, K. (1983). What your eyes do while your mind is reading. In *Eye movements in reading: Perceptual and language processes*. New York: Academic.

Chang, W., Hwang, W., & Ji, Y. G. (2011). Haptic Seat Interfaces for Driver Information and Warning Systems. *International Journal of Human-Computer Interaction*, *27*(12), 1119–1132.

Chen, X., Barnes, T. J., Childs, T. H. C., Henson, B., & Shao, F. (2009). Materials' tactile testing and characterisation for consumer products' affective packaging design, *30*, 4299–4310.

Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, *25*(5), 975–979. http://doi.org/http://dx.doi.org/10.1121/1.1907229

Chun, M. M., & Wolfe, J. M. (2000). Visual Attention. In *Blackwell Handbook of Perception* (p. 66).

Conway, A. R., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review*, *8*(2), 331–335. Retrieved from http://link.springer.com/article/10.3758/BF03196169

Cox, R. W. (n.d.). 1986, Electronics Developed for Lotus Active Suspension Technology. Retrieved July 5, 2013, from http://history.gmheritagecenter.com/wiki/index.php/1986,_Electronics_Developed_for_Lotus_Active_Suspension_Technology

Craig, J. C. (1972). Difference threshold for intensity of tactile stimuli. *Perceptional Psychophysiology*, *11*, 150–152.

Craig, J. C. (1974). Vibrotactile difference thresholds for intensity and the effect of a masking stimulus. *Perception & Psychophysics*, *15*(1), 123–127. Retrieved from http://link.springer.com/article/10.3758/BF03205839

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334.

Dabic, S., Navarro, J., Tissot, J.-M., & Versace, R. (2013). User perceptions and evaluations of short vibrotactile feedback. *Journal of Cognitive Psychology*, *25*(3).

David, H. (1959). The method of paired comparisons. Proceedings of the fifth conference on the design of experiments in army research development and testing, 1–16.

Del-Colle, A. (2013, May 7). Carchaeology: 1986 Buick Riviera Introduces the Touchscreen. Retrieved May 7, 2013, from http://www.popularmechanics.com/cars/news/auto-blog/carchaeology-1986-buick-riviera-introduces-the-touchscreen-15437094

Delprat, B. (2013, June 17). Composition of the cochlear fluids. Retrieved July 24, 2013, from http://www.cochlea.eu/en/cochlea/cochlear-fluids

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum-Likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*.

Devanandan, M.S., Ghosh, S., & John, K.T. (1983). A quantitative study of muscle spindles and tendon organs in some intrinsic muscles of the hand in the bonnet monkey (Macaca radiata). Anatomical Record, 207, 263-266

Dilger, M. . (2000). *Grounded Theory. Ein Überblick über ihre charakteristischen Merkmale.* (Verstehende Forschungsmethoden: Ethnographische Methoden in der Erziehungswissenschaft). Berlin: Freie Universität Berlin.

DIN EN 894. DIN EN 894 3:2000-02, Sicherheit von Maschinen - Ergonomische Anforderungen an die Gestaltung von Anzeigen und Stellteilen - Teil 3: Stellteile; Deutsche Fassung EN 894-3:2000 (2000).

DIN EN ISO 3958. DIN EN ISO 03958 - Personenkraftwagen - Handreichweiten des Fahrzeugführers (1978).

DIN EN ISO 04040. DIN EN ISO 04040 Personenkraftwagen - Anordnung der Handbedienteile, Anzeige- und Kontrollgeräte (1986).

DIN EN ISO 9241-210. DIN EN ISO 9241-210 Ergonomie der Mensch-System-Interaktion – Teil 210: Prozess zur Gestaltung gebrauchstauglicher interaktiver Systeme (2010).

DIN EN ISO 9241-920:2009. DIN EN ISO 09241-920 Ergonomics of human-system interaction — Part 920: Guidance on tactile and haptic interactions (2009).

DIN EN ISO 15005:2012. DIN EN ISO 15005:2012 Road vehicles - Ergonomic aspects of transportation information and control systems - Dialogue management principles and compliance procedures (2002).

DIN EN ISO 15008:2003. DIN EN ISO 15008 - Anforderungen und Bewertungsmethoden der Visuellen Informationsdarstellung im Fahrzeug (2003).

DIN EN ISO 16352. DIN EN ISO 16352 Road vehicles - Ergonomic aspects of in-vehicle information presentation for transport information and control systems - Warning systems, Pub. L. No. 16352 (2005).

DIN EN ISO 17287. DIN EN ISO 17287 Verfahren zur Bewertung der Gebrauchstauglichkeit beim Führen eines Kraftfahrzeugs (2003).

Dinse, H.R. (2008). Haptic Banknote Design. In Martin Grunwald (Ed.), *Human Haptic Perception* (537–547). Berlin: Birkhäuser Basel.

Diwischek, L., Eßers, S., & Lisseman, J. (2014). Evaluation of and design implications for force-sensitive touch input devices. In N. Stanton, G. D. Bucchianico, A. Vallicelli, & S. Landry (Eds.), *Advances in Human Aspects of Transportation* (Vol. Part 1, pp. 327–336). Krakow, Poland: AHFE Conference.

Diwischek, L., & Lisseman, J. (2015). Tactile feedback for virtual automotive steering wheel switches. In *Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 31–38). ACM.

Doerrer, C., & Werthschuetzky, R. (2002). Simulating Push-Buttons Using a Haptic Display: Requirements on Force Resolution and Force-Displacement Curve. Presented at the EuroHaptics, Edinburgh.

Dunn, W., Cahill, A., Dukes, M., & Ginsberg, A. (1986). The policy grid: A cognitive methodology for assessing change dynamics. *Policy Analysis: Perspectives, Concepts, and Methods*.

Easterby-Smith, M. (1980). The design, analysis and interpretation of repertory grids. *International Journal of Man-Machine Studies*, *13*(1), 3–24.

Edwards, M., & Cunefare, K. A. (2015). *Jury Test Summary of Results*. Internal Report.

Elias, N. (1969). *Über den Prozeß der Zivilisation* (2. ed.). Berlin: Suhrkamp.

EN 61672-1:2003. Schallmessgeräte. Elektroakustik - Schallpegelmesser - Teil 1: Anforderungen (IEC 61672-1:2002); Deutsche Fassung EN 61672-1:2003 (2003).

Faeth, A. J. (2012). *Expressive feedback from virtual buttons* (Dissertation). Iowa State University.

Fastl, H. (1997). The Psychoacoustics of Sound-Quality Evaluation. *Acta Acustica*, *83*(5), 754–763.

Fastl, H. (2006a). Advanced procedures for psychoacoustic noise evaluation. (p. 10). In: Proceedings of Euronoise 2006, Tampere, Finland.

Fastl, H. (2006b). Psychoacoustic basis of sound quality evaluation and sound engineering. In *The Thirteenth International Congress on Sound and Vibration* (Vol. 37).

Fastl, H., & Zwicker, E. (2007). *Psychoacoustics: Facts and Models*. Berlin Heidelberg: Springer.

Field, A., Miles, J., & Field, Z. (2012). *Discovering Statistics Using R*. SAGE Publications. Retrieved from https://books.google.de/books?id=Q9GCAgAAQBAJ

Flanagan, J. R., & Tresilian, J. (1994). Grip-load force coupling: a general control strategy for transporting objects. *Journal of Experimental Psychology. Human Perception and Performance*, *20*(5), 944–957.

Fransella, F., Bell, R., & Bannister, D. (2004). *A Manual for Repertory Grid Technique*. Chichester: Wiley. Retrieved from http://public.eblib.com/choice/publicfullrecord.aspx?p=219766

Fricke, N. (2009). *Gestaltung zeit- und sicherheitskritischer Warnungen im Fahrzeug*. Dissertation. Technische Universität Berlin.

Fromm, M. (1995). *Repertory Grid Methodik: Ein Lehrbuch.* Weinheim: Deutscher Studien Verlag.

Fucci, D., Small, L. H., & Petrosino, L. (1982). Intensity difference limens for lingual vibrotactile stimuli. *Journal of Magnetic Resonance*, *1*, 54–56.

Gallace, A., Tan, H. Z., & Spence, C. (2006). The failure to detect tactile change: A tactile analogue of visual change blindness. *Psychonomic Bulletin & Review*, *13*(2), 300–303. Retrieved from http://link.springer.com/article/10.3758/BF03193847

Garrett, R. (2012). *Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehicle Electronic Devices* (No. NHTSA-2010-0053) (p. 177). Washington, DC, USA: National Highway Traffic Safety Administration (NHTSA).

Geldard, A. (1957). Adventures in tactile literacy. *Journal of American Psychology*, *12*, 115–124.

Gescheider, G. A., Bolanowski, S. J., Verrillo, R. T., Arpajian, D. J., & Ryna, T. F. (1990). Vibrotactile intensity discrimination measured by three methods. *The Journal of the Acoustical Society of America*, *87*(1), 330–338.

Gescheider, G. A., Zwislocki, J. J., & Rasmussen, A. (1996). Effects of stimulus duration on the amplitude difference limen for vibrotaction. *The Journal of the Acoustical Society of America*, *100*(4), 2312–2319.

Gescheider, G., Bolanowski, S. J., Hall, K., Hoffman, K., & Verrillo, R. (1994). The Effects of Aging on Information-Processing Channels in the Sense of Touch: I. *Absolute Sensitivity. Somatosensory & Motor Research*, *11*(4).

Gescheider, G., Bolanowski, S. J., Pope, J., & Verrillo, R. T. (2002). A four-channel analysis of the tactile sensitivity of the fingertip: frequency selectivity, spatial summation, and temporal summation. *Somatosensory & Motor Research*, *19*(2).

Gibson, J. J. (1962). Observations on active touch. *Psychological Review*, *69*(6), 477–491.

Gilson, R. D. (1969). Vibrotactile masking: Some spatial and temporal aspects. *Perception & Psychophysics*, *5*(3), 176–180. Retrieved from http://link.springer.com/article/10.3758/BF03209553

Gläser, J., & Laudel, G. (2006). *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen* (2., durchges. Aufl). Wiesbaden: VS, Verl. für Sozialwiss.

Globe, A. K., & Hollins, M. (1993). Vibrotactile adaptation enhances amplitude discrimination. *The Journal of the Acoustical Society of America*, *93*(1), 418–424.

Goldstein, E. B. (2008). *Wahrnehmungspsychologie*. Berlin: Spektrum Akademischer Verlag.

Goodwin, A.W. & Wheat, H. E. (2008). Physiological mechanisms of the receptor system. In Grunwald, M. (Ed.), *Human Haptic Perception* (pp.93-102). Berlin: Birkhäuser Basel.

Grane, C. (2012). *HAPTIC PERCEPTION, ATTENTION, AND EFFECTS ON PERFORMANCE IN A DRIVING CONTEXT USE OF HAPTIC ROTARY DEVICE IN A VISUAL MENU SELECTION TASK*. LULEÅ UNIVERSITY OF TECHNOLOGY, Lulea.

Grane, C., & Bengtsson, P. (2012). Haptic Addition to a Visual Menu Selection Interface Controlled by an In-Vehicle Rotary Device. *Advances in Human-Computer Interaction*, (2012), 12.

Gray, H. (1918). *Anatomy of the Human Body.* Lea and Febiger: New York. Retrieved September 12, 2016 from https://commons.wikimedia.org/wiki/File:Gray928.png

Halata, Z., & Baumann, K. (2008). Anatomy of receptors. In M. Grunwald (Ed.), *Human Haptic Perception* (pp. 85–92). Berlin: Birkhäuser Basel.

Hameed, S., Ferris, T., Jayaraman, S., & Sarter, N. B. (2006). Supporting interruption management through informative tactile and peripheral visual cues. In *Proceedings of the Human*

*Factors and Ergonomics Society 50th Annual Meeting* (pp. 376–380). Santa Monica, CA: Human Factors and Ergonomics Society.

Harbluk, Noy, I. L., & Eizenman, M. (2002). *The impact of cognitive distraction on driver visual behaviour and vehicle control.* (p. 29). Ontario K1A 0N5 Canada: Transport Canada.

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting,* Vol. 50, No. 9, pp. 904-908). Sage Publications.

Hartmann, W. M. (1997). *Signals, Sound, and Sensation*. East Lansing, Michigan: Springer-Verlag. Retrieved from http://books.google.de/books?id=3N72rIoTHiEC&printsec=frontcover&hl=de#v=onepage&q&f=false

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

Hashimoto, T., & Hatano, S. (2001). Effects of factors other than sound to the perception of sound quality. *17th ICA Rome, CD-ROM*.

Hatzfeld, C. (2013, Mai). *Experimentelle Analyse der menschlichen Kraftwahrnehmung als ingenieurtechnische Entwurfsgrundlage für haptische Systeme*. Technical University of Darmstadt.

Haverkamp, M. (2012). *Synesthetic Design: Handbook for a Multi-Sensory Approach*. Birkhäuser Basel.

Hecht, D., & Reiner, M. (2008). Sensory dominance in combinations of audio, visual and haptic stimuli. *Exp Brain Res*, *193*, 307–314. http://doi.org/DOI 10.1007/s00221-008-1626-z

He, J. (2010). *IDENTIFY MIND-WANDERING BEHIND THE WHEEL* (Masterarbeit). University of Illinois, Urbana, Illinois.

Hemmecke, J. (2012). *Repertory Grids als Methode zum Explizieren impliziten Wissens in Organisationen: Ein Beitrag zur Methodenentwicklung im Wissensmanagement*. Universität Wien.

Hoggan, E., Crossan, A., Brewster, S. A., & Kaaresoja, T. (2009). Audio or tactile feedback: which modality when? In *Proceedings of the SIGCHI Conference on Human Factors in*

*Computing Systems* (pp. 2253–2256). Retrieved from http://dl.acm.org/citation.cfm?id=1519045

Hoggan, E., Kaaresoja, T., Laitinen, P., & Brewster, S. (2008). Crossmodal congruence: The Look, Feel and Sound of Touchscreen Widgets. In *ICMI '08 Proceedings of the 22nd British HCI Group Annual Conference on People and Computers* (pp. 157–164). New York, N. Y.: ACM.

Honda. (2002, February 8). 2003 Acura 3.2 TL -- Interior. Retrieved July 5, 2013, from http://www.honda.com/newsandviews/article.aspx?id=658-en

Honey, P. (1979). The repertory grid in action: How to use it to conduct an attitude survey. *Industrial and Commercial Training*, *11*(11), 452–459. http://doi.org/10.1108/eb003756

Hsiao, S., & Yau, J. (2008). Neural basis of haptic perception. In Grunwald, M. (Ed.), *Human Haptic Perception* (pp. 103–112). Berlin: Birkhäuser Basel.

Human Performance Group at NASA's Ames Research Center. (1986). NASA_TLX_manual.

Hunter, M. G. (1997). The use of RepGrids to gather interview data about information systems analysts. *Information Systems Journal*, *7*(1), 67–81.

Huotari, V. (2012, December). *Tactile actuator technologies*. Presented at the Seminar Haptic Interaction in mobile environments. Retrieved from http://www.uta.fi/sis/tie/him/schedule/Vesa_Huotari_presentation.pdf

ISO 14915-3:2002. DIN EN ISO 14915-3:2002 Software ergonomics for multimedia user interfaces - Part 3: Media selection and combination (2003).

ISO/TC 22, Road vehicles, Subcommittee SC 13, Ergonomics. DIN EN ISO 26022 Road vehicles — Ergonomic aspects of transport information and control systems — Simulated lane change test to assess in-vehicle secondary task demand (2009).

Jankowicz, D. (2003). *The Easy Guide to Repertory Grids*. New York: Wiley.

Jeffries, R., & Desuvire, H. (1992). Usability Testing vs. Heuristic Evaluation: Was there a contest? *SIGCHI Bulletin*, *24*(4), 39–41. Retrieved from http://csi.ufs.ac.za/resres/files/p39-jeffries.pdf

Jones, L. A. (2000). Kinesthetic Sensing. In *Human and Machine Haptics*. MIT Press.

Jones, L. A., & Lederman, S. J. (2006). *Human Hand Function*. Oxford: University Press.

Jones, L. A., & Sarter, N. B. (2008). Tactile Displays: Guidance for Their Design and Application. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *50*(1), 90–111. http://doi.org/10.1518/001872008X250638

Joos, M., Rötting, M., & Velichkovsky, B. (2002). Bewegungen des menschlichen Auges: Fakten, Methoden und innovative Anwendungen. In G. Rickheit, T. Herrmann, & W. Deutsch (Eds.), *Psycholinguistik/ Psycholinguistics. Ein internationales Handbuch/ An International Handbook* (pp. 142–168). Berlin: de Gruyter.

Jürgensohn, T., & Timpe, K.-P. (2001). *Kraftfahrzeugführung*. Berlin: Springer-Verlag.

Kaas, A. L., Stoeckel, C. & Goebel, R. (2008). The neural bases of haptic working memory. In Grunwald, M. (Ed.), *Human Haptic Perception* (113-130). Berlin: Birkhäuser Basel.

Kamps, J., Marx, M., Mokken, R. J., & de Rijke, M. (2004). Using WordNet to measure semantic orientations of adjectives. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation* (pp. 1115–1118).

Kelly, G. (1955). *The psychology of Personal constructs* (Vol. 2).

Kemp, B. J., & Bryan, J. (1973). Reaction time of young and elderly subjects in relation to perceptual deprivation and signal-on versus signal-off condition. *Developmental Psychology*, *8*, 268–272.

Kirschbaum, C. (2008). *Biopsychologie von A bis Z*. Berlin: Springer Verlag.

Klopp, E. (2015, March 7). Explorative Faktorenanalyse. Retrieved from http://www.eric-klopp.de/texte/methoden/14-explorative-faktorenanalyse

Kosinski, J. R. (2012, September). A Literature Review on Reaction Time. Retrieved July 8, 2013, from http://biae.clemson.edu/bpc/bp/lab/110/reaction.htm#Mean%20Times

Koskinen, E., Kaaresoja, T., & Laitinen, P. (2008). Feel-good touch: Finding the Most Pleasant Tactile Feedback for a Mobile Touch Screen Button. In *ICMI ’08 Proceedings of the 10th international conference on Multimodal interfaces* (pp. 297–304). New York, NY: ACM.

Krüger, H.-P., Küçükay, F., Neukum, A., Frömmig, L., Paulig, J., & Henze, R. (2009). *Untersuchung zur Wahrnehmung von Lenkmomenten bei PKW* (Forschungsvereinigung Automobiltechnik e.V. (FAT) No. 222) (p. 130). VDA.

Kuwano, S., & Namba, S. (2001). Dimension of sound quality and their measurement. In *Proceedings of the 17th International Congress on Acoustics (ICA)*.

Lancet, Y. (2013, July). What Are The Differences Between Capacitive & Resistive Touchscreens? Retrieved July 5, 2013, from http://www.makeuseof.com/tag/differences-capacitive-resistive-touchscreens-si/

Larsson, P., Opperud, A., Fredriksson, K., & Västfjäll, D. (2009). Emotional and behavioural response to auditory icons and earcons in driver-vehicle interfaces. In *21st International Technical Conference On the Enhanced Safety of Vehicles*.

Last, M. (2013). *Psychophysically based vibrotaktile and acoustic feedback of virtual buttons on a multi functional steering wheel – a driving simulation study* (Masterarbeit, internal). TU Berlin.

Latorella, K. A. (1999). *Investigating Interruptions: Implications for Flightdeck Performance*. Langley Research Center, Hampton, Virginia, USA.

Lécuyer, A. (2007). *Pseudo-Haptic Feedback*. Presented at the Tutorial 3 - Integration of Haptics in VE.

Lederman, S. J., & Jones, L. A. (2011). Tactile and Haptic Illusions. *IEEE Transactions on Haptics*, *4*(4), 273–295.

Lederman, S. J., & Klatzky, R. (1998). The hand as a perceptual system. In K. Connolly (Ed.), *The psychobiology of the hand* (pp. 16–35). London: MacKeith Press.

Lederman, S. J., & Klatzky, R. L. (2009). Haptic perception: A tutorial. *Attention, Perception, & Psychophysics*, *71*(7), 1439–1459.

Lee, G. C. H., Younghak, Y., & Jones, S. (1997). Investigation of driving performance, vection, postural sway, and simulator Sickness in a fixed-based driving simulator. *Computers and Industrial Engineering*, *33*(3), 533–536.

Loomis, J. M., & Lederman, S. J. (1986). TACTUAL PERCEPTION. In Kenneth R. Boff & Lloyd Kaufman (Eds.), *Handbook of Perception and Human Performance: Sensory processes and perception* (Vol. 1, p. 1568). Wiley.

Ludden, G. D. S., Schifferstein, H. N. J., & Miquelon, P. (2012). Beyond Surprise: A Longitudinal Study on the Experience of Visual-Tactual Incongruities in Products. *International Journal of Design*, *6*(1), 1–10.

Madhero88 (n.d.). Skin Layers. Retrieved September 12, 2016 from https://en.wikipedia.org/wiki/File:Skin_layers.png

Mahns, D. A., Perkins, N. M., Sahai, V., Robinson, L., & Rowe, M. J. (2006). Vibrotactile frequency discrimination in human hairy skin. *Journal of Neurophysiology*, *95*, 1442–1450.

Mangiafico, S. S. (2015). Two-way Anova with Robust Estimation. Retrieved from http://rcompanion.org/rcompanion/d_08a.html

Marshall, W. H., Talbot, S. A., & Ades, H. W. (1943). Cortical response of the anaesthesized cat to gross photic and electrical afferent stimulation. *Journal of Nerophysiology*, *6*, 1–15.

Mayring, P. (2000). Qualitative Inhaltsanalyse. *Philipp (2000, Juni). Qualitative Inhaltsanalyse [28 Abs?tze]. Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [On-Line Journal]*, *9*(31). Retrieved from http://nbn-resolving.de/urn:nbn:de:0114-fqs0803314.

McCracken, J. H., & Aldrich, T. B. (1984). *Analyses of selected LHX mission functions: Implications for operator workload and system automation goals* (No. ASI-479-024-84B). Anacapa Sciences Inc. Retrieved from http://www.google.de/url?q=http://www.dtic.mil/cgi-bin/GetTRDoc%3FAD%3DADA232330&sa=U&ei=7_YzU9-3M-HnygP2uYCYDw&ved=0CCcQFjAB&usg=AFQjCNG5kCfUCpGNBf2qYiRd9yLw0dEwIQ

Metz, B., & Krüger, H.-P. (2011). Distribution of visual attention during distraction: Influence of demands of the driving task and of the secondary task. In *The Second International Conference on Driver Distraction and Inattention* (p. 12). Gothenburg, Sweden.

Mike.lifeguard  (2005). Anatomy of Human Ear with Cochlear Frequency Mapping. Retrieved September 12, 2016 from https://commons.wikimedia.org/wiki/File:10.1371_journal.pbio.0030137.g001-L-A.jpg

Milham, M.P., Erickson, K. I., Banich, M. T., Kramer, A. F., Webb, A., Wszalek, T. & Cohen, N. J. (2002). Attentional Control in the Aging Brain: Insights from an fMRI Study of the Stroop Task. *Brain and Cognition*, *49*, 277–296.

Morley, J., & Rowe, M. J. (1990). Perceived pitch of vibrotactile stimuli: effects of vibration amplitude, and implications for vibration frequency coding. *Journal of Phyiology*, *431*, 403–416.

Müller, G., & Möser, M. (2003). *Taschenbuch Der Technischen Akustik*. Berlin: Springer.

Müller, M., & Giabbiconi, C.-M. (2008). Attention in the sense of touch. In M. Grunwald (Ed.), *Human Haptic Perception* (pp. 199–206). Berlin: Birkhäuser Basel.

NA 001-01-03 GA Gemeinschaftsarbeitsausschuss NALS/DKE: Schallmessgeräte. Elektroakustik - Schallpegelmesser - Teil 1: Anforderungen (IEC 61672-1:2002); Deutsche Fassung EN 61672-1:2003 (2003). Retrieved from http://www.beuth.de/de/norm/din-en-61672-1/65087188?SearchID=647653610

Nielsen, J. (1992). FINDING USABILITY PROBLEMS THROUGH HEURISTIC EVALUATION. In *Proceedings of CHI '92* (pp. 373–380). ACM.

Nielsen, J. (2010). *Usability engineering* (Nachdr.). Amsterdam: Kaufmann.

Nielsen, J., & Molich, R. (1990). HEURISTIC EVALUATION OF USER INTERFACES. In *CHI 90 Proceedings* (pp. 249–256). ACM.

Nikolic, M. I., & Sarter, N. B. (2001). Peripheral Visual Feedback: A Powerful Means of Supporting Effective Attention Allocation in Event-Driven, Data-Rich Environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *43*(1), 30–38. http://doi.org/10.1518/001872001775992525

Norman, D. (2013). Foreword. In D. Saffer, *Microinteractions*. Sebastopol, CA: O'Reilly Media.

Norman, D., Miller, J., & Henderson, A. (1995). What you see, some of what's in the future, and how we go about doing it: HI at Apple Computer (pp. 7–11).

Okamoto, S., Nagano, H., & Yamada, Y. (2013). Psychophysical Dimensions of Tactile Perception of Textures. *JOURNAL OF IEEE TRANSACTIONS ON HAPTICS*, *6*(1), 81 – 93.

Oviatt, S. (2003). The human-computer interaction handbook. In J. A. Jacko & A. Sears (Eds.), (pp. 286–304). Hillsdale, NJ, USA: L. Erlbaum Associates Inc. Retrieved from http://dl.acm.org/citation.cfm?id=772072.772093

Park, M.-S., Fukuda, T., Kim, T.-G., & Maeda, S. (2013). Health Risk Evaluation of Whole-body Vibration by ISO 2631-5 and ISO 2631-1 for Operators of Agricultural Tractors and Recreational Vehicles. *Industrial Health*, *51*, 364–370. Retrieved from https://www.jniosh.go.jp/en/indu_hel/pdf/IH_51_3_364.pdf

Pfeiffer, M., Kern, D., Schöning, J., Döring, T., Krüger, A., & Schmidt, A. (2010). A Multi-Touch Enabled Steering Wheel – Exploring the Design Speech. In *CHI '10 Extended abstracts on human factors in computing systems (S* (pp. 3355–3360). New York: ACM New York.

Philipp, H. (2013, April). Please Touch! Explore The Evolving World Of Touchscreen Technology. Retrieved July 25, 2013, from http://electronicdesign.com/components/please-touch-explore-evolving-world-touchscreen-technology

Pitts, M. J., Burnett, G. E., Williams, M. A., & Wellings, T. (2010). Does haptic feedback change the way we view touchscreens in cars? Presented at the International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction.

Pitts, M. J., Skrypchuk, L., Wellings, T., Attridge, A., & Williams, M. A. (2012). Evaluating User Response to In-Car Haptic Feedback Touchscreens Using the Lane Change Test. *Advances in Human-Computer Interaction*, *2012*, 1–13. http://doi.org/10.1155/2012/598739

Posner, M. I., & Nissen, M. J. (1976). Visual Dominance: An Information-Processing Account of Its Origins and Significance. *Psychological Review*, *83*(2), 157–171.

Provost, G., & Robert, J.-M. (2013). The Dimensions of Positive and Negative User Experiences with Interactive Products. In A. Marcus (Ed.), *Design, User Experience, and Usability. Design Philosophy, Methods, and Tools* (Vol. 8012). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from http://link.springer.com/10.1007/978-3-642-39229-0

Pujol, R. (2013). Inner Hair Cells (IHCs): overview. Retrieved July 24, 2013, from http://www.cochlea.eu/en/hair-cells/inner-hair-cells-ihcs-stucture/physiologie

Ramsey, M. (2013, June 16). Ford to Add Back Dashboard Buttons After Complaints. *June 2013). Wall Street Journal. Abgerufen Am*. Retrieved from http://online.wsj.com/article/SB10001424127887323566804578549351972660468.html

Reason, J. (1990). *Human Error*. New York, NY: Cambridge University Press.

Raichel, D. R. (2006). *THE SCIENCE AND APPLICATIONS OF ACOUSTICS*. New York: Springer.

Revelle, W. (2011). *An introduction to psychometric theory with applications in R*. Springer.

Richter, H., Ecker, R., Deisler, C., & Butz, A. (2010). HapTouch and the 2+1 State Model: Potentials of Haptic Feedback on Touch Based In-Vehicle Information Systems (p. 8). Presented at the AutomotiveUI, Pittsburgh, Pennsylvania: ACM New York.

Riemann, R. (1983). Die Erfassung individueller Einstellungen mit Hilfe der Gridtechnik. *Zeitschrift Für Sozialpsychologie*, *14*.

Rinker, M. A., Craig, J. C., & Bernstein, L. E. (1998). Amplitude and period discrimination of haptic stimuli. *The Journal of the Acoustical Society of America*, *104*(1), 453–463.

Robinson, E. S. (1934). Work of the integrated organism. In C. Murchison (Ed.), *Handbook of General Experimental Psychology*. Worcester, MA: Clark University Press.

Rohrmann, B. (1978). Antwortskalen in der empirischen Sozialforschung. *Zeitschrift für Sozialpsychologie*, *9*, 222–245.

Rosenberg, J. (2003/2004): Studien zur sinnlichen Bewertung von Tastern, Unveröffentlichtes Manuskript, Forschungsbericht der DaimlerChrysler AG, Berlin

Rosenberger, M. (2015). *Vademecum Repgrid: Leitfaden zum Professionellen Einsatz der Repertory Grid Technik - Band 1 Legitimation, Theorie, Methologie und Methodik*. Books on Demand. Retrieved from https://books.google.de/books?id=Fz8QBgAAQBAJ

Rosenthal, R. (1993). Cumulating Evidence. In G. E. Keren & C. E. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Methodological issues.* Lawrence Erlbaum Associates, Inc.

Rothenberg, M., Verrillo, R. T., Zahorian, S. A., Brachman, M. L., & Bolanowski, S. J. (1977). Vibrotactile Frequency for encoding a speech parameter. *Journal of the Acoustical Society of America*, *62*(4), 1003–1008.

Samborac (2008). Lenkrad des Audi R8. Retrieved September 12, 2016 from https://commons. wikimedia.org/wiki/File:Audi_R8_Lenkrad_20080225.jpg

Sarodnick, F., & Brau, H. (2011). *Methoden der Usability Evaluation: wissenschaftliche Grundlagen und praktische Anwendung* (2., überarb. und aktualisierte Aufl). Bern: Huber.

Sarter, N. B. (2006). Multimodal information presentation: Design guidance and research challenges. *International Journal of Industrial Ergonomics*, *36*(5), 439 – 445. http://doi.org/http://dx.doi.org/10.1016/j.ergon.2006.01.007

Schandry, R. (2006). *Biologische Psychologie: Ein Lehrbuch*. Bad Langensalza: Beltz.

Scheer, J. W., & Catina, A. (1993). *Einführung in die Repertory-Grid-Technik.* Huber.

Schmidt, G. (2005, August 4). Techniken der Erhebung. Retrieved from http://www.org-portal.org/index.php?id=12&tx_ttnews%5Bpointer%5D=9&tx_ttnews%5Btt_news%5D=360&tx_ttnews%5BbackPid%5D=4&cHash=dda6c0753cda2a8931419b1101132bab

Schmitt, M. & Altstötter-Gleich, C. (2010). Die Repertory Grid Technik. In *Differentielle Psychologie und Persönlichkeitspsychologie kompakt*. Weinheim: Beltz.

Schulte-Fortkamp, B., Quehl, B., Mellert, V., & Remmers, H. (2001). Evaluation of reaction to noise and vibration–a survey of comfort in airplanes. In *Proceedings of 17th International Congress on Acoustics*.

Sears, A., Plaisant, C., & Shneiderman, B. (1992). Expanding the scope of Touchscreen application: High Precision, Dragging Icons, and Refined Feedback. In H. R. Hartson & D. Hix (Eds.), *Advances in Human-Computer Interaction* (Vol. 3, p. 288). Intellect Books.

Sherrick, C., & Craig, J. (1982). The Psychophysics of Touch. In W. Schiff & E. Foulke (Eds.), *Tactual Perception - A Sourcebook*. New York: Cambridge University Press.

Simons, D. J. (2000). Current approaches to change blindness. *Visual Cognition*, *7*(1-3), 1–15. Retrieved from http://www.tandfonline.com/doi/abs/10.1080/135062800394658

Simons, D. J., & Levin, D. T. (1997). Change Blindness. *Trends in Cognitive Sciences*, *1*(7), 261–267.

Snell, L. D. (2008). *Force and moment analysis of stacked counter rotating eccentric mass tree shaker energy-wheel system* (Masterarbeit). Iowa State University, Iowa, USA.

Spence, C., & Ho, C. (2008). Multisensory interface design for drivers: Past, present and future. *Ergonomics 51(1)*, *51*(11), 65–70.

Spence, C., Parise, C., & Chen, Y.-C. (2012). Chapter 27 The Colavita Visual Dominance Effect. In M. Murray & M. Wallace (Eds.), *The Neural Bases of Multisensory Processes*. Boca Raton: CRC Press.

Stewart, V., Stewart, A., & Fonda, N. (1981). *Business applications of repertory grid*. McGraw-Hill Companies.

Suied, C., Susini, P., Misdariis, N., Langlois, S., Smith, B. K., & McAdams, S. (2005). Toward a sound design methodology: Application to electronic automotive sound.

Sungjae Hwang, & Jung-hee Ryu. (2010). The Haptic Steering Wheel: Vibro-tactile based Navigation for the Driving Environment. In *2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)* (pp. 660–665). Mannheim, Germany.

Susini, P., Lemaitre, G., & McAdams, S. (2011). Psychological measurement for sound description and evaluation. In B. Berglund, G. B. Rossi, J. T. Townsend, & L. R. Pendrill (Eds.), *Measurement with Persons: Theory, Methods, and Implementation Areas* (pp. 227–253). New York: Psychology Press.

Tan, F. B., & Hunter, G. M. (2002). The repertory grid technique: A method for the study of cognition in information systems. *MIS Quarterly*.

Tan, H. Z., Pang, X. D., & Durlach, N. (1992). Manual resolution of length, force, and compliance. *Advances in Robotics*, *42*.

Tan, H. Z., Srinivasan, M. A., Eberman, B., & Cheng, B. (1994). Human factors for the design of force-reflecting haptic interfaces. *Dynamic Systems and Control*, *55*(1), 353–359.

Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review*, *34*, 278–286.

Vagias, V. M. (2006). Likert-Type Scale Response Anchors. Clemson University.

Verrillo, R. T. (1963). Effect of Contactor Area on the Vibrotactile Threshold. *The Journal of the Acoustical Society of America*, *35*, 1962–1966.

Verrillo, R. T., & Chamberlain, S. C. (1972). The effect of neural density and contactor surround on vibrotactile sensation magnitude. *Perception & Psychophysics*, *11*(B), 117–120.

Verrillo, R. T., Gescheider, G. A., Calman, B. G., & Van Doren, C. L. (1983). Vibrotactile masking: Effects of one and two-site stimulation. *Perception & Psychophysics*, *33*(4), 379–387. Retrieved from http://link.springer.com/article/10.3758/BF03205886

Vogelpohl, T. (2014). *Material Design for Surfaces of Virtual Buttons with Haptic Feedback on the Steering Wheel* (Master thesis, internal). TU Berlin.

Vollrath, M., & Totzke, I. (2003). Möglichkeiten der Nutzung unterschiedlicher Ressourcen für die Fahrer-Fahrzeug-Interaktion: Using Multiple Resources for the Drivier-Vehicle-Interaction. Düsseldorf: VDI-Verlag.

von Békésy, G. (1959). Synchronism of Neural Discharges and Their Demultiplication in Pitch Perception on the Skin and in Hearing. *The Journal of the Acoustical Society of America*, *31*(3). http://doi.org/http://dx.doi.org/10.1121/1.1907722

Wagner, K., Drescher, S. & Reichel, S. (2010). *Faktorenanalyse*. Presented at the Multivariate Analysemethoden, Fakultät für Human- und Sozialwissenschaften Professur für Forschungsmethodik und Evaluation in der Psychologie.

Wellings, T., Williams, M., & Tennant, C. (2009). Understanding customers' holistic perception of switches in automotive human–machine interfaces. *Applied Ergonomics*, *41*(1), 8–17.

Werz, S. (2006, September). *Repertory Grid - Untersuchung eines Datenanalyseverfahrens* (Diplomarbeit). Universität Konstanz, Konstanz.

Wickens, C. D. (2002). Multiple resources and performance prediction Theoretical Issues in Ergonomics Science, 3(2). *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177.

Wickens, C. D. (2007). Attention to attention and its applications: A concluding view. In A. F. Kramer, D. A. Wiegmann, & A. Kirlik (Eds.), *Attention: From theory to practice*. New York: Oxford University Press.

Wickens, C. D. (2008). Multiple Resources and Mental Workload. *Human Factors*, *50*(3), 449–455.

Wickens, C., & Hollands, J. (1999). Attention, Time-Sharing, and Workload. In *C*. Wickens & J. Hollands. Engineering Psychology and Human Performance. Boston: Addison Wesley Pub Co Inc.

Wilson, E. C., Reed, C. M., & Braida, L. D. (2010). Integration of auditory and vibrotactile stimuli: Effects of frequency. *Journal of the Acoustical Society of America*, *127*(5), 3044–3059.

Wood, N., & Cowan, N. (1995). The Cocktail Party Phenomenon Revisited: How Frequent Are Attention Shifts to One's Name in an Irrelevant Auditory Channel? *Hournal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 255–260.

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2p2), 1.

Zeilinger, S. (2005). *Aktive haptische Bedienelemente zur Interaktion mit Fahrerinformationssystemen*. Universität der Bundeswehr München, München.

Zeitler, A., Ellermeier, W., & Fastl, H. (2004). Significance of meaning in sound quality evaluation. *Fortschritte Der Akustik, CFA/DAGA*, *4*, 781–782.

Zeitler, A., & Hellbrück, J. (2001). Semantic attributes of environmental sounds and their correlations with psychoacoustic magnitude. In *Proc. of the 17th International Congress on Acoustics [CDROM], Rome, Italy* (Vol. 28).

Zeller, P. (2009). Psychoakustik. In P. Zeller (Ed.), *Handbuch Fahrzeugakustik*. Wiesbaden: Vieweg+Teubner. Retrieved from http://www.springerlink.com/index/10.1007/978-3-8348-9322-2

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and McDonald's ω: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70*(1), 123–133.

## Appendix


a) Second preliminary experiment ……………………………………………………...152

a. Demographic questionnaire (German) ........................................................................ 152

b. Repertory Grid (German with item translations below) ............................................ 154

c. Driving performance results........................................................................................ 155

d. Post-hoc Nemenyi results for the skip function ......................................................... 159

e. Post-hoc Nemenyi and Tukey HSDs for good/bad rating........................................... 159

f. Principal component analysis ...................................................................................... 161

b) Third preliminary experiment ...................................................................................... 163

a. Repertory Grid Design (German with item translations below) ................................ 163

b. Driving performance results........................................................................................ 164

c. Preference data ........................................................................................................... 165

d. Principal component analysis ...................................................................................... 167

e. Cronbach’s α ............................................................................................................... 169

c) Main experiment ......................................................................................................... 170

a. Results of the pretest ................................................................................................... 170

b. Grid design for main experiment (German with item translations below) ................. 171

c. Switch Perception Questionnaire (adapted after Wellings et al., 2009)...................... 172

d. Twofactorial ANOVA of the distance to ideal............................................................ 172

e. Posthoc tests for the main effects ................................................................................ 173

f. Factor determination .................................................................................................... 175

g. Principal component analyses with three, four and five factors ................................ 176

f. Cronbach’s α of the main experiment ......................................................................... 178

h. McDonald‘s omega ..................................................................................................... 179

i. Promax rotation with five factors ................................................................................ 180

j. Interitem Correlations .................................................................................................. 181

d. Final Instrument........................................................................................................... 181

a. German Version ........................................................................................................... 182

b. English Version ........................................................................................................... 183

## a) *Second preliminary experiment*

### a. Demographic questionnaire (German)

TAKATA

FAHRSIMULATOR-STUDIE

| Vom Versuchsleiter auszufüllen | Datum: | |
|---|---|---|
| | Uhrzeit: | |
| VP-Nummer: | Konfiguration: O Skip First | O Mode first |
| | Beginn | Ende |
| Fingertemperatur: | °C | °C |
| Raumtemperatur: | °C | °C |

**Nun haben wir noch ein paar Fragen an Sie.** Ihre Daten werden selbstverständlich anonym und vertrauensvoll behandelt.

1. Alter: ______
2. Geschlecht: O männlich O weiblich
3. Händigkeit: O linkshändig O rechtshändig

**Wir möchten Sie nun bitten, die für Sie zutreffenden Antworten anzukreuzen.**

4. Wie viele Kilometer legen Sie etwa in einem Jahr mit einem Fahrzeug zurück?
   - O <5 000
   - O 5 000 – 10 000
   - O >10 000 – 20 000
   - O >20 000 – 30 000
   - O > 30 000
5. Haben Sie schon einmal an einem Fahrsicherheitstraining teilgenommen?

   O Ja O Nein

**Information:** *Ein Multifunktionslenkrad ist ein Lenkrad, das zum Beispiel einen Knopf zum Einstellen der Radiolautstärke hat.*

6. Fahren Sie regelmäßig ein Fahrzeug, das über ein Multifunktionslenkrad verfügt?

   O Ja O Nein

TAKATA

VP_NR:_____

7. Wenn ja, benutzen Sie die Knöpfe auf dem Lenkrad?

   O Ja O Nein

8. Wie viele Stunden haben Sie in der vergangenen Nacht geschlafen?

   _______ Stunden

9. Sind Sie Raucher?

   O Ja O Nein

10. Haben Sie in den letzten 24 Stunden Medikamente eingenommen, die Ihre Fahrleistung beeinträchtigen?

    O Ja O Nein

    Wenn ja, welche?

    ______________________________________________

    ______________________________________________

    ______________________________________________

11. Haben Sie bereits Fahrsimulator-Erfahrungen?

    O Ja O Nein

    Wenn ja, welche?

    - O PC mit Lenkrad
    - O Fahrsimulator
    - O PC mit Joystick
    - O Sonstige: ________________________

      ________________________

Falls Sie schon einmal in einem Fahrsimulator gefahren sind, was für ein Typ Simulator ist dies gewesen?

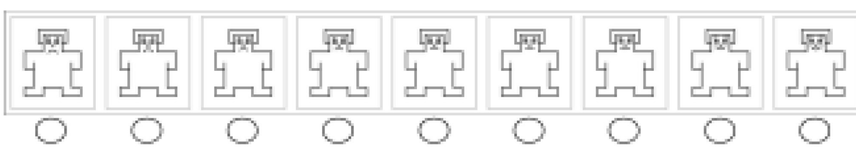
- ○ Echtfahrzeug
- ○ Fahrstand
- ○ Sonstige: ________________________________________________
________________________________________________________________

**Aktuelles Befinden**

Bitte Setzen Sie ein Kreuz entsprechend Ihrer aktuellen Stimmung.

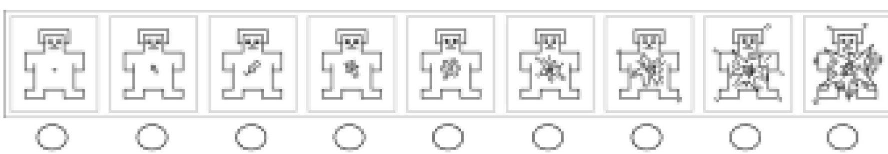
12. Wie fühlen Sie sich gerade? Bitte kreuzen Sie in jeder Reihe das „Männchen" an, welches Ihrem aktuellen Empfinden am nächsten kommt.

sehr traurig/ deprimiert



sehr glücklich/ fröhlich

sehr ruhig



aktiv/ voller Taten drang

## b. Repertory Grid (German with item translations below)

| Was haben Knopf 1 und 2 gemeinsam im Vergleich zu Knopf 3? | | Worin unterscheidet sich Knopf 3? |
|---|---|---|
| | vs. | |
| | vs. | |
| | vs. | |
| | **Grid-Generierung** | |
| 3 und 7 | vs. | 4 |
| 1 und 5 | vs. | 2 |
| 2 und 6 | vs. | 3 |

| Konstruktpol | Antwortskala | | | | | Gegenpol |
|---|---|---|---|---|---|---|
| | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | |
| 0 | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | 0 |
| 0 | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | 0 |
| kraftlos | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | beherrschend |
| gleichmäßig | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | rauh |
| kurz | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | lang |
| stark | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | schwach |
| angenehm | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | klickig |
| langsam | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | schnell |
| hart | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | weich |
| mechanisch | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | buzzy |
| gut | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | schlecht |

Item translations (order reversed compared to grid):

| English | German |
|---|---|
| Predominant – Faint | Beherrschend – Kraftlos |
| Rough – Smooth | Rauh – Gleichmäßig |
| Long – Short | Lang – Kurz |
| Strong – Weak | Stark – Schwach |
| Clicky – Pleasant | Klickig – Angenehm |
| Fast – Slow | Schnell – Langsam |
| Hard – Soft | Hart – Weich |
| Mechanical – Buzzy | Mechanisch – Buzzy |
| Bad – Good | Schlecht – Gut |

## c. Driving performance results

| TWO-FACTORIAL ANOVA for mean velocity | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Effect | | DFn | DFd | SSn | SSd | F | p | p<.05 |
| 1 | (Intercept) | 1 | 38 | 7281358.44 | 5710.936 | 4.84E+04 | 1.25E-60 | 0.98722039 |
| 2 | Drive | 4 | 152 | 3126.7468 | 19358.733 | 6.14E+00 | 1.32E-04 | 0.03210732 |
| 3 | Switch setting | 7 | 266 | 585.5835 | 12389.745 | 1.80E+00 | 8.82E-02 | 0.00617424 |
| 4 | Drive: Switch setting | 28 | 1064 | 1103.8572 | 56798.072 | 7.39E-01 | 8.36E-01 | 0.01157552 |
| Mauchly's Test for Sphericity | | | | | | | | |
| Effect | W | p | p<.05 | | | | | |
| 2 | Drive | 1.13E-01 | 2.47E-13 | * | | | | |
| 3 | Switch setting | 4.73E-01 | 4.92E-01 | | | | | |
| 4 | Drive: Switch setting | 1.60E-14 | 7.37E-39 | * | | | | |
| Sphericity Corrections | | | | | | | | |
| Effect | GGe | p[GG] | p[GG]<.05 | HFe | p[HF] | p[HF]<.05 | | |
| 2 | Drive | 0.5098545 | 0.00316845 | * | 0.5390437 | 0.00261494 | * | |
| 3 | Switch setting | 0.8284145 | 0.1035992 | 0.9944522 | 0.08864218 | | | |
| 4 | Drive: Switch setting | 0.3993969 | 0.70375369 | 0.5781824 | 0.75681841 | | | |

General eta square for effect „Drive“= 0.032107321

Explanation: Main effect “Drive” describes the five different types of courses, F = Familiarization, MG = Mode Repertory Grid, MP = Mode Performance, SG = Skip Repertory Grid, and SP = Skip Performance.
The main effect “Switch setting” describes the different switch factor variations.

| Bonferroni corrected pairwise t-test for mean velocity in main effect “Drive” | | | | |
|---|---|---|---|---|
| | F | MG | MP | SG |
| MG | 0.03181 | - | - | - |
| MP | 0.00372 | 5.00E-08 | - | - |
| SG | 0.0516 | 1 | 6.90E-08 | - |
| SP | 1 | 0.0176 | 0.00032 | 0.0217 |

| MANOVA for standard deviation of mean velocity | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Effect | | DFn | DFd | SSn | SSd | F | p | p<.05 |
| 1 | (Intercept) | 1 | 38 | 18548.1353 | 6.94E+02 | 1.02E+03 | 5.03E-29 | 0.314918 |
| 2 | Drive | 4.00E+00 | 152 | 10590.0612 | 2.41E+03 | 1.67E+02 | 1.41E-54 | 0.20789191 |
| 3 | Switch setting | 7 | 266 | 438.3793 | 6.81E+03 | 2.45E+00 | 1.91E-02 | 0.01074761 |
| 4 | Drive: Switch setting | 28 | 1064 | 1210.7485 | 3.04E+04 | 1.51E+00 | 4.33E-02 | 0.0291319 |
| Mauchly's Test for Sphericity | | | | | | | | |
| Effect | W | p | p<.05 | | | | | |
| 2 | Drive | 4.80E-02 | 1.42E-19 | * | | | | |
| 3 | Switch setting | 1.56E-01 | 4.70E-05 | * | | | | |
| 4 | Drive: Switch setting | 2.85E-34 | 9.71E-242 | * | | | | |
| Sphericity Corrections | | | | | | | | |
| Effect | GGe | p[GG] | p[GG]<.05 | HFe | p[HF] | p[HF]<.05 | | |
| 2 | Drive | 0.5334108 | 2.71E-30 | * | 0.5660742 | 5.38E-32 | * | |
| 3 | Switch setting | 0.6631385 | 3.98E-02 | * | 0.7667325 | 3.17E-02 | * | |
| 4 | Drive: Switch setting | 0.2717578 | 1.57E-01 | 0.3463921 | 1.36E-01 | | | |

General eta square for effect „Drive“= 0.20789191
General eta square for effect „Switch setting“= 0.01074761

Explanation: Main effect "Drive" describes the five different types of courses, F = Familiarization, MG = Mode Repertory Grid, MP = Mode Performance, SG = Skip Repertory Grid, and SP = Skip Performance.
The main effect "Switch setting" describes the different switch factor variations.

| Bonferroni corrected pairwise t-test for standard deviation of velocity for main effect "Drive" | | | | |
|---|---|---|---|---|
| | F | MG | MP | SG |
| MG | 2.00E-16 | - | - | - |
| MP | 2.00E-16 | 2.00E-16 | - | - |
| SG | 7.00E-15 | 1 | 2.00E-16 | - |
| SP | 2.00E-16 | 2.00E-16 | 4.40E-13 | 2.00E-16 |

| Bonferroni corrected pairwise t-test for standard deviation of velocity for main effect "Switch Setting" | | | | | | | |
|---|---|---|---|---|---|---|---|
| | $\bar{Y}_{11}$ | $\bar{Y}_{21}$ | $\bar{Y}_{31}$ | $\bar{Y}_{41}$ | $\bar{Y}_{12}$ | $\bar{Y}_{22}$ | $\bar{Y}_{32}$ |
| $\bar{Y}_{21}$ | 1 | - | - | - | - | - | - |
| $\bar{Y}_{31}$ | 1 | 1 | - | - | - | - | - |
| $\bar{Y}_{41}$ | 1 | 1 | 1 | - | - | - | - |
| $\bar{Y}_{12}$ | 1 | 1 | 1 | 1 | - | - | - |
| $\bar{Y}_{22}$ | 0.829 | 0.459 | 0.6 | 0.230 | 0.538 | - | - |
| $\bar{Y}_{32}$ | 1 | 1 | 1 | 1 | 1 | 0.035 | - |
| $\bar{Y}_{42}$ | 1 | 1 | 1 | 0.829 | 1 | 1 | 0.168 |

| TWO-FACTORIAL ANOVA (type 1) for standard deviation of lane position | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Effect | | DFn | DFd | SSn | SSd | F | p | p<.05 |
| 1 | Drive | 4 | 148 | 11.023272 | 2.61565 | 155.93107 | 5.14E-52 | 0.56960222 |
| 2 | Switch setting | 7 | 259 | 0.03746064 | 1.420767 | 0.9755601 | 4.49E-01 | 0.00447731 |
| 3 | Drive: Switch setting | 28 | 1036 | 0.10289895 | 4.292891 | 0.8868759 | 6.36E-01 | 0.01220309 |

General eta square for main effect „Drive“= 0.569602215

Due to the data, no investigation of sphericity was possible. Results are therefore not reliable.

Explanation: Main effect “Drive” describes the five different types of courses, F = Familiarization, MG = Mode Repertory Grid, MP = Mode Performance, SG = Skip Repertory Grid, and SP = Skip Performance.
The main effect “Switch setting” describes the different switch factor variations.

| Bonferroni corrected pairwise t-test for standard deviation lane position for main effect “Drive” | | | | |
|---|---|---|---|---|
| | F | MG | MP | SG |
| MG | 2.00E-16 | - | - | - |
| MP | 2.00E-16 | - | - | - |
| SG | 1 | 2.00E-16 | 2.00E-16 | - |
| SP | 2.00E-16 | 1.70E-13 | 1.70E-13 | 2.00E-16 |

| TWO-FACTORIAL ANOVA for lane keeping quality | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Effect | | DFn | DFd | SSn | SSd | F | p | p<.05 |
| 1 | (Intercept) | 1 | 37 | 10149.0113 | 323.5433 | 1.16E+03 | 1.53E-29 | 0.77440226 |
| 2 | Drive | 4 | 148 | 1521.15644 | 678.6964 | 8.29E+01 | 8.40E-37 | 0.33971433 |
| 3 | Switch setting | 7.00E+00 | 2.59E+02 | 14.03866 | 383.6474 | 1.3539266 | 2.25E-01 | 0.00472581 |
| 4 | Drive: Switch setting | 2.80E+01 | 1.04E+03 | 2.82E+01 | 1570.7081 | 0.6647656 | 9.08E-01 | 0.00945464 |
| Mauchly's Test for Sphericity | | | | | | | | |
| Effect | W | p | p<.05 | | | | | |
| 2 | Drive | 3.92E-01 | 1.26E-04 | * | | | | |
| 3 | Switch setting | 1.26E-01 | 8.20E-06 | * | | | | |
| 4 | Drive: Switch setting | 1.17E-14 | 4.38E-36 | * | | | | |
| Sphericity Corrections | | | | | | | | |
| Effect | GGe | p[GG] | p[GG]<.05 | HFe | p[HF] | p[HF]<.05 | | |
| 2 | Drive | 0.7531891 | 2.48E-28 | * | 0.8274042 | 7.02E-31 | * | |
| 3 | Switch setting | 0.5820703 | 2.52E-01 | 0.6631017 | 2.47E-01 | | | |
| 4 | Drive: Switch setting | 0.2068486 | 6.73E-01 | 0.2495756 | 7.02E-01 | | | |

General eta square for main effect „Drive“= 0.339714334

Explanation: Main effect “Drive” describes the five different types of courses, F = Familiarization, MG = Mode Repertory Grid, MP = Mode Performance, SG = Skip Repertory Grid, and SP = Skip Performance.
The main effect “Switch setting” describes the different switch factor variations.

| Bonferroni corrected pairwise t-test for lane keeping quality for main effect "Drive" | | | | |
|---|---|---|---|---|
| | F | MG | MP | SG |
| MG | 1 | - | - | - |
| MP | 2.00E-16 | 2.00E-16 | - | - |
| SG | 1.00E+00 | 1 | 2.00E-16 | - |
| SP | 2.00E-16 | 2.00E-16 | 1.8e-10 | 2.00E-16 |

| TWO-FACTORIAL ANOVA for steering wheel velocity | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Effect | | DFn | DFd | SSn | SSd | F | p | p<.05 |
| 1 | (Intercept) | 1.00E+00 | 37 | 9.574279 | 154.6419 | 2.290766 | 1.39E-01 | 0.00337777 |
| 2 | Drive | 4 | 1.48E+02 | 113.846648 | 655.7285 | 6.42E+00 | 8.56E-05 | 0.03873951 |
| 3 | Switch setting | 7.00E+00 | 259 | 1.57E+01 | 460.4825 | 1.26E+00 | 2.72E-01 | 0.00551267 |
| 4 | Drive: Switch setting | 2.80E+01 | 1.04E+03 | 81.153663 | 1554.074 | 1.932138 | 2.64E-03 | 0.02792547 |
| Mauchly's Test for Sphericity` | | | | | | | | |
| Effect | W | p | p<.05 | | | | | |
| 2 | Drive | 5.51E-03 | 1.16E-34 | * | | | | |
| 3 | Switch setting | 5.54E-02 | 3.58E-10 | * | | | | |
| 4 | Drive: Switch setting | 2.61E-30 | 9.66E-185 | * | | | | |
| Sphericity Corrections | | | | | | | | |
| Effect | GGe | p[GG] | p[GG]<.05 | HFe | p[HF] | p[HF]<.05 | | |
| 2 | Drive | 0.3636925 | 0.00698135 | * | 0.3747427 | 0.0064569 | * | |
| 3 | Switch setting | 0.4858802 | 2.91E-01 | 0.5410227 | 0.29001657 | | | |
| 4 | Drive: Switch setting | 0.1825441 | 8.92E-02 | 0.2152873 | 7.64E-02 | | | |

General eta square for main effect „Drive"= 0.038739510

Explanation: Main effect "Drive" describes the five different types of courses, F = Familiarization, MG = Mode Repertory Grid, MP = Mode Performance, SG = Skip Repertory Grid, and SP = Skip Performance.
The main effect "Switch setting" describes the different switch factor variations.

| Bonferroni corrected pairwise t-test for steering wheel velocity for main effect "Drive" | | | | |
|---|---|---|---|---|
| | E | MG | MP | SG |
| MG | 1 | - | - | - |
| MP | 1.10E-04 | 4.30E-05 | - | - |
| SG | 1.67E-01 | 0.31155 | 1.60E-05 | - |
| SP | 1.00E+00 | 1.00E+00 | 0.00017 | 1.93E-01 |

| TWO-FACTORIAL ANOVA for steering wheel angle velocity | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Effect | | DFn | DFd | SSn | SSd | F | p | p<.05 |
| 1 | (Intercept) | 1.00E+00 | 37 | 0.03774381 | 290.5006 | 0.00480729 | 9.45E-01 | 2.52E-06 |
| 2 | Drive | 4 | 1.48E+02 | 18.4186238 | 1007.0928 | 6.77E-01 | 6.09E-01 | 1.23E-03 |
| 3 | Switch setting | 7.00E+00 | 259 | 7.34E+01 | 2873.3532 | 9.46E-01 | 4.72E-01 | 4.88E-03 |
| 4 | Drive: Switch setting | 2.80E+01 | 1.04E+03 | 299.198732 | 10814.9759 | 1.02361329 | 4.32E-01 | 1.96E-02 |
| Mauchly's Test for Sphericity | | | | | | | | |
| Effect | W | p | p<.05 | | | | | |
| 2 | Drive | 2.03E-02 | 3.85E-25 | * | | | | |
| 3 | Switch setting | 1.95E-01 | 8.46E-04 | * | | | | |
| 4 | Drive: Switch setting | 2.88E-32 | 5.40E-206 | * | | | | |
| Sphericity Corrections | | | | | | | | |
| Effect | GGe | p[GG] | p[GG]<.05 | HFe | p[HF] | p[HF]<.05 | | |
| 2 | Drive | 0.4229866 | 0.488269 | 0.4410747 | 0.494054 | | | |
| 3 | Switch setting | 0.7274931 | 4.54E-01 | 0.8574467 | 0.4631044 | | | |
| 4 | Drive: Switch setting | 0.248402 | 4.15E-01 | 0.3117965 | 4.20E-01 | | | |

Explanation: Main effect "Drive" describes the five different types of courses, F = Familiarization, MG = Mode Repertory Grid, MP = Mode Performance, SG = Skip Repertory Grid, and SP = Skip Performance.
The main effect "Switch setting" describes the different switch factor variations.

## d. Post-hoc Nemenyi results for the skip function

Pairwise comparisons using Nemenyi post-hoc test with q approximation for unreplicated blocked data

| | $\bar{Y}_{11}$ | $\bar{Y}_{21}$ | $\bar{Y}_{31}$ | $\bar{Y}_{41}$ | $\bar{Y}_{12}$ | $\bar{Y}_{22}$ | $\bar{Y}_{32}$ |
|---|---|---|---|---|---|---|---|
| $\bar{Y}_{21}$ | 1 | 1 | | | | | |
| $\bar{Y}_{31}$ | 1 | 0.998 | 1 | | | | |
| $\bar{Y}_{41}$ | 0.288 | 0.097 | 0.404 | 1 | | | |
| $\bar{Y}_{12}$ | 1 | 1 | 0.998 | 0.097 | 1 | | |
| $\bar{Y}_{22}$ | 0.995 | 0.916 | 0.999 | 0.786 | 0.916 | 1 | |
| $\bar{Y}_{32}$ | 0.837 | 0.534 | 0.916 | 0.989 | 0.534 | 0.998 | 1 |
| $\bar{Y}_{42}$ | 0.916 | 0.666 | 0.965 | 0.965 | 0.666 | 1.000 | 1 |

## e. Post-hoc Nemenyi and Tukey HSDs for good/bad rating

### Mode functionality

Post-hoc Nemenyi Test for good/bad ratings for each switch setting for mode function

| | $\bar{Y}_{11}$ | $\bar{Y}_{21}$ | $\bar{Y}_{31}$ | $\bar{Y}_{41}$ | $\bar{Y}_{12}$ | $\bar{Y}_{22}$ | $\bar{Y}_{32}$ |
|---|---|---|---|---|---|---|---|
| $\bar{Y}_{21}$ | 0.955 | - | - | - | - | - | - |
| $\bar{Y}_{31}$ | 0.3131 | 9.44E-01 | - | - | - | - | - |
| $\bar{Y}_{41}$ | 0.0058 | 1.69E-01 | 0.8531 | - | - | - | - |
| $\bar{Y}_{12}$ | 0.7939 | 1.43E-01 | 3.80E-03 | 5.50E-06 | - | - | - |
| $\bar{Y}_{22}$ | 0.9898 | 1 | 8.53E-01 | 9.01E-02 | 0.2508 | - | - |
| $\bar{Y}_{32}$ | 0.9683 | 1.00E+00 | 9.25E-01 | 1.43E-01 | 1.69E-01 | 1 | - |
| $\bar{Y}_{42}$ | 0.0058 | 0.1686 | 0.8531 | 1 | 5.50E-06 | 0.0901 | 0.1434 |

| Tukey HSD post-hoc test for good/bad rating for mode function | | | | |
|---|---|---|---|---|
| | diff | lwr | upr | p adj |
| $\bar{Y}_{21}$-$\bar{Y}_{11}$ | 3.33E-01 | -0.46568245 | 1.13234911 | 0.9080746 |
| $\bar{Y}_{31}$-$\bar{Y}_{11}$ | 7.44E-01 | -5.54E-02 | 1.54E+00 | 0.0890537 |
| $\bar{Y}_{41}$-$\bar{Y}_{11}$ | 1.05E+00 | 2.52E-01 | 1.85029783 | 1.89E-03 |
| $\bar{Y}_{12}$-$\bar{Y}_{11}$ | -3.85E-01 | -1.18E+00 | 4.14E-01 | 8.23E-01 |
| $\bar{Y}_{22}$-$\bar{Y}_{11}$ | 2.05E-01 | -0.59388757 | 1.00E+00 | 9.94E-01 |
| $\bar{Y}_{32}$-$\bar{Y}_{11}$ | 3.08E-01 | -4.91E-01 | 1.11E+00 | 9.38E-01 |
| $\bar{Y}_{42}$-$\bar{Y}_{11}$ | 1.05E+00 | 0.25226627 | 1.85029783 | 0.0018943 |
| $\bar{Y}_{31}$-$\bar{Y}_{21}$ | 4.10E-01 | -0.38875937 | 1.21E+00 | 7.70E-01 |
| $\bar{Y}_{41}$-$\bar{Y}_{21}$ | 7.18E-01 | -0.08106706 | 1.52E+00 | 1.14E-01 |
| $\bar{Y}_{12}$-$\bar{Y}_{21}$ | -7.18E-01 | -1.5169645 | 8.11E-02 | 1.14E-01 |
| $\bar{Y}_{22}$-$\bar{Y}_{21}$ | -1.28E-01 | -0.92722091 | 6.71E-01 | 1.00E+00 |
| $\bar{Y}_{32}$-$\bar{Y}_{21}$ | -2.56E-02 | -0.82465681 | 0.77337475 | 1 |
| $\bar{Y}_{42}$-$\bar{Y}_{21}$ | 7.18E-01 | -0.08106706 | 1.5169645 | 0.1140881 |
| $\bar{Y}_{41}$-$\bar{Y}_{31}$ | 3.08E-01 | -0.49132347 | 1.10670809 | 9.38E-01 |
| $\bar{Y}_{12}$-$\bar{Y}_{31}$ | -1.13E+00 | -1.92722091 | -0.32918935 | 5.79E-04 |
| $\bar{Y}_{22}$-$\bar{Y}_{31}$ | -5.38E-01 | -1.33747732 | 0.26055424 | 4.45E-01 |
| $\bar{Y}_{32}$-$\bar{Y}_{31}$ | -4.36E-01 | -1.23491322 | 0.36311834 | 0.7097758 |
| $\bar{Y}_{42}$-$\bar{Y}_{31}$ | 3.08E-01 | -0.49132347 | 1.10670809 | 0.9383979 |
| $\bar{Y}_{12}$-$\bar{Y}_{41}$ | -1.44E+00 | -2.23491322 | -0.63688166 | 0.0000024 |
| $\bar{Y}_{22}$-$\bar{Y}_{41}$ | -8.46E-01 | -1.64516963 | -0.04713807 | 0.0293166 |
| $\bar{Y}_{32}$-$\bar{Y}_{41}$ | -7.44E-01 | -1.54260552 | 0.05542604 | 0.0890537 |
| $\bar{Y}_{42}$-$\bar{Y}_{41}$ | -4.22E-15 | -0.79901578 | 0.79901578 | 1 |
| $\bar{Y}_{22}$-$\bar{Y}_{12}$ | 5.90E-01 | -0.20927219 | 1.38875937 | 0.3235466 |
| $\bar{Y}_{32}$-$\bar{Y}_{12}$ | 6.92E-01 | -0.10670809 | 1.49132347 | 0.1443217 |
| $\bar{Y}_{42}$-$\bar{Y}_{12}$ | 1.44E+00 | 0.63688166 | 2.23491322 | 0.0000024 |
| $\bar{Y}_{32}$-$\bar{Y}_{22}$ | 1.03E-01 | -0.69645168 | 0.90157988 | 0.999933 |
| $\bar{Y}_{42}$-$\bar{Y}_{22}$ | 8.46E-01 | 0.04713807 | 1.64516963 | 0.0293166 |
| $\bar{Y}_{42}$-$\bar{Y}_{32}$ | 7.44E-01 | -0.05542604 | 1.54260552 | 0.0890537 |

## Skip functionality

| Post-hoc Nemenyi Test for good/bad ratings for skip function | | | | | | | |
|---|---|---|---|---|---|---|---|
| | $\bar{Y}_{11}$ | $\bar{Y}_{21}$ | $\bar{Y}_{31}$ | $\bar{Y}_{41}$ | $\bar{Y}_{12}$ | $\bar{Y}_{22}$ | $\bar{Y}_{32}$ |
| $\bar{Y}_{21}$ | 5.38E-01 | - | - | - | - | - | - |
| $\bar{Y}_{31}$ | 1.52E-01 | 9.97E-01 | - | - | - | - | - |
| $\bar{Y}_{41}$ | 8.90E-03 | 7.25E-01 | 0.9813 | - | - | - | - |
| $\bar{Y}_{12}$ | 1.00E+00 | 6.02E-01 | 1.87E-01 | 1.23E-02 | - | - | - |
| $\bar{Y}_{22}$ | 1.00E+00 | 0.6652 | 2.28E-01 | 1.69E-02 | 1 | - | - |
| $\bar{Y}_{32}$ | 9.84E-01 | 9.75E-01 | 6.96E-01 | 1.43E-01 | 9.91E-01 | 0.9958 | - |
| $\bar{Y}_{42}$ | 1.14E-01 | 0.9927 | 1 | 0.9913 | 1.43E-01 | 0.1777 | 0.6182 |

| Tukey HSD test for good/bad rating for skip function | | | | |
|---|---|---|---|---|
| | diff | lwr | upr | p adj |
| $\bar{Y}_{21}$-$\bar{Y}_{11}$ | 5.64E-01 | -0.27154321 | 1.39974834 | 0.4431778 |
| $\bar{Y}_{31}$-$\bar{Y}_{11}$ | 8.46E-01 | 1.05E-02 | 1.68E+00 | 0.0447668 |
| $\bar{Y}_{41}$-$\bar{Y}_{11}$ | 1.05E+00 | 2.16E-01 | 1.88692782 | 3.70E-03 |
| $\bar{Y}_{12}$-$\bar{Y}_{11}$ | -2.56E-02 | -8.61E-01 | 8.10E-01 | 1.00E+00 |
| $\bar{Y}_{22}$-$\bar{Y}_{11}$ | 2.05E-01 | -0.63051757 | 1.04E+00 | 9.95E-01 |
| $\bar{Y}_{32}$-$\bar{Y}_{11}$ | 2.31E-01 | -6.05E-01 | 1.07E+00 | 9.90E-01 |
| $\bar{Y}_{42}$-$\bar{Y}_{11}$ | 7.69E-01 | -0.066415 | 1.60487654 | 0.0964118 |
| $\bar{Y}_{31}$-$\bar{Y}_{21}$ | 2.82E-01 | -0.55359449 | 1.12E+00 | 9.70E-01 |
| $\bar{Y}_{41}$-$\bar{Y}_{21}$ | 4.87E-01 | -0.34846629 | 1.32E+00 | 6.35E-01 |
| $\bar{Y}_{12}$-$\bar{Y}_{21}$ | -5.90E-01 | -1.42538936 | 2.46E-01 | 3.83E-01 |
| $\bar{Y}_{22}$-$\bar{Y}_{21}$ | -3.59E-01 | -1.19462013 | 4.77E-01 | 8.94E-01 |
| $\bar{Y}_{32}$-$\bar{Y}_{21}$ | -3.33E-01 | -1.16897911 | 0.50231244 | 0.9263291 |
| $\bar{Y}_{42}$-$\bar{Y}_{21}$ | 2.05E-01 | -0.63051757 | 1.04077398 | 0.9953436 |
| $\bar{Y}_{41}$-$\bar{Y}_{31}$ | 2.05E-01 | -0.63051757 | 1.04077398 | 9.95E-01 |
| $\bar{Y}_{12}$-$\bar{Y}_{31}$ | -8.72E-01 | -1.70744065 | -0.0361491 | 3.39E-02 |
| $\bar{Y}_{22}$-$\bar{Y}_{31}$ | -6.41E-01 | -1.47667141 | 0.19462013 | 2.75E-01 |
| $\bar{Y}_{32}$-$\bar{Y}_{31}$ | -6.15E-01 | -1.45103039 | 0.22026116 | 0.3264862 |
| $\bar{Y}_{42}$-$\bar{Y}_{31}$ | -7.69E-02 | -0.91256885 | 0.7587227 | 0.9999931 |
| $\bar{Y}_{12}$-$\bar{Y}_{41}$ | -1.08E+00 | -1.91256885 | -0.2412773 | 0.0026025 |
| $\bar{Y}_{22}$-$\bar{Y}_{41}$ | -8.46E-01 | -1.68179962 | -0.01050807 | 0.0447668 |
| $\bar{Y}_{32}$-$\bar{Y}_{41}$ | -8.21E-01 | -1.65615859 | 0.01513295 | 0.0584452 |
| $\bar{Y}_{42}$-$\bar{Y}_{41}$ | -2.82E-01 | -1.11769706 | 0.55359449 | 0.9695508 |
| $\bar{Y}_{22}$-$\bar{Y}_{12}$ | 2.31E-01 | -0.60487654 | 1.066415 | 0.9904498 |
| $\bar{Y}_{32}$-$\bar{Y}_{12}$ | 2.56E-01 | -0.57923552 | 1.09205603 | 0.9822424 |
| $\bar{Y}_{42}$-$\bar{Y}_{12}$ | 7.95E-01 | -0.04077398 | 1.63051757 | 0.0754826 |
| $\bar{Y}_{32}$-$\bar{Y}_{22}$ | 2.56E-02 | -0.81000475 | 0.8612868 | 1 |
| $\bar{Y}_{42}$-$\bar{Y}_{22}$ | 5.64E-01 | -0.27154321 | 1.39974834 | 0.4431778 |
| $\bar{Y}_{42}$-$\bar{Y}_{32}$ | 5.38E-01 | -0.29718423 | 1.37410731 | 0.5061964 |

## f. Principal component analysis

### Mode condition

| Correlation Matrix for mode condition | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Predominan t - Faint | Rough - Smooth | Long - Short | Strong - Weak | Clicky - Pleasant | Fast - Slow | Hard - Soft | Mechanica l - Buzzy | Bad - Good |
| Predominan t - Faint | 1 | | | | | | | | |
| Rough - Smooth | -6.00E-02 | 1 | | | | | | | |
| Long - Short | -1.30E-01 | 3.10E-01 | 1 | | | | | | |
| Strong - Weak | -0.58 | 1.50E-01 | 2.40E-01 | 1 | | | | | |
| Clicky - Pleasant | 2.00E-02 | 4.40E-01 | 2.10E-01 | 7.00E-02 | 1 | | | | |
| Fast - Slow | 0.13 | -0.16 | -0.47 | -1.50E-01 | -0.19 | 1 | | | |
| Hard - Soft | -0.32 | 2.90E-01 | 1.40E-01 | 0.51 | 0.4 | -0.23 | 1 | | |
| Mechanical - Buzzy | -0.05 | -1.00E-02 | -8.00E-02 | 0 | 0.16 | 0.01 | 0.22 | 1 | |
| Bad - Good | -0.01 | 4.40E-01 | 3.00E-01 | 0.08 | 0.62 | -0.23 | 0.29 | 0.06 | 1 |

| Principal Component Analysis of mode data | | | | | | | |
|---|---|---|---|---|---|---|---|
| | item | RC1 | RC2 | RC3 | RC4 | h2 | u2 |
| Clicky - Pleasant | 5 | 8.40E-01 | | | | 0.76 | 2.40E-01 |
| Bad - Good | 9 | 0.81 | | | | 6.90E-01 | 3.10E-01 |
| Rough - Smooth | 2 | 7.40E-01 | | | | 6.10E-01 | 3.90E-01 |
| Strong - Weak | 4 | | 8.90E-01 | | | 0.81 | 1.90E-01 |
| Predominant - Faint | 1 | | -8.30E-01 | | | 0.71 | 2.90E-01 |
| Hard - Soft | 7 | 0.43 | 0.62 | | 0.33 | 0.68 | 0.32 |
| Fast - Slow | 6 | | | -0.88 | | 0.8 | 0.2 |
| Long - Short | 3 | | | 0.79 | | 0.73 | 0.27 |
| Mechanical - Buzzy | 8 | | | | 0.94 | 0.88 | 0.12 |

| | RC1 | RC2 | RC3 | RC4 |
|---|---|---|---|---|
| SS loadings | 2.17E+00 | 1.9 | 1.47E+00 | 1.12 |
| Proportion Var | 2.40E-01 | 2.10E-01 | 1.60E-01 | 1.20E-01 |
| Cumulative Var | 0.24 | 4.50E-01 | 6.20E-01 | 0.74 |
| Proportion Explained | 3.30E-01 | 2.90E-01 | 2.20E-01 | 1.70E-01 |
| Cumulative Proportion | 0.33 | 0.61 | 0.83 | 1.00E+00 |

Test of the hypothesis that 4 components are sufficient.
The degrees of freedom for the null model are 36 and the objective function was 2.31
The degrees of freedom for the model are 6 and the objective function was 1.13
Fit based upon off diagonal values = 0.91

## Skip condition

| Principal Component Analysis of skip data | | | | | | | |
|---|---|---|---|---|---|---|---|
| | item | RC1 | RC3 | RC2 | RC4 | h2 | u2 |
| Clicky - Pleasant | 5 | 0.85 | | | | 7.50E-01 | 2.49E-01 |
| Bad - Good | 9 | 8.00E-01 | 3.10E-01 | | | 7.50E-01 | 2.54E-01 |
| Rough - Smooth | 2 | 7.30E-01 | | | | 0.6 | 4.02E-01 |
| Strong - Weak | 7 | 6.60E-01 | | 3.30E-01 | | 0.71 | 2.95E-01 |
| Predominant - Faint | 6 | | -0.87 | | | 0.78 | 0.222 |
| Hard - Soft | 3 | | 7.80E-01 | | | 0.7 | 0.303 |
| Fast - Slow | 1 | | | -0.87 | | 0.76 | 0.236 |
| Long - Short | 4 | | | 0.82 | | 0.73 | 0.27 |
| Mechanical - Buzzy | 8 | | | | 0.95 | 0.92 | 0.085 |

| | RC1 | RC3 | RC2 | RC4 |
|---|---|---|---|---|
| SS loadings | 2.42 | 1.59 | 1.59 | 1.09 |
| Proportion Var | 0.27 | 0.18 | 0.18 | 0.12 |
| Cumulative Var | 0.27 | 0.44 | 0.62 | 0.74 |
| Proportion Explained | 0.36 | 0.24 | 0.24 | 0.16 |
| Cumulative Proportion | 0.36 | 0.6 | 0.84 | 1 |

Test of the hypothesis that 4 components are sufficient.
The degrees of freedom for the null model are 36 and the objective function was 2.46
The degrees of freedom for the model are 6 and the objective function was 1.01

Fit based upon off diagonal values = 0.93

### b) *Third preliminary experiment*

#### a. Repertory Grid Design (German with item translations below)

| **Was haben Knopf 1 und 2 gemeinsam im Vergleich zu Knopf 3?** | | **Worin unterscheidet sich Knopf 3?** |
|---|---|---|
| | vs. | |
| | vs. | |
| | vs. | |
| | **Grid-Generierung** | |
| 3 und 7 | vs. | 4 |
| 1 und 5 | vs. | 2 |
| 2 und 6 | vs. | 3 |

| **Konstruktpol** | | | **Antwortskala** | | | **Gegenpol** |
|---|---|---|---|---|---|---|
| 0 | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | 0 |
| 0 | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | 0 |
| 0 | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | 0 |
| schwergängig | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | leichtgängig |
| dominant | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | kraftlos |
| rauh | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | gleichmäßig |
| lang | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | kurz |
| stark | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | schwach |
| klickig | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | angenehm |
| langsam | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | schnell |
| hart | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | weich |

| | | | | | | |
|---|---|---|---|---|---|---|
| mechanisch | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | buzzy |
| schlecht | ☐ *sehr* | ☐ *ein wenig* | ☐ *weder-noch* | ☐ *ein wenig* | ☐ *sehr* | gut |

## Item translations

| English | German |
|---|---|
| Sluggish – Responsive | Schwergängig – Leichtgängig |
| Predominant – Faint | Dominant – Kraftlos |
| Rough – Smooth | Rauh – Gleichmäßig |
| Long – Short | Lang – Kurz |
| Strong – Weak | Stark – Schwach |
| Clicky – Pleasant | Klickig – Angenehm |
| Fast – Slow | Schnell – Langsam |
| Hard – Soft | Hart – Weich |
| Mechanical – Buzzy | Mechanisch – Buzzy (Summend/vibrierend) |
| Bad – Good | Schlecht – Gut |

### b. Driving performance results

| Summary Shapiro-Wilk Tests for normal distribution of data per variable | | |
|---|---|---|
| Variable | Condition | |
| Mean Velocity | Mode Repertory Grid | $p < 0.01$ |
| Standard deviation of velocity | Mode Repertory Grid | $p < 0.01$ |
| Lane keeping quality | Mode Repertory Grid | $p < 0.01$ |
| Standard deviation of lane position | Mode Repertory Grid | $p = 0.2502$ |
| Steering wheel velocity | Mode Repertory Grid | $p < 0.01$ |
| Steering wheel angle velocity | Mode Repertory Grid | $p < 0.01$ |
| Mean Velocity | Mode Performance | $p < 0.01$ |
| Standard deviation of velocity | Mode Performance | $p < 0.01$ |
| Lane keeping quality | Mode Performance | $p < 0.01$ |
| Standard deviation of lane position | Mode Performance | $p < 0.01$ |
| Steering wheel velocity | Mode Performance | $p < 0.01$ |
| Steering wheel angle velocity | Skip Performance | $p < 0.01$ |
| Mean Velocity | Skip Repertory Grid | $p < 0.01$ |
| Standard deviation of velocity | Skip Repertory Grid | $p < 0.01$ |
| Lane keeping quality | Skip Repertory Grid | $p < 0.01$ |
| Standard deviation of lane position | Skip Repertory Grid | $p < 0.01$ |
| Steering wheel velocity | Skip Repertory Grid | $p < 0.01$ |
| Steering wheel angle velocity | Skip Repertory Grid | $p < 0.01$ |
| Mean Velocity | Skip Repertory Grid | $p < 0.01$ |
| Standard deviation of velocity | Skip Performance | $p < 0.01$ |
| Lane keeping quality | Skip Performance | $p < 0.01$ |
| Standard deviation of lane position | Skip Performance | $p < 0.01$ |
| Steering wheel velocity | Skip Performance | $p < 0.01$ |
| Steering wheel angle velocity | Skip Performance | $p < 0.01$ |

| Summary significant Friedman test of data per variable | | |
|---|---|---|
| Variable | Condition | p |
| Mean Velocity | Mode Repertory Grid | 0.2956 |
| Standard deviation of velocity | Mode Repertory Grid | 0.4578 |
| Lane keeping quality | Mode Repertory Grid | $F(1/31) = 1.247451$; $p = 0.2726221$ |
| Standard deviation of lane position | Mode Repertory Grid | 0.4301 |
| Steering wheel velocity | Mode Repertory Grid | 0.9649 |
| Steering wheel angle velocity | Mode Repertory Grid | 0.2679 |
| Mean Velocity | Mode Performance | 0.1969 |
| Standard deviation of velocity | Mode Performance | 0.5523 |
| Lane keeping quality | Mode Performance | 0.7548 |
| Standard deviation of lane position | Mode Performance | 0.5188 |
| Steering wheel velocity | Mode Performance | 0.7128 |
| Steering wheel angle velocity | Mode Performance | 0.6619 |
| Mean Velocity | Skip Repertory Grid | $\chi^2(2) = 56.875$ $p = 0.05821$ |
| Standard deviation of velocity | Skip Repertory Grid | 0.1684 |
| Lane keeping quality | Skip Repertory Grid | 0.7548 |
| Standard deviation of lane position | Skip Repertory Grid | 0.6065 |
| Steering wheel velocity | Skip Repertory Grid | $\chi^2(2) = 152.987$ $p = 0.0004764$ |
| Steering wheel angle velocity | Skip Repertory Grid | $\chi^2(2) = 12.5$ $p = 0.00193$ |
| Mean Velocity | Skip Performance | 0.5523 |
| Standard deviation of velocity | Skip Performance | 0.6065 |
| Lane keeping quality | Skip Performance | $\chi^2(2) = 60.625$ $p = 0.04826$ |
| Standard deviation of lane position | Skip Performance | 0.3247 |
| Steering wheel velocity | Skip Performance | 0.1953 |
| Steering wheel angle velocity | Skip Performance | 0.07956 |

| Posthoc Nemenyi Test for mode Repertory Grid Lane keeping quality | | |
|---|---|---|
| | $\bar{Y}_{11}$ | $\bar{Y}_{12}$ |
| $\bar{Y}_{12}$ | 0.046 | - |
| $\bar{Y}_{13}$ | 0.187 | 0.806 |

| Posthoc Nemenyi Test for skip Repertory Grid StW Velocity | | |
|---|---|---|
| | $\bar{Y}_{21}$ | $\bar{Y}_{22}$ |
| $\bar{Y}_{22}$ | 0.0093 | - |
| $\bar{Y}_{32}$ | 0.0848 | 0.6953 |

| Posthoc Nemenyi Test for skip Repertory Grid Lane StW Angle Velocity | | |
|---|---|---|
| | $\bar{Y}_{21}$ | $\bar{Y}_{22}$ |
| $\bar{Y}_{22}$ | 0.073 | - |
| $\bar{Y}_{32}$ | 0.033 | 0.948 |

| Posthoc Nemenyi Test for skip Performance Lane keeping quality | | |
|---|---|---|
| | $\bar{Y}_{21}$ | $\bar{Y}_{22}$ |
| $\bar{Y}_{22}$ | 0.806 | - |
| $\bar{Y}_{32}$ | 0.046 | 0.187 |

## c. Preference data

### Snap ratio results (mode)

| Summary precondition testing for good/bad scale |
|---|

| | |
|---|---|
| Shapiro Wilk normal distribution | W = 0.854. p= 2.741e-08 |
| Levene's Test for Homogeneity | F (2/93) = 0.7809. p = 0.461 |

| ANOVA for good/bad scale for snap ratio (mode) | | | | | |
|---|---|---|---|---|---|
| | Df | Sum Sq | Mean Sq | F | p |
| Setting | 2 | 0.40 | 0.1979 | 0.196 | 0.823 |
| Residuals | 93 | 94.09 | 10.118 | | |

| Summary precondition testing for distance to ideal (mode) | |
|---|---|
| Shapiro Wilk normal distribution | W = 0.8801. p < 2.2e-16 |
| Levene's Test for Homogeneity | F (3/1276) = 0.5171. p = 0.6706 |

| ANOVA for distance to ideal (mode) | | | | | |
|---|---|---|---|---|---|
| | Df | Sum Sq | Mean Sq | F | p |
| Setting | 3 | 14.70 | 4.903 | 3.15 | 0.0242* |
| Residuals | 1276 | 1986.3 | 1.557 | | |

| Tukey HSD test for distance to ideal (mode) | | | | |
|---|---|---|---|---|
| | diff | lwr | upr | p |
| Ideal-$\overline{Y}_{31}$ | 2.59E-01 | 0.005637836 | 0.513112164 | 0.043* |
| $\overline{Y}_{11}$-$\overline{Y}_{31}$ | 4.06E-02 | -2.13E-01 | 2.94E-01 | 0.976 |
| $\overline{Y}_{21}$-$\overline{Y}_{31}$ | 3.13E-03 | -2.51E-01 | 0.256862164 | 1.000 |
| $\overline{Y}_{11}$-Ideal | -2.19E-01 | -4.72E-01 | 3.50E-02 | 0.119 |
| $\overline{Y}_{21}$-Ideal | -2.56E-01 | -0.509987164 | -2.51E-03 | 0.047* |
| $\overline{Y}_{21}$-$\overline{Y}_{11}$ | -3.75E-02 | -2.91E-01 | 2.16E-01 | 0.981 |

# Time between pulses results (skip)

| Summary precondition testing for good/bad scale (skip) | |
|---|---|
| Shapiro Wilk normal distribution | W = 0.8269. p = 5.742e-11 |
| Levene's Test for Homogeneity | F (3/124) = 17.259. p = 1.988e-09 |

| ANOVA for good/bad scale for Time between Pulses (skip) | | | | | |
|---|---|---|---|---|---|
| | Df | Sum Sq | Mean Sq | F | p |
| Setting | 2 | 7.65 | 3.823 | 2.967 | 0.0564 |
| Residuals | 93 | 119.84 | 1.289 | | |

| Summary precondition testing for distance to ideal (skip) | |
|---|---|
| Shapiro Wilk normal distribution | W = 0.8796. p < 2.2e-16 |
| Levene's Test for Homogeneity | F (3/1275) = 0.9074. p = 0.4368 |

| ANOVA for distance to ideal (skip) | | | | | |
|---|---|---|---|---|---|
| | Df | Sum Sq | Mean Sq | F | p |
| Setting | 3 | 28.2 | 9.410 | 5.13 | 0.00157 ** |
| Residuals | 1275 | 2338.6 | 1.834 | | |

| Tukey HSD test for distance to ideal (skip) | | | | |
|---|---|---|---|---|
| | diff | lwr | upr | p |
| Ideal-$\bar{Y}_{32}$ | 0.31863832 | 0.04299645 | 0.5942802 | 0.0158455* |
| $\bar{Y}_{12}$-$\bar{Y}_{32h}$ | 0.39587500 | 0.12044890 | 0.6713011 | 0.0012972** |
| $\bar{Y}_{22}$-$\bar{Y}_{32}$ | 0.25312500 | -0.02230110 | 0.5285511 | 0.0846308 |
| $\bar{Y}_{12}$-Ideal | 0.07723668 | -0.19840519 | 0.3528785 | 0.8888266 |
| $\bar{Y}_{22}$-Ideal | -0.06551332 | -0.34115519 | 0.2101285 | 0.9284768 |
| $\bar{Y}_{22}$-$\bar{Y}_{12}$ | -0.14275000 | -0.41817610 | 0.1326761 | 0.5418313 |

## d. Principal component analysis

| Principal Component Analysis with 3 Factors and oblimin rotation | | | | | |
|---|---|---|---|---|---|
| | TC1 | TC2 | TC3 | h2 | u2 |
| Responsive - Sluggish | 0.32 | 0.21 | 0.59 | 0.55 | 0.45 |
| Predominant - Faint | 0.8 | -0.19 | 0.11 | 0.69 | 0.31 |
| Rough - Smooth | 0.02 | 0.51 | 0.36 | 0.39 | 0.61 |
| Long - Short | 0 | -0.06 | 0.81 | 0.66 | 0.34 |
| Strong - Weak | 0.76 | -0.05 | -0.04 | 0.57 | 0.43 |
| Clicky - Pleasant | 0.12 | 0.83 | -0.08 | 0.73 | 0.27 |
| Fast - Slow | -0.09 | -0.04 | 0.81 | 0.65 | 0.35 |
| Hard - Soft | 0.66 | 0.34 | -0.11 | 0.58 | 0.42 |
| Mechanical - Buzzy | 0.31 | 0.29 | 0.07 | 0.2 | 0.8 |
| Bad - Good | -0.24 | 0.79 | 0.02 | 0.66 | 0.34 |

| | TC1 | TC2 | TC3 |
|---|---|---|---|
| SS loadings | 1.96 | 1.87 | 1.85 |
| Proportion Var | 0.2 | 0.19 | 0.18 |
| Cumulative Var | 0.2 | 0.38 | 0.57 |
| Proportion Explained | 0.35 | 0.33 | 0.33 |
| Cumulative Proportion | 0.35 | 0.67 | 1 |

| Intercorellations between Factors | | | |
|---|---|---|---|
| | TC1 | TC2 | TC3 |
| TC1 | 1.00 | 0.08 | 0.13 |
| TC2 | 0.08 | 1.00 | -0.02 |
| TC3 | 0.13 | -0.02 | 1.00 |

The degrees of freedom for the null model are 45 and the objective function was 1.88
The degrees of freedom for the model are 18 and the objective function was 0.7
Fit based upon off diagonal values = 0.8

| Principal Component Analysis with 4 Factors and oblimin rotation | | | | | | | |
|---|---|---|---|---|---|---|---|
| | item | TC1 | TC3 | TC2 | TC4 | h2 | u2 |
| Strong - Weak | 5 | 0.82 | | | | 0.64 | 0.36 |
| Predominant - Faint | 2 | 0.78 | | | | 0.69 | 0.31 |
| Hard - Soft | 8 | 0.61 | | | | 0.58 | 0.42 |
| Long - Short | 4 | | 0.81 | | | 0.67 | 0.33 |
| Fast - Slow | 7 | | 0.81 | | | 0.65 | 0.35 |
| Responsive - Sluggish | 1 | 0.36 | 0.60 | | | 0.60 | 0.40 |
| Bad - Good | 10 | | | 0.88 | | 0.77 | 0.23 |
| Clicky - Pleasant | 6 | | | 0.82 | | 0.76 | 0.24 |
| Mechanical - Buzzy | 9 | | | | 0.87 | 0.74 | 0.26 |
| Rough - Smooth | 3 | | | | 0.61 | 0.56 | 0.44 |

| | TC1 | TC3 | TC2 | TC4 |
|---|---|---|---|---|
| SS loadings | 1.87 | 1.82 | 1.71 | 1.27 |
| Proportion Var | 0.19 | 0.18 | 0.17 | 0.13 |
| Cumulative Var | 0.19 | 0.37 | 0.54 | 0.67 |
| Proportion Explained | 0.28 | 0.27 | 0.26 | 0.19 |
| Cumulative Proportion | 0.28 | 0.55 | 0.81 | 1 |

| Intercorellations between Factors | | | | |
|---|---|---|---|---|
| | TC1 | TC3 | TC2 | TC4 |
| TC1 | 1.00 | 0.11 | 0.04 | 0.19 |
| TC3 | 0.11 | 1.00 | 0.00 | 0.09 |
| TC2 | 0.04 | 0.00 | 1.00 | 0.16 |

The degrees of freedom for the null model are 45 and the objective function was 1.88
The degrees of freedom for the model are 11 and the objective function was 1.11
Fit based upon off diagonal values = 0.81

| Principal Component Analysis with 5 Factors and oblimin rotation | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | item | TC1 | TC3 | TC2 | TC5 | TC4 | h2 | u2 |
| Strong - Weak | 5 | 0.83 | | | | | 0.66 | 0.34 |
| Predominant - Faint | 2 | 0.74 | | | | | 0.71 | 0.29 |
| Hard - Soft | 8 | 0.66 | | | | | 0.63 | 0.37 |
| Fast - Slow | 7 | | 0.83 | | | | 0.68 | 0.32 |
| Long - Short | 4 | | 0.73 | | 0.35 | | 0.74 | 0.26 |
| Responsive - Sluggish | 1 | | 0.67 | | | | 0.69 | 0.31 |
| Bad - Good | 10 | | | 0.92 | | | 0.82 | 0.18 |
| Clicky - Pleasant | 6 | | | 0.77 | | | 0.76 | 0.24 |
| Rough - Smooth | 3 | | | | 0.87 | | 0.84 | 0.16 |
| Mechanical - Buzzy | 9 | | | | | 0.97 | 0.94 | 0.06 |

| | TC1 | TC3 | TC2 | TC5 | TC4 |
|---|---|---|---|---|---|
| SS loadings | 1.87 | 1.78 | 1.63 | 1.12 | 1.09 |
| Proportion Var | 0.19 | 0.18 | 0.16 | 0.11 | 0.11 |
| Cumulative Var | 0.19 | 0.36 | 0.53 | 0.64 | 0.75 |
| Proportion Explained | 0.25 | 0.24 | 0.22 | 0.15 | 0.15 |
| Cumulative Proportion | 0.25 | 0.49 | 0.7 | 0.85 | 1 |

| Intercorrelations between Factors | | | | | |
|---|---|---|---|---|---|
| | TC1 | TC3 | TC2 | TC5 | TC4 |
| TC1 | 1.00 | 0.12 | 0.06 | 0.08 | 0.18 |
| TC3 | 0.12 | 1.00 | -0.04 | 0.09 | 0.06 |
| TC2 | 0.06 | -0.04 | 1.00 | 0.16 | 0.11 |

The degrees of freedom for the null model are 45 and the objective function was 1.88
The degrees of freedom for the model are 5 and the objective function was 1.33
Fit based upon off diagonal values = 0.86

| Principal Component Analysis with 5 Factors and varimax rotation | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | item | RC1 | RC3 | RC2 | RC5 | RC4 | h2 | u2 |
| Strong - Weak | 5 | 0.81 | | | | | 0.66 | 0.34 |
| Predominant - Faint | 2 | 0.76 | | | | | 0.71 | 0.29 |
| Hard - Soft | 8 | 0.67 | | | 0.33 | | 0.63 | 0.37 |
| Fast - Slow | 7 | | 0.82 | | | | 0.68 | 0.32 |
| Long - Short | 4 | | 0.75 | | 0.34 | | 0.74 | 0.26 |
| Responsive - Sluggish | 1 | 0.35 | 0.68 | | | | 0.69 | 0.31 |
| Bad - Good | 10 | | | 0.89 | | | 0.82 | 0.18 |
| Clicky - Pleasant | 6 | | | 0.81 | | | 0.76 | 0.24 |
| Rough - Smooth | 3 | | | | 0.88 | | 0.84 | 0.16 |
| Mechanical - Buzzy | 9 | | | | | 0.95 | 0.94 | 0.06 |

| | RC1 | RC3 | RC2 | RC5 | RC4 |
|---|---|---|---|---|---|
| SS loadings | 1.87 | 1.79 | 1.65 | 1.11 | 1.07 |
| Proportion Var | 0.19 | 0.18 | 0.16 | 0.11 | 0.11 |
| Cumulative Var | 0.19 | 0.37 | 0.53 | 0.64 | 0.75 |
| Proportion Explained | 0.25 | 0.24 | 0.22 | 0.15 | 0.14 |
| Cumulative Proportion | 0.25 | 0.49 | 0.71 | 0.86 | 1 |

The degrees of freedom for the null model are 45 and the objective function was 1.88
The degrees of freedom for the model are 5 and the objective function was 1.33
Fit based upon off diagonal values = 0.86

## e. Cronbach's α

| Cronbach's α comparison | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | ase | mean | sd | lower | alpha | upper |
| Potency (RC1) | 0.64 | 0.64 | 0.55 | 0.370 | 1.8 | 0.07 | -0.21 | 0.86 | 0.5 | 0.64 | 0.78 |
| Valence (RC2) | 0.67 | 0.69 | 0.53 | 0.53 | 2.20 | 0.095 | 0.62 | 0.095 | 0.49 | 0.67 | 0.86 |
| Dynamics (RC3) | 0.66 | 0.66 | 0.56 | 0.39 | 1.90 | 0.068 | 0.39 | 0.9 | 0.52 | 0.66 | 0.79 |
| Combined | 0.63 | 0.63 | 0.69 | 0.15 | 1.70 | 0.042 | 0.23 | 0.57 | 0.55 | 0.63 | 0.71 |

| Reliability if an item is dropped | | | | | | |
|---|---|---|---|---|---|---|
| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | ase |
| Responsive - Sluggish | 0.57 | 0.57 | 0.63 | 0.13 | 1.3 | 0.048 |
| Predominant - Faint | 0.61 | 0.61 | 0.66 | 0.150 | 1.5 | 0.045 |
| Rough - Smooth | 0.6 | 0.6 | 0.670 | 0.15 | 1.50 | 0.046 |
| Long - Short | 0.62 | 0.62 | 0.67 | 0.15 | 1.60 | 0.044 |
| Strong - Weak | 0.61 | 0.61 | 0.67 | 0.15 | 1.60 | 0.045 |
| Clicky - Pleasant | 0.6 | 0.6 | 0.64 | 0.140 | 1.5 | 0.046 |
| Fast - Slow | 0.63 | 0.63 | 0.68 | 0.160 | 1.7 | 0.044 |
| Hard - Soft | 0.59 | 0.59 | 0.65 | 0.140 | 1.5 | 0.047 |
| Mechanical - Buzzy | 0.62 | 0.62 | 0.68 | 0.15 | 1.6 | 0.044 |
| Bad - Good | 0.64 | 0.65 | 0.68 | 0.17 | 1.8 | 0.043 |

| Item statistics | | | | | | |
|---|---|---|---|---|---|---|
| | n | r | r.cor | r.drop | mean | sd |
| Responsive - Sluggish | 256.00 | 0.65 | 0.61 | 0.49 | 0.60 | 1.10 |
| Predominant - Faint | 256.00 | 0.50 | 0.44 | 0.31 | -0.35 | 1.10 |
| Rough - Smooth | 256.00 | 0.51 | 0.40 | 0.34 | 0.95 | 1.20 |
| Long - Short | 256.00 | 0.45 | 0.37 | 0.25 | 0.38 | 1.20 |
| Strong - Weak | 256.00 | 0.48 | 0.39 | 0.29 | -0.29 | 1.10 |
| Clicky - Pleasant | 256.00 | 0.51 | 0.47 | 0.33 | 0.31 | 1.50 |
| Fast - Slow | 256.00 | 0.41 | 0.32 | 0.22 | 0.20 | 1.20 |
| Hard - Soft | 256.00 | 0.55 | 0.48 | 0.39 | 0.02 | 1.10 |
| Mechanical - Buzzy | 256.00 | 0.44 | 0.31 | 0.25 | -0.49 | 1.10 |
| Bad - Good | 256.00 | 0.33 | 0.24 | 0.16 | 0.94 | 1.10 |

## *c) Main experiment*

### a. Results of the pretest

| Subject | 1st Element | 2nd Element | 3rd Element | Construct Pole 1 | Construct Pole 2 |
|---|---|---|---|---|---|
| 1 | Capacitive | Mechanical | 3D | No communality | No opposition |
| 2 | Mechanical | Capacitive | 3D | quiet | loud |
| 3 | 3D | Capacitive | Mechanical | strange | normal |
| 4 | Capacitive | 3D | Mechanical | touch | Pressure switch |
| 5 | Mechanical | 3D | Capacitive | Feedback | No feedback |
| 6 | 3D | Mechanical | Capacitive | Feedback | No feedback |
| 7 | 3D | Capacitive | Mechanical | moving | sluggish |
| 8 | 3D | Mechanical | Capacitive | Feedback | No feedback |

## b. Grid design for main experiment (German with item translations below)

| | *sehr* | *ein wenig* | *weder-noch* | *ein wenig* | *sehr* | |
|---|---|---|---|---|---|---|
| gibt Feedback | O | O | O | O | O | gibt kein Feedback |
| schwergängig | O | O | O | O | O | leichtgängig |
| kraftlos | O | O | O | O | O | dominant |
| rauh | O | O | O | O | O | gleichmäßig |
| kurz | O | O | O | O | O | lang |
| stark | O | O | O | O | O | schwach |
| angenehm | O | O | O | O | O | klickig |
| langsam | O | O | O | O | O | schnell |
| weich | O | O | O | O | O | hart |
| mechanisch | O | O | O | O | O | summend/ vibrierend |
| scharf | O | O | O | O | O | stumpf |
| niedrigfrequent | O | O | O | O | O | hochfrequent |
| dynamisch | O | O | O | O | O | träge |
| schlecht | O | O | O | O | O | gut |

Item translations (order with polarity for analysis applied):

| English | German | Item polarity changed before analysis (yes/no) |
|---|---|---|
| Feedback – No Feedback | Gibt Feedback – Gibt kein Feedback | no |
| Responsive – Sluggish | Schwergängig – Leichtgängig | yes |
| Predominant – Faint | Dominant – Kraftlos | yes |
| Smooth – Rough | Rauh – Gleichmäßig | yes |
| Short – Long | Lang – Kurz | no |
| Weak – Strong | Schwach - Stark | yes |
| Pleasant – Clicky | Klickig – Angenehm | no |
| Fast – Slow | Schnell – Langsam | yes |
| Soft – Hard | Weich – Hart | no |
| Mechanical – Buzzy | Mechanisch – Summend/vibrierend | no |
| Sharp – Dull | Scharf – Stumpf | no |
| High Frequency – Low Frequency | Hochfrequent – Niedrigfrequent | yes |
| Dynamic – Tranquil | Dynamisch – Träge | no |
| Good – Bad | Schlecht – Gut | yes |

### c. Switch Perception Questionnaire (adapted after Wellings et al., 2009)

**<u>Bei der Bedienung des Schalters verwenden Sie bitte zunächst folgende Strategien in dieser Reihenfolge:</u>**

1. Drücken Sie den Schalter mindestens drei Mal.
2. Variieren Sie die Geschwindigkeit mit der Sie den Schalter drücken.
3. Drücken Sie den Schalter bewusst ein und wieder aus.

Es sollen Ihre Erfahrungen bei der **<u>Bedienung des Schalters</u>** eingefangen werden. Bitte markieren Sie auf der Skala wie gut die gegensätzlichen Wörter den Schalter beschreiben.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| schwer | O | O | O | O | O | O | O | leicht |
| präzise | O | O | O | O | O | O | O | unpräzise |
| billig | O | O | O | O | O | O | O | teuer |
| kultiviert | O | O | O | O | O | O | O | unkultiviert |
| klickig | O | O | O | O | O | O | O | geschmeidig |
| angenehm | O | O | O | O | O | O | O | lästig |
| lose | O | O | O | O | O | O | O | fest |
| solide | O | O | O | O | O | O | O | fadenscheinig |
| langweilig | O | O | O | O | O | O | O | interessant |
| modern | O | O | O | O | O | O | O | altmodisch |

### d. Twofactorial ANOVA of the distance to ideal

The Main effect Switch describes the different switch types: 3D, capacitive, and mechanical. The main effect course describes the four different driving maneuvers: Highway straight, highway overtaking, rural road straight, and rural road curves.

| | Effect | DFn | DFd | SSn | SSd | F | p | p<.05 ges |
|---|---|---|---|---|---|---|---|---|
| 1 | (Intercept) | 1 | 40 | 22691 | 106.2 | 8546 | 0.000 | * 0.92861069 |
| 2 | Switch | 3 | 120 | 100.3 | 282.7 | 121.1471 | 0.000 | * 0.05434661 |
| 3 | Course | 13 | 520 | 236.6 | 400.9 | 274.1267 | 0.000 | * 0.11941975 |
| 4 | Switch: Course | 39 | 1560 | 354.9 | 954.5 | 566.3774 | 0.000 | * 0.16906569 |

| Mauchly's Te st for Sphericity | | | | | | |
|---|---|---|---|---|---|---|
| | Effect | DFn | DFd | W | p p< | 0.05 |
| 2 | Switch | 5.98 | 1378e | -01 1.31683 | 0.009 | * |
| 3 | Course | 8.28 | 0509e | -03 1.05918 | 0.000 | * |
| 4 | Switch: Course | 2.24 | 5116e | -23 3.51729 | 0.000 | * |

| Sphericity Corrections | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Effect | | GGe | p[G | G] p[GG] | <.05 | HFe p[H | F] p[HF]<.05 | |
| 2 | Switch | 0.81 | 92241 | 6.643264e- | 7 | * 0.8769 | 174 3.027552e- | 07 | * |
| 3 | Course | 0.52 | 14806 | 2.061558e- | 24 | * 0.6389 | 639 2.245053e- | 29 | * |
| 4 | Switch: Course | 0.35 | 65050 | 1.707884e- | 30 | * 0.5581 | 412 2.204891e- | 46 | * |

## e. Posthoc tests for the main effects

Hommel-corrected robust pairwise test for switch configuration

| Comparison | Statistic | p.value | p.adjust |
|---|---|---|---|
| IDEAL - Capacitive | 0.4165 | 0 | 0 |
| IDEAL - Mechanical | 0.002033 | 0.9 | 0.9 |
| IDEAL - 3D | 0.4213 | 0 | 0 |
| Capacitive - Mechanical | -0.4145 | 0 | 0 |
| Capacitive - 3D | 0.004791 | 0.896 | 0.9 |

| Hommel-corrected robust pairwise test for measurement scale | | | | |
|---|---|---|---|---|
| Comparison | | Statistic | p.value | p.adjust |
| Feedback – No Feedback | Responsive – Sluggish | 0.2603 | 0.236 | 0.984 |
| Feedback – No Feedback | Predominant – Faint | 1.065 | 0 | 0 |
| Feedback – No Feedback | Smooth – Rough | 0.215 | 0.216 | 0.984 |
| Feedback – No Feedback | Short – Long | 0.781 | 0 | 0 |
| Feedback – No Feedback | Weak – Strong | 0.6702 | 0 | 0 |
| Feedback – No Feedback | Pleasant – Clicky | -0.406 | 0.032 | 0.736 |
| Feedback – No Feedback | Fast – Slow | 0.6565 | 0 | 0 |
| Feedback – No Feedback | Soft – Hard | 0.6946 | 0 | 0 |
| Feedback – No Feedback | Mechanical – Buzzy | 0.3648 | 0.032 | 0.736 |
| Feedback – No Feedback | Sharp – Dull | 0.957 | 0 | 0 |
| Feedback – No Feedback | High Frequency – Low Frequency | 0.721 | 0 | 0 |
| Feedback – No Feedback | Dynamic – Tranquil | -1.231 | 0 | 0 |
| Feedback – No Feedback | Good – Bad | 0.2169 | 0.3 | 0.984 |
| Responsive – Sluggish | Predominant – Faint | -0.8047 | 0 | 0 |
| Responsive – Sluggish | Smooth – Rough | 0.04527 | 0.96 | 0.984 |
| Responsive – Sluggish | Short – Long | -0.5207 | 0 | 0 |
| Responsive – Sluggish | Weak – Strong | 0.4099 | 0.008 | 0.264 |
| Responsive – Sluggish | Pleasant – Clicky | -0.1457 | 0.216 | 0.984 |
| Responsive – Sluggish | Fast – Slow | -0.3962 | 0 | 0 |
| Responsive – Sluggish | Soft – Hard | 0.4343 | 0 | 0 |
| Responsive – Sluggish | Mechanical – Buzzy | -0.1045 | 0.364 | 0.984 |
| Responsive – Sluggish | Sharp – Dull | 0.6967 | 0 | 0 |
| Responsive – Sluggish | High Frequency – Low Frequency | -0.4607 | 0 | 0 |
| Responsive – Sluggish | Dynamic – Tranquil | -0.9705 | 0 | 0 |
| Responsive – Sluggish | Good – Bad | 0.04332 | 0.984 | 0.984 |
| Predominant – Faint | Smooth – Rough | -0.85 | 0 | 0 |
| Predominant – Faint | Short – Long | -0.284 | 0.016 | 0.464 |
| Predominant – Faint | Weak – Strong | -0.3948 | 0 | 0 |
| Predominant – Faint | Pleasant – Clicky | 0.659 | 0 | 0 |
| Predominant – Faint | Fast – Slow | -0.4085 | 0 | 0 |
| Predominant – Faint | Soft – Hard | -0.3704 | 0 | 0 |
| Predominant – Faint | Mechanical Buzzy | -0.7002 | 0 | 0 |
| Predominant – Faint | Sharp – Dull | -0.108 | 0.296 | 0.984 |
| Predominant – Faint | High Frequency – Low Frequency | 0.344 | 0 | 0 |
| Predominant – Faint | Dynamic – Tranquil | -0.1658 | 0.484 | 0.984 |
| Predominant – Faint | Good – Bad | -0.8481 | 0 | 0 |
| Smooth – Rough | Short – Long | -0.566 | 0 | 0 |
| Smooth – Rough | Weak – Strong | 0.4552 | 0.004 | 0.156 |
| Smooth – Rough | Pleasant – Clicky | -0.191 | 0.208 | 0.984 |
| Smooth – Rough | Fast – Slow | -0.4415 | 0.004 | 0.156 |
| Smooth – Rough | Soft – Hard | 0.4796 | 0.004 | 0.156 |
| Smooth – Rough | Mechanical – Buzzy | -0.1498 | 0.26 | 0.984 |
| Smooth – Rough | Sharp – Dull | 0.742 | 0 | 0 |
| Smooth – Rough | High Frequency – Low Frequency | -0.506 | 0 | 0 |
| Smooth – Rough | Dynamic – Tranquil | -1.016 | 0 | 0 |
| Smooth – Rough | Good – Bad | 0.001945 | 0.908 | 0.984 |
| Short – Long | Weak – Strong | -0.1108 | 0.308 | 0.984 |
| Short – Long | Pleasant – Clicky | 0.375 | 0.008 | 0.264 |
| Short – Long | Fast – Slow | -0.1245 | 0.228 | 0.984 |
| Short – Long | Soft – Hard | -0.08638 | 0.448 | 0.984 |
| Short – Long | Mechanical – Buzzy | -0.4162 | 0 | 0 |
| Short – Long | Sharp – Dull | 0.176 | 0.22 | 0.984 |
| Short – Long | High Frequency – Low Frequency | 0.05999 | 0.488 | 0.984 |
| Short – Long | Dynamic – Tranquil | -0.4498 | 0.008 | 0.264 |
| Short – Long | Good – Bad | -0.564 | 0 | 0 |
| Weak – Strong | Pleasant – Clicky | 0.2642 | 0.036 | 0.7669 |
| Weak – Strong | Fast – Slow | 0.01372 | 0.86 | 0.984 |
| Weak – Strong | Soft – Hard | 0.02439 | 0.88 | 0.984 |
| Weak – Strong | Mechanical – Buzzy | 0.3054 | 0.012 | 0.372 |
| Weak – Strong | Sharp – Dull | 0.2868 | 0.028 | 0.7 |
| Weak – Strong | High Frequency – Low Frequency | -0.05078 | 0.436 | 0.984 |
| Weak – Strong | Dynamic – Tranquil | -0.5606 | 0 | 0 |
| Weak – Strong | Good – Bad | 0.4533 | 0 | 0 |
| Pleasant – Clicky | Fast – Slow | 0.2505 | 0.044 | 0.836 |

| | | | | |
|---|---|---|---|---|
| Pleasant – Clicky | Soft – Hard | 0.2886 | 0.032 | 0.736 |
| Pleasant – Clicky | Mechanical – Buzzy | -0.04116 | 0.884 | 0.984 |
| Pleasant – Clicky | Sharp – Dull | 0.551 | 0 | 0 |
| Pleasant – Clicky | High Frequency – Low Frequency | 0.315 | 0 | 0 |
| Pleasant – Clicky | Dynamic – Tranquil | 0.8248 | 0 | 0 |
| Pleasant – Clicky | Good – Bad | -0.189 | 0.188 | 0.984 |
| Fast – Slow | Soft – Hard | 0.03811 | 0.784 | 0.984 |
| Fast – Slow | Mechanical – Buzzy | -0.2917 | 0.02 | 0.58 |
| Fast – Slow | Weak – Strong | 0.3005 | 0.008 | 0.264 |
| Fast – Slow | High Frequency – Low Frequency | -0.0645 | 0.376 | 0.984 |
| Fast – Slow | Dynamic – Tranquil | -0.5743 | 0 | 0 |
| Fast – Slow | Good – Bad | -0.4395 | 0 | 0 |
| Soft – Hard | Mechanical – Buzzy | 0.3298 | 0.008 | 0.264 |
| Soft – Hard | Sharp – Dull | -0.2624 | 0.06 | 0.96 |
| Soft – Hard | High Frequency – Low Frequency | -0.02639 | 0.672 | 0.984 |
| Soft – Hard | Dynamic – Tranquil | -0.5362 | 0 | 0 |
| Soft – Hard | Good – Bad | 0.4776 | 0.004 | 0.156 |
| Mechanical – Buzzy | Sharp – Dull | 0.5922 | 0 | 0 |
| Mechanical – Buzzy | High Frequency – Low Frequency | -0.3562 | 0 | 0 |
| Mechanical – Buzzy | Dynamic – Tranquil | -0.866 | 0 | 0 |
| Mechanical – Buzzy | Good – Bad | -0.1479 | 0.272 | 0.984 |
| Sharp – Dull | High Frequency – Low Frequency | 0.236 | 0.052 | 0.884 |
| Sharp – Dull | Dynamic – Tranquil | -0.2738 | 0.06 | 0.96 |
| Sharp – Dull | Good – Bad | 0.74 | 0 | 0 |
| High Frequency – Low frequency | Dynamic – Tranquil | -0.5098 | 0 | 0 |
| High Frequency – Low frequency | Good – Bad | -0.504 | 0 | 0 |
| Dynamic – Tranquil | Good – Bad | -1.014 | 0 | 0 |

## f. Factor determination

Number of Factors retrieved after different criteria

| vss1 | vss2 | map | dof | chisq | prob | sqresid | fit | RMSEA | BIC | SABIC | complex | eChisq | eRMS | eCRMS | eBIC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.58 | 0.00 | 0.05 | 77.00 | 1200.00 | 0.00 | 12.50 | 0.58 | 0.17 | 718.70 | 963.10 | 1.00 | 2100.00 | 0.15 | 0.17 | 1628.00 |
| 0.65 | 0.80 | 0.03 | 64.00 | 400.00 | 0.00 | 6.10 | 0.80 | 0.11 | 7.30 | 210.46 | 1.20 | 370.00 | 0.06 | 0.08 | -29.80 |
| 0.55 | 0.78 | 0.03 | 52.00 | 150.00 | 0.00 | 4.50 | 0.85 | 0.06 | -170.20 | -5.17 | 1.50 | 100.00 | 0.03 | 0.05 | -218.60 |
| 0.52 | 0.74 | 0.04 | 41.00 | 83.00 | 0.00 | 3.80 | 0.87 | 0.05 | -171.10 | -40.99 | 1.70 | 42.00 | 0.02 | 0.03 | -212.30 |
| 0.49 | 0.71 | 0.06 | 31.00 | 40.00 | 0.13 | 3.50 | 0.88 | 0.03 | -152.00 | -53.63 | 1.80 | 23.00 | 0.02 | 0.03 | -168.90 |
| 0.49 | 0.70 | 0.08 | 22.00 | 28.00 | 0.18 | 3.30 | 0.89 | 0.02 | -108.60 | -38.77 | 1.90 | 14.00 | 0.01 | 0.03 | -122.10 |
| 0.49 | 0.70 | 0.10 | 14.00 | 20.00 | 0.13 | 3.20 | 0.89 | 0.03 | -66.80 | -22.39 | 1.90 | 12.00 | 0.01 | 0.03 | -74.40 |
| 0.47 | 0.66 | 0.13 | 7.00 | 12.00 | 0.09 | 3.00 | 0.90 | 0.04 | -31.20 | -8.95 | 2.00 | 9.20 | 0.01 | 0.04 | -34.20 |
| 0.52 | 0.67 | 0.18 | 1.00 | 2.40 | 0.12 | 2.50 | 0.92 | 0.06 | -3.80 | -0.59 | 2.10 | 1.70 | 0.00 | 0.04 | -4.50 |
| 0.38 | 0.61 | 0.25 | -4.00 | 0.14 | NA | 2.30 | 0.92 | NA | NA | NA | 2.20 | 0.12 | 0.00 | NA | NA |
| 0.47 | 0.64 | 0.32 | -8.00 | 0.00 | NA | 2.20 | 0.92 | NA | NA | NA | 2.10 | 0.00 | 0.00 | NA | NA |
| 0.43 | 0.62 | 0.49 | -11.00 | 0.00 | NA | 2.20 | 0.93 | NA | NA | NA | 2.10 | 0.00 | 0.00 | NA | NA |
| 0.46 | 0.61 | 1.00 | -13.00 | 0.00 | NA | 2.10 | 0.93 | NA | NA | NA | 2.10 | 0.00 | 0.00 | NA | NA |
| 0.46 | 0.61 | NA | -14.00 | 0.00 | NA | 2.10 | 0.93 | NA | NA | NA | 2.10 | 0.00 | 0.00 | NA | NA |

## g. Principal component analyses with three, four and five factors

| Principal Component Analysis with 3 Factors and varimax rotation | | | | | | |
|---|---|---|---|---|---|---|
| | item | RC1 | RC2 | RC3 | h2 | u2 |
| Sharp – Dull | 11 | 0.79 | | | 0.62 | 0.38 |
| Fast – Slow | 8 | 0.76 | | | 0.61 | 0.39 |
| Dynamic – Tranquil | 13 | 0.74 | | 0.36 | 0.71 | 0.29 |
| High Frequency – Low Frequency | 12 | 0.66 | | | 0.44 | 0.56 |
| Short – Long | 5 | 0.50 | | 0.33 | 0.37 | 0.63 |
| Weak – Strong | 6 | | 0.84 | | 0.75 | 0.25 |
| Predominant – Faint | 3 | | 0.84 | | 0.71 | 0.29 |
| Feedback – No Feedback | 1 | | 0.71 | | 0.59 | 0.41 |
| Good – Bad | 14 | 0.40 | 0.57 | 0.52 | 0.76 | 0.24 |
| Mechanical – Buzzy | 10 | | | | 0.04 | 0.96 |
| Pleasant – Clicky | 7 | | 0.35 | 0.72 | 0.68 | 0.32 |
| Soft – Hard | 9 | | -0.34 | 0.72 | 0.63 | 0.37 |
| Smooth – Rough | 4 | | | 0.7 | 0.51 | 0.49 |
| Responsive – Sluggish | 2 | 0.34 | | 0.69 | 0.66 | 0.34 |

| | RC1 | RC2 | RC3 |
|---|---|---|---|
| SS loadings | 2.83 | 2.63 | 2.63 |
| Proportion Var | 0.20 | 0.19 | 0.19 |
| Cumulative Var | 0.20 | 0.39 | 0.58 |
| Proportion Explained | 0.35 | 0.33 | 0.32 |
| Cumulative Proportion | 0.35 | 0.68 | 1.00 |

Test of the hypothesis that 3 components are sufficient.
The degrees of freedom for the null model are 91 and the objective function was 5.15
The degrees of freedom for the model are 52 and the objective function was 0.74
Fit based upon off diagonal values = 0.95
> sum (large.resid) = 41
> sum (large.resid)/nrow (residuals) = 0.4505495
> sqrt (mean(residuals^2)) = 0.06675841

| Principal Component Analysis with 4 Factors (according to BIC value) and varimax rotation | | | | | | | |
|---|---|---|---|---|---|---|---|
| | item | RC1 | RC2 | RC3 | RC4 | h2 | u2 |
| Sharp – Dull | 11 | 0.78 | | | | 0.62 | 0.38 |
| Fast – Slow | 8 | 0.77 | | | | 0.64 | 0.36 |
| Dynamic – Tranquil | 13 | 0.73 | | 0.35 | | 0.71 | 0.29 |
| High Frequency – Low Frequency | 12 | 0.61 | | | -0.4 | 0.55 | 0.45 |
| Short – Long | 5 | 0.54 | | 0.32 | | 0.46 | 0.54 |
| Weak – Strong | 6 | | 0.84 | | | 0.75 | 0.25 |
| Predominant – Faint | 3 | | 0.84 | | | 0.71 | 0.29 |
| Feedback – No Feedback | 1 | | 0.73 | | | 0.59 | 0.41 |
| Good – Bad | 14 | 0.37 | 0.62 | 0.5 | | 0.76 | 0.24 |
| Soft – Hard | 9 | | -0.31 | 0.73 | | 0.65 | 0.35 |
| Pleasant – Clicky | 7 | | 0.39 | 0.71 | | 0.68 | 0.32 |
| Responsive – Sluggish | 2 | 0.36 | | 0.69 | | 0.67 | 0.33 |
| Smooth – Rough | 4 | | | 0.69 | | 0.51 | 0.49 |
| Mechanical – Buzzy | 10 | | | | 0.89 | 0.8 | 0.2 |

| | RC1 | RC2 | RC3 | RC4 |
|---|---|---|---|---|
| SS loadings | 2.76 | 2.71 | 2.6 | 1.04 |
| Proportion Var | 0.2 | 0.19 | 0.19 | 0.07 |
| Cumulative Var | 0.2 | 0.39 | 0.58 | 0.65 |
| Proportion Explained | 0.3 | 0.3 | 0.29 | 0.11 |
| Cumulative Proportion | 0.3 | 0.6 | 0.89 | 1 |

Test of the hypothesis that 4 components are sufficient.
The degrees of freedom for the null model are 91 and the objective function was 5.15
The degrees of freedom for the model are 41 and the objective function was 1.05
Fit based upon off diagonal values = 0.94
sum (large.resid) = 44
sum (large.resid)/nrow(residuals) = 0.4835165
sqrt(mean(residuals^2)) = 0.07085327

| Principal Component Analysis with 5 Factors (according to empirical BIC value) and varimax rotation | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | item | RC2 | RC3 | RC1 | RC5 | RC4 | h2 | u2 |
| Weak – Strong | 6 | 0.84 | | | | | 0.75 | 0.25 |
| Predominant – Faint | 3 | 0.84 | | | | | 0.72 | 0.28 |
| Feedback – No Feedback | 1 | 0.74 | | | | | 0.63 | 0.37 |
| Good – Bad | 14 | 0.57 | 0.53 | 0.37 | | | 0.77 | 0.23 |
| Pleasant – Clicky | 7 | 0.32 | 0.76 | | | | 0.73 | 0.27 |
| Soft – Hard | 9 | -0.37 | 0.75 | | | | 0.71 | 0.29 |
| Smooth – Rough | 4 | | 0.64 | | 0.36 | | 0.55 | 0.45 |
| Responsive – Sluggish | 2 | | 0.63 | | 0.42 | | 0.67 | 0.33 |
| High Frequency – Low Frequency | 12 | | | 0.83 | | | 0.74 | 0.26 |
| Sharp – Dull | 11 | | | 0.74 | | | 0.64 | 0.36 |
| Dynamic – Tranquil | 13 | | 0.34 | 0.65 | 0.36 | | 0.71 | 0.29 |
| Fast – Slow | 8 | | | 0.57 | 0.56 | | 0.66 | 0.34 |
| Short – Long | 5 | | | | 0.85 | | 0.77 | 0.23 |
| Mechanical – Buzzy | 10 | | | | | 0.99 | 0.98 | 0.02 |

| | RC2 | RC3 | RC1 | RC5 | RC4 |
|---|---|---|---|---|---|
| SS loadings | 2.66 | 2.47 | 2.24 | 1.64 | 1.03 |
| Proportion Var | 0.19 | 0.18 | 0.16 | 0.12 | 0.07 |
| Cumulative Var | 0.19 | 0.37 | 0.53 | 0.64 | 0.72 |
| Proportion Explained | 0.26 | 0.25 | 0.22 | 0.16 | 0.1 |
| Cumulative Proportion | 0.26 | 0.51 | 0.73 | 0.9 | 1 |

Test of the hypothesis that 5 components are sufficient.
The degrees of freedom for the null model are 91 and the objective function was 5.15
The degrees of freedom for the model are 31 and the objective function was 1.03
Fit based upon off diagonal values = 0.96
sum (large.resid) = 31
sum (large.resid)/nrow(residuals) = 0.3406593
sqrt (mean(residuals^2)) = 0.05916282

## h. Cronbach's α of the main experiment

| Cronbach's α comparison | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | ase | mean | sd | lower | alpha | upper |
| Potency (RC1) | 0.78 | 0.79 | 0.72 | 0.55 | 3.7 | 0.042 | 3 | 1 | 0.7 | 0.78 | 0.86 |
| Valence (RC2) | 0.72 | 0.73 | 0.68 | 0.4 | 2.6 | 0.037 | 3.2 | 0.87 | 0.65 | 0.72 | 0.8 |
| Dynamics (RC3) | 0.77 | 0.77 | 0.75 | 0.4 | 3.3 | 0.03 | 2.9 | 0.74 | 0.71 | 0.77 | 0.83 |
| Combined | 0.79 | 0.78 | 0.84 | 0.21 | 3.6 | 0.018 | 3.1 | 0.58 | 0.75 | 0.79 | 0.82 |

| Reliability if an item is dropped: | | | | | | |
|---|---|---|---|---|---|---|
| | raw_alpha | std.alpha | G6(smc) | average_r | S/N | alpha se |
| Feedback – No Feedback | 0.77 | 0.77 | 0.83 | 0.2 | 3.3 | 0.02 |
| Responsive – Sluggish | 0.77 | 0.77 | 0.82 | 0.2 | 3.3 | 0.02 |
| Predominant – Faint | 0.79 | 0.79 | 0.84 | 0.22 | 3.7 | 0.019 |
| Smooth – Rough | 0.77 | 0.77 | 0.83 | 0.2 | 3.3 | 0.02 |
| Short – Long | 0.78 | 0.77 | 0.83 | 0.21 | 3.4 | 0.019 |
| Weak – Strong | 0.79 | 0.79 | 0.84 | 0.22 | 3.8 | 0.018 |
| Pleasant – Clicky | 0.76 | 0.76 | 0.82 | 0.19 | 3.1 | 0.021 |
| Fast – Slow | 0.76 | 0.76 | 0.82 | 0.2 | 3.2 | 0.02 |
| Soft – Hard | 0.79 | 0.79 | 0.84 | 0.22 | 3.7 | 0.018 |
| Mechanical – Buzzy | 0.8 | 0.8 | 0.86 | 0.24 | 4.1 | 0.018 |
| Sharp – Dull | 0.77 | 0.76 | 0.83 | 0.2 | 3.2 | 0.02 |
| High Frequency – Low Frequency | 0.78 | 0.77 | 0.84 | 0.21 | 3.4 | 0.019 |
| Dynamic – Tranquil | 0.75 | 0.74 | 0.81 | 0.18 | 2.9 | 0.021 |
| Good – Bad | 0.74 | 0.74 | 0.81 | 0.18 | 2.9 | 0.02 |

| Item statistics | | | | | | |
|---|---|---|---|---|---|---|
| | n | r | r.cor | r.drop | mean | sd |
| Feedback – No Feedback | 492 | 0.55 | 0.512 | 0.451 | 3.3 | 1.34 |
| Responsive – Sluggish | 492 | 0.55 | 0.527 | 0.437 | 3.4 | 1.28 |
| Predominant – Faint | 492 | 0.33 | 0.281 | 0.209 | 2.8 | 1.13 |
| Smooth – Rough | 492 | 0.54 | 0.488 | 0.443 | 3.3 | 1 |
| Short – Long | 492 | 0.51 | 0.446 | 0.39 | 3.2 | 1.12 |
| Weak – Strong | 492 | 0.31 | 0.272 | 0.187 | 2.9 | 1.18 |
| Pleasant – Clicky | 492 | 0.66 | 0.644 | 0.581 | 3.1 | 1.2 |
| Fast – Slow | 492 | 0.63 | 0.597 | 0.523 | 2.9 | 1.11 |
| Soft – Hard | 492 | 0.33 | 0.271 | 0.196 | 3.2 | 1.2 |
| Mechanical – Buzzy | 492 | 0.14 | 0.007 | -0 | 3.4 | 0.92 |
| Sharp – Dull | 492 | 0.58 | 0.538 | 0.469 | 2.7 | 0.91 |
| High Frequency – Low Frequency | 492 | 0.51 | 0.441 | 0.383 | 2.9 | 0.81 |
| Dynamic – Tranquil | 492 | 0.76 | 0.767 | 0.683 | 2.9 | 1.13 |
| Good – Bad | 492 | 0.78 | 0.797 | 0.727 | 3.1 | 1.21 |

## i. McDonald's omega

| Omega with four factors | |
|---|---|
| Alpha: | 0.78 |
| G.6: | 0.84 |
| Omega Hierarchical: | 0.56 |
| Omega H asymptotic: | 0.64 |
| Omega Total | 0.87 |

| Schmid Leiman Factor loadings greater than 0.2 | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | g | F1* | F2* | F3* | F4* | h2 | u2 | p2 |
| Feedback – No Feedback | 0.32 | 0.59 | | | | 0.5 | 0.5 | 0.21 |
| Responsive – Sluggish | 0.5 | -0.29 | 0.28 | | 0.35 | 0.54 | 0.46 | 0.43 |
| Predominant – Faint | | 0.78 | | | | 0.62 | 0.38 | 0.01 |
| Smooth – Rough | 0.42 | | 0.34 | | 0.25 | 0.36 | 0.64 | 0.48 |
| Short – Long | 0.45 | | | | 0.52 | 0.47 | 0.53 | 0.42 |
| Weak – Strong | | 0.81 | | | | 0.67 | 0.33 | 0.01 |
| Pleasant – Clicky | 0.51 | | 0.62 | | | 0.66 | 0.34 | 0.39 |
| Fast – Slow | 0.55 | | | 0.35 | 0.28 | 0.51 | 0.49 | 0.59 |
| Soft – Hard | 0.29 | -0.45 | 0.44 | | | 0.48 | 0.52 | 0.17 |
| Mechanical – Buzzy | | | | | | 0.02 | 0.98 | 0.02 |
| Sharp – Dull | 0.5 | | | 0.48 | | 0.48 | 0.52 | 0.51 |
| High Frequency – Low Frequency | 0.39 | | | 0.48 | | 0.42 | 0.58 | 0.38 |
| Dynamic – Tranquil | 0.67 | | | 0.44 | | 0.67 | 0.33 | 0.66 |
| Good – Bad | 0.59 | 0.38 | 0.45 | | | 0.73 | 0.27 | 0.47 |

| With eigenvalues of: | | | | |
|---|---|---|---|---|
| g | F1* | F2* | F3* | F4* |
| 2.58 | 2.09 | 1.03 | 0.82 | 0.62 |

General/max 1.23 max/min = 3.4
Mean percent general = 0.34 with sd = 0.22 and cv of 0.64
Explained Common Variance of the general factor = 0.36

The degrees of freedom are 41 and the fit is 0.17
The number of observations was 492 with Chi Square = 83.02 with prob < 0.00011
The root mean square of the residuals is 0.02
The df corrected root mean square of the residuals is 0.03
RMSEA index = 0.046 and the 90 % confidence intervals are 0.031 0.06
BIC = -171.12

Compare this with the adequacy of just a general factor and no group factors
The degrees of freedom for just the general factor are 77 and the fit is 2.66
The number of observations was 492 with Chi Square = 1290.45 with prob < 1.8e-219
The root mean square of the residuals is 0.17
The df corrected root mean square of the residuals is 0.19
RMSEA index = 0.18 and the 90 % confidence intervals are 0.17 and 0.188
BIC = 813.17

| Measures of factor score adequacy | | | | | |
|---|---|---|---|---|---|
| | g | F1* | F2* | F3* | F4* |
| Correlation of scores with factors | 0.79 | 0.93 | 0.76 | 0.65 | 0.66 |
| Multiple R square of scores with factors | 0.62 | 0.86 | 0.58 | 0.42 | 0.44 |
| Minimum correlation of factor score estimates | 0.24 | 0.72 | 0.16 | -0.15 | -0.13 |

| Total. General and Subset omega for each subset | | | | | |
|---|---|---|---|---|---|
| | g | F1* | F2* | F3* | F4* |
| Omega total for total scores and subscales | 0.87 | 0.49 | 0.73 | 0.79 | 0.58 |
| Omega general for total scores and subscales | 0.56 | 0.1 | 0.39 | 0.47 | 0.32 |
| Omega group for total scores and subscales | 0.16 | 0.39 | 0.34 | 0.32 | 0.26 |

## j. Promax rotation with five factors

| Principal Component Analysis with 5 Factors (according to empirical BIC value) and promax rotation | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | item | PC2 | PC3 | PC1 | PC5 | PC4 | h2 | u2 |
| Predominant – Faint | 3 | 0.87 | | | | | 0.72 | 0.28 |
| Weak – Strong | 6 | 0.87 | | | | | 0.75 | 0.25 |
| Feedback – No Feedback | 1 | 0.75 | | | | | 0.63 | 0.37 |
| Good – Bad | 14 | 0.52 | 0.51 | | | | 0.77 | 0.23 |
| Soft – Hard | 9 | -0.42 | 0.85 | | | | 0.71 | 0.29 |
| Pleasant – Clicky | 7 | | 0.83 | | | | 0.73 | 0.27 |
| Smooth – Rough | 4 | | 0.61 | | 0.31 | | 0.55 | 0.45 |
| Responsive – Sluggish | 2 | | 0.57 | | 0.34 | | 0.67 | 0.33 |
| High Frequency – Low Frequency | 12 | | | 0.97 | -0.37 | | 0.74 | 0.26 |
| Sharp – Dull | 11 | | | 0.75 | | | 0.64 | 0.36 |
| Dynamic – Tranquil | 13 | | | 0.55 | | | 0.71 | 0.29 |
| Short – Long | 5 | | | | 0.97 | | 0.77 | 0.23 |
| Fast – Slow | 8 | | | 0.46 | 0.54 | | 0.66 | 0.34 |
| Mechanical – Buzzy | 10 | | | | | 1 | 0.98 | 0.02 |

| | PC2 | PC3 | PC1 | PC5 | PC4 |
|---|---|---|---|---|---|
| SS loadings | 2.66 | 2.49 | 2.13 | 1.73 | 1.03 |
| Proportion Var | 0.19 | 0.18 | 0.15 | 0.12 | 0.07 |
| Cumulative Var | 0.19 | 0.37 | 0.52 | 0.64 | 0.72 |
| Proportion Explained | 0.26 | 0.25 | 0.21 | 0.17 | 0.1 |
| Cumulative Proportion | 0.26 | 0.51 | 0.73 | 0.9 | 1 |

| With component correlations of | | | | | |
|---|---|---|---|---|---|
| | PC2 | PC3 | PC1 | PC5 | PC4 |
| PC2 | 1 | 0.13 | 0.28 | 0.13 | 0.15 |
| PC3 | 0.13 | 1 | 0.4 | 0.5 | -0.05 |
| PC1 | 0.28 | 0.4 | 1 | 0.48 | 0.04 |
| PC5 | 0.13 | 0.5 | 0.48 | 1 | 0.02 |
| PC4 | 0.15 | -0.05 | 0.04 | 0.02 | 1 |

Test of the hypothesis that 5 components are sufficient.
The degrees of freedom for the null model are 91 and the objective function was 5.15
The degrees of freedom for the model are 31 and the objective function was 1.03
Fit based upon off diagonal values = 0.96

### k. Interitem Correlations

| | Feedback - No Feedback | Responsive - Sluggish | Predominant - Faint | Smooth - Rough | Short - Long | Weak - Strong | Pleasant - Clicky | Fast - Slow | Soft - Hard | Mechanical - Buzzy | Sharp - Dull | High Frequency - Low frequency | Dynamic - Tranquil | Good - Bad |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Feedback - No Feedback | 1 | | | | | | | | | | | | | |
| Responsive - Sluggish | 0.06 | 1 | | | | | | | | | | | | |
| Predominant - Faint | 0.49 | -0.22 | 1 | | | | | | | | | | | |
| Smooth - Rough | 0.13 | 0.43 | 0.01 | 1 | | | | | | | | | | |
| Short - Long | 0.14 | 0.41 | -0.06 | 0.32 | 1 | | | | | | | | | |
| Weak - Strong | 0.49 | -0.23 | 0.67 | -0.06 | -0.06 | 1 | | | | | | | | |
| Pleasant - Clicky | 0.35 | 0.36 | 0.16 | 0.41 | 0.21 | 0.13 | 1 | | | | | | | |
| Fast - Slow | 0.23 | 0.37 | 0.07 | 0.25 | 0.44 | 0.04 | 0.26 | 1 | | | | | | |
| Soft - Hard | -0.06 | 0.54 | -0.28 | 0.3 | 0.15 | -0.36 | 0.35 | 0.12 | 1 | | | | | |
| Mechanical - Buzzy | 0.02 | -0.08 | 0.1 | -0.04 | -0.02 | 0.09 | 0.04 | 0.01 | -0.07 | 1 | | | | |
| Sharp - Dull | 0.23 | 0.27 | 0.04 | 0.16 | 0.29 | 0.1 | 0.23 | 0.47 | 0.09 | 0.02 | 1 | | | |
| High Frequency - Low frequency | 0.12 | 0.17 | 0.07 | 0.13 | 0.12 | 0.1 | 0.24 | 0.35 | 0.1 | -0.09 | 0.39 | 1 | | |
| Dynamic - Tranquil | 0.28 | 0.43 | 0.12 | 0.37 | 0.4 | 0.12 | 0.43 | 0.54 | 0.2 | -0.03 | 0.55 | 0.44 | 1 | |
| Good - Bad | 0.51 | 0.29 | 0.37 | 0.4 | 0.26 | 0.38 | 0.65 | 0.4 | 0.15 | 0.04 | 0.35 | 0.33 | 0.56 | 1 |

### *d) Final Instrument*

Below are the final versions of the developed instrument in German and English. The items marked with an asterisk (*) need to be recoded for the final analysis. There is some extra space left above if it is aimed at developming new items with the Repertory Grid technique. It is recommended to use triads as setting presentation to the subjects for the technology under investigation. Instructions are provided on top of the questionnaire, but may be adapted depending on setting or technology.

### a. German Version

Im Folgenden werden Sie gebeten, den Schalter anhand der mit Ihnen entwickelten und einigen zusätzlichen Items zu bewerten. Bitte probieren Sie den Schalter dafür mehrfach aus. Variieren Sie dabei bitte die Geschwindigkeit und den Druck, mit dem Sie drücken. Versuchen Sie, so spontan wie möglich zu antworten. Dabei gibt es kein „richtig“ oder „falsch“- es kommt auf Ihre persönliche Meinung an! Der Schalter ist…

| | ***sehr*** | ***ein wenig*** | ***weder-noch*** | ***ein wenig*** | ***sehr*** | |
|---|---|---|---|---|---|---|
| | **O** | O | O | O | **O** | |
| Feedback gebend | **O** | O | O | O | **O** | nicht Feedback gebend |
| schwergängig | **O** | O | O | O | **O** | leichtgängig* |
| dominant | **O** | O | O | O | **O** | kraftlos |
| rauh | **O** | O | O | O | **O** | gleichmäßig* |
| kurz | **O** | O | O | O | **O** | lang |
| stark | **O** | O | O | O | **O** | schwach* |
| angenehm | **O** | O | O | O | **O** | klickig |
| langsam | **O** | O | O | O | **O** | schnell* |
| weich | **O** | O | O | O | **O** | hart |
| summend/ vibrierend | **O** | O | O | O | **O** | mechanisch* |
| scharf | **O** | O | O | O | **O** | stumpf |
| niedrigfrequent | **O** | O | O | O | **O** | hochfrequent* |
| dynamisch | **O** | O | O | O | **O** | träge |
| schlecht | **O** | O | O | O | **O** | gut* |

**Vielen Dank für Ihre Teilnahme!**

**b. English Version**

Now you'll be kindly asked to rate the switch on the constructed and some previously developed items. Please press the switch several times and vary speed and pressure during this. Try to answer as spontaneously as possible. There is no "right" or "wrong" about this- your personal impression is what matters to us.

The switch is…

| | *Very much* | *A little* | *Neither* | *A little* | *Very much* | |
|---|---|---|---|---|---|---|
| | **O** | O | O | O | **O** | |
| Providing Feedback | **O** | O | O | O | **O** | Providing no Feedback |
| Sluggish | **O** | O | O | O | **O** | Responsive* |
| Predominant | **O** | O | O | O | **O** | Faint |
| Rough | **O** | O | O | O | **O** | Smooth* |
| Short | **O** | O | O | O | **O** | Long |
| Strong | **O** | O | O | O | **O** | Weak* |
| Pleasant | **O** | O | O | O | **O** | Clicky |
| Slow | **O** | O | O | O | **O** | Fast* |
| Soft | **O** | O | O | O | **O** | Hard |
| Buzzy | **O** | O | O | O | **O** | Mechanical* |
| Sharp | **O** | O | O | O | **O** | Dull |
| Low Frequency | **O** | O | O | O | **O** | High Frequency* |
| Dynamic | **O** | O | O | O | **O** | Tranquil |
| Bad | **O** | O | O | O | **O** | Good* |

**Thank you very much for your participation!**