Grant Agreement N°101092861

Making the Invisible Visible for Off-Highway Machinery by Conveying Extended Reality Technologies

DELIVERABLE 4.2 – INTERACTION MODALITY ASSIGNMENT

This work is part of the THEIA[XR] project that received funding from the European Union's Horizon Europe research and innovation program under grant agreement no. 101092861

Document Identification

Project	THEIA^{XR}
Project Full Title	Making the Invisible Visible for Off-Highway Machinery by Conveying Extended Reality Technologies
Project Number	101092861
Starting Date	January 1st, 2023
Duration	3 years
Horizon Europe Programme	Horizon Europe Framework Program (HORIZON)
Topic	HORIZON-CL4-2022-HUMAN-01-14: “eXtended Reality Technologies”
Call for proposal	HORIZON-CL4-2022-HUMAN-01
Type of Action	RIA-Research & Innovation Action
Website	www.theia-xr.eu
Work Package	WP4
WP Leader	TUD
Responsible Partner(s)	TUD
File Name	D 4.2 – Interaction Modality Assignment
Contractual delivery date	30.09.2023
Actual delivery date	29.09.2023
Version	1
Status	Final
Dissemination level	Public
Author(s)	Sebastian Lorenz (TUD), Volker Waurich (TUD)
Contact details of the coordinator	Martijn Rooker, martijn.rooker@tttech.com
Internal Reviewer	Kaj Helin (VTT) Pekka Yli-Paunu (KAL)

Disclaimer

Funded by the European Union	Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European union or the European Commission. Neither the European Union nor the granting authority can be held responsible for them.

Executive summary

Deliverable 4.2 provides the interaction modality framework with criteria-based recommendations for information presentation. It is rather a methodology for the development of multimodal user interfaces, but at the same time, it offers the question of how modern Human-Machine Interfaces (HMI) can be improved. It discusses different interaction modalities and synthesizes the results into work materials guiding the decision-making process of appropriate information flow between the operator and system. In doing so, work packages 5 and 6 can benefit from the results. It includes a short analysis of general interaction models and modalities in literature. D4.2 evaluates their respective capabilities, advantages, disadvantages, and operator-specific requirements. The result is a methodology for a creative yet systematic ideation of multimodal interaction. For this purpose, the necessary working materials were developed. It consists of a template for the description of operating challenges, inspiring modality cards, two handbooks and solution templates for a comparable variant documentation. Moreover, the use of an interlinked process was described. With the attached detail cards, the focus is on promoting multimodal thinking in the field of HMI elements and their use for off-highway commercial vehicles and machines.

Table of Contents

Document Identification 2
Executive summary 4
1 Introduction 7
1.1 Design Principles 7
1.2 Interaction Vocabulary 8
2 Understanding the Basics of Interactions 9
2.1 HMI Models in Literature 9
2.2 Summary of modalities 11
3 Analysing criteria for designing multimodal interactions 13
4 Framework Concepts 15
4.1 Drafts and approaches of the design tool 15
4.2 Final draft of the tool components 16
5 Results 21
5.1 Introducing the design toolkit 21
5.2 Brief description how to use the toolkit 22
6 Discussion 24
6.1 General 24
6.2 Limitations 24
7 Further Work 25
References 26

List of Figures

Figure 1: The Seven Stages of Action according to Norman Cycle Model [1] (Left), The Information model according to Gitt Syntax, Semantics, Pragmatics and Apobetics [6] (Right) 9
Figure 2: Interaction circuit according to Dumas [7] 10
Figure 3: Human sensory modalities and possibilities for input 12
Figure 4: Concept of Evaluation Charts 15
Figure 5: Draft of the Digital Configurator as a modular tool for describing and evaluating multimodal interactions 15
Figure 6: Layout and design of a Modality Card 16
Figure 7: Final solution sheet, handbook and task definition template of the design tool 17
Figure 8: Synergy of the developed design tool components 20
Figure 9: Developed toolkit for multimodal interaction design | Combining interaction challenge, modality cards, handbook I+II and solution sheets with detail cards 21
Figure 10: Result of the modality card set, the visual modality "text" and the auditory modality “sound” .. 21
Figure 11: Colour scheme of the modality cards 22

List of Tables

Table 1: Translated interaction vocabulary according to Diefenbach[5] 8
Table 2: Summary of the key characteristics of the most important feedback channels 11
Table 3: Full list of criteria for designing multimodal interactions.. 14

1 Introduction

1.1 Design Principles

Designing multimodal systems is challenging. On the one hand typical design choices and intuitions from standard computing environments may not adapt to multimodal contexts. On the other hand, the intricate interplay of available modalities, specific application tasks, and user constraints often necessitates unique design approaches [4]. Nevertheless, established design guidelines and approaches help and provide the starting point for designing a supporting framework for the interaction modality assignment.
Considering general design guidelines, there are numerous forms and expressions according to their founders. Steve Jobs or Dieter Rams shall be mentioned only briefly by name. Kim Goodwin is one of these people, but her work Designing for The Digital Age is of particular importance, which bundles all her experience and expertise about the relevant facets that have to be taken into human-centred product and service development. At the same time, it focuses on the digital age and not on physical products. It sets out four requirements for good design[2]:

> **Good design is ethical.** Firstly, it highlights the importance of privacy and data protection, ensuring transparency and protecting sensitive information. Secondly, the principle underscores the need for inclusivity and responsibility, providing accessibility and social well-being to all users, regardless of their abilities. Lastly, sustainability and accountability are integral components of ethical design. Sustainable design choices in terms of hardware and energy consumption contribute to reducing the environmental footprint. By adhering to these ethical principles, designers can develop technologies that are not only innovative for more efficient and safer task completion from an economic perspective but also socially responsible and environmentally conscious. That brings us to the key aspects of the following principles.
>
> **Good design is purposeful and pragmatic.** A product should be designed to perform the task in an optimized way. The user-centred and problem-solving approach prioritizes practicality, efficiency, and adaptability. It encourages designers to focus on delivering solutions that effectively address real-world needs rather than getting caught up in theoretical or ornamental aspects of design. Concerning the HMI this principle strives for the best usability of the product.
>
> **Good design is elegant.** An elegant solution is usually the simplest and a complete solution that satisfies all the necessary needs but from which nothing can be further removed. They are perceived as lean or efficient. Due to their simplicity and consistency, they tend to maximize the user satisfaction.

The UX Design Institute guidelines describe similar principles with a total of seven core points, adding concrete aspects for the user experience of HMI. One of the important principles is to provide the user a right amount of control where they can interact with the product efficiently. A balanced degree of control and freedom has positive influences on the UX. At this point, the existing standard (ISO 9241-110:2020-05) is also mentioned, which specifies the design of user interfaces based on task appropriateness, self-descriptiveness, conformity to expectations, learnability, controllability, robustness to user error, and adaptability in customized usage.
There are also principles for multimodal design in the cognitive science literature on intersensory perception and intermodal coordination. These principles direct the design of information presented to users or describe the integration of multiple modalities and supporting multiple user inputs. Reeves provides a brief summary of some general guiding principles essential to the design of effective multimodal interaction [3]:

Maximize human cognitive and physical abilities. *Designers need to determine how to support intuitive, streamlined interactions based on users' human information processing abilities (including attention, working memory, and decision making) for example:*

Avoid unnecessarily presenting information *in two different modalities in cases where the user must simultaneously attend to both sources to comprehend the material being presented; such redundancy can increase cognitive load.*

Maximize the advantages of each modality *to reduce user's memory load in certain tasks and situations.*
System visual presentation *coupled with user manual input for spatial information and parallel processing.*

Integrate modalities in a manner compatible with user preferences*, context, and system functionality. Additional modalities should be added to the system only if they improve satisfaction, efficiency, or other aspects of performance for a given user and context.*

Modalities with very different characteristics – e.g., speech and eye gaze or facial expression and haptics input – may not have obvious points of similarity and straight forward ways to connect. Turk [4] concludes that perhaps the most challenging aspect is the temporal dimension. Different modalities may have different temporal constraints and different signal and semantic endurance. Some modalities provide information at sparse, discrete points in time while others generate continuous but less time-specific output. Some modal combinations are intended to be interpreted in parallel, which others may typically be offered sequentially.

These dependencies are the basis of the framework to be developed and must be taken into account when designing multimodal interactions.

1.2 Interaction Vocabulary

Simply describing interaction is difficult. Evaluating which interaction variant is the best for a given information transfer would offer great added value in the process. Sarah Diefenbach and colleagues also asked themselves this question. They developed the "Interaction Vocabulary, a set of dimensions to describe and distinguish forms of interaction for use in the design and evaluation phases." Eleven dimensions were established, with each dimension described by its two extremes. For example, one of these polarity fields describes the speed of information exchange from slow to fast [5]. Application of this tool can now not only be found in the evaluation of interaction forms, but it can already support development. By breaking down a task into its elementary actions and reactions, the designer can exploratively generate different variants. Diefenbach concludes that "Interaction [is] often characterized only by modality or even a technology, but not by its aesthetic quality. [These] dimensions offer a way to describe interactions now in a technology-free way, and also a vocabulary that usability professionals can use in their daily work" [5].

slow	*fast*
stepped	*stepless*
stable	*changing*
approximate	*precise*
mediated	*direct*
hidden	*obvious*
cautious	*powerful*
spatially separation	*spatial proximity*
undemanding	*in need of attention*
immediate	*delayed*
casual	*targeted*

Table 1: Translated interaction vocabulary according to Diefenbach[5]

2 Understanding the Basics of Interactions

2.1 HMI Models in Literature

The Seven Stages of Action cycle model, as shown in Figure 1, will serve as a starting point [1]. The execution and evaluation of an action are connected via the environment of the interaction (the World) and the goal of the interaction. Norman sees the role of the designer in two points: The reduction in evaluation time or evaluation complexity and in the use of an understandable conceptual model for operating concepts. This refers to the interaction concept, which should help people understand the interaction itself, for example, with the help of analogies. The quality of the conceptual model is a first indication of the degree of intuitiveness of the interaction. The reduction of the evaluation time can be realized by appropriate feedback that supports the desired interaction even during the execution stages.

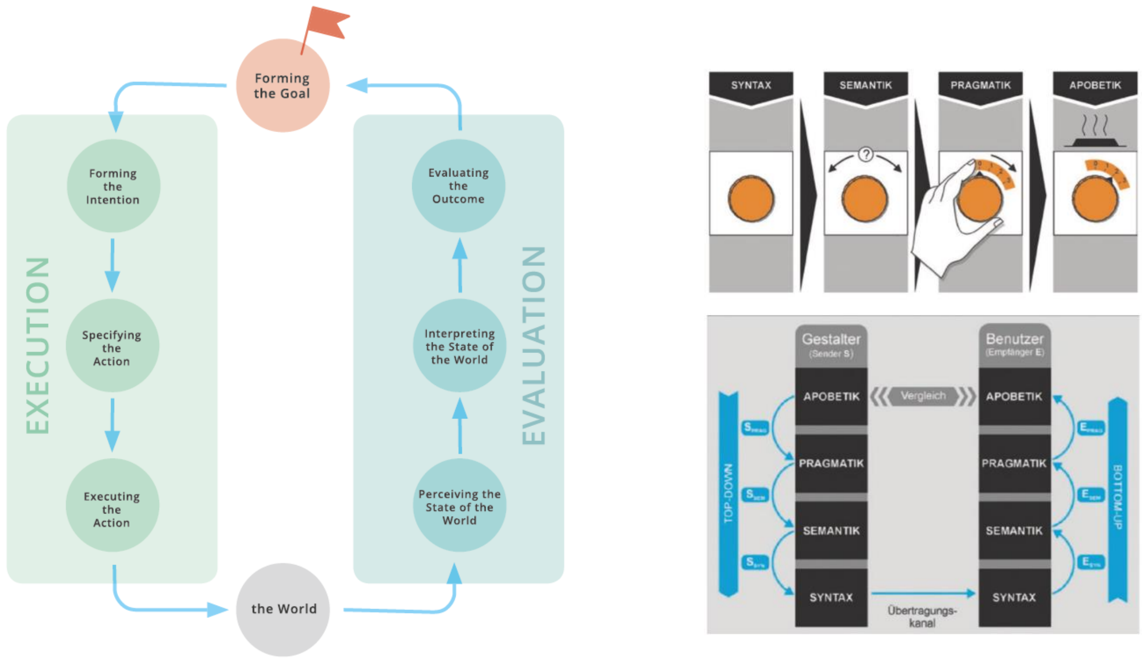

Figure 1: The Seven Stages of Action according to Norman Cycle Model [1] (Left), The Information model according to Gitt Syntax, Semantics, Pragmatics and Apobetics [6] (Right)

The starting point of the model is either the formulation of the goal or the environment. The model is strongly simplified, since here no temporal components are integrated and also no iteration loops are indicated. Most of the seven stages occur subconsciously. Actions that seem difficult at first and require conscious decisions, such as riding a bicycle, become automated after some practice and are thus performed subconsciously. These insights should also be incorporated into interaction design considerations. Thus, certain aspects may require deliberate conscious decisions and be less intuitive in design than aspects that are automated. To describe this Norman defines three levels of processing of human cognition and emotion.

- **The visceral level** includes reflexes and protective instincts, decisions that are made quickly without a detour via logic.
- **The behavioural level** also covers subconscious processes and abilities, but which are learned such as speaking or riding a bike.
- On **the reflective level** humans react slowly, deliberately and consciously. Evaluations and causal relationships are processed or higher emotions such as guilt or responsibility take place here.

In comparison to Norman's model, Schmid and Maier give an information model according to Gitt including the design considerations of the designer [6]. It connects the sender and receiver via the levels of apobetics, pragmatics, semantics and syntax. While the designer starts from a goal and thinks about how the interaction should look like, the user only sees the syntactic level. From there, in the best case, the same goal is reached that was specified by the designer in the apobetics.

These models describe interactions quite generally and do not incorporate the different channels of perception or modes of executive action. Dumas et al. summarize the steps of an interaction circuit with modalities by connecting several possible items of the human perception and action with perception and action technologies (see Figure 1) [7].

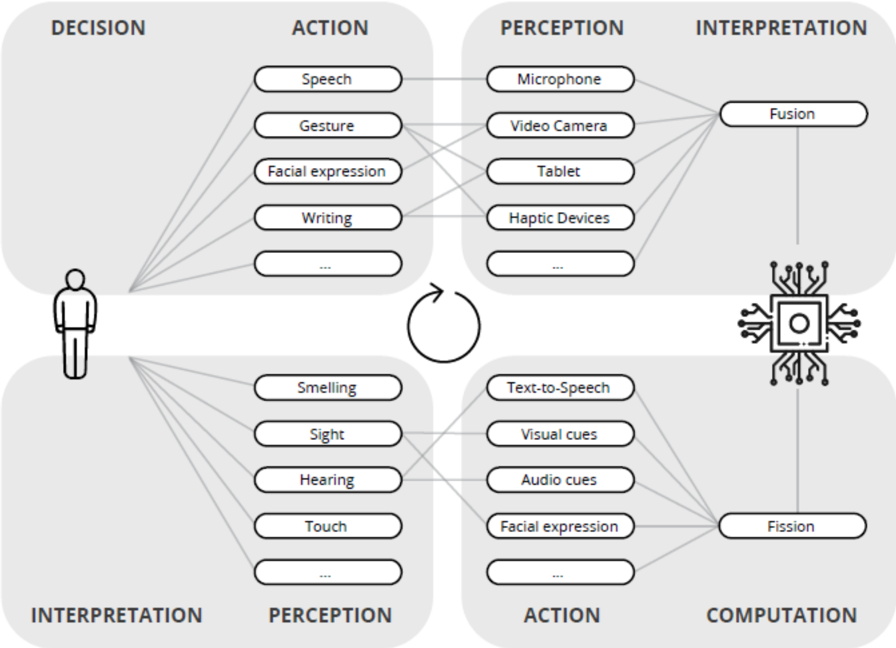

Figure 2: Interaction circuit according to Dumas [7]

One advantage of this description of multimodal interactions is a possible identification of resource conflicts, where complex information exchanges take place simultaneously over the same communication channel. The growing number of information sources poses ever greater challenges for stimulus filtering. As a result, information is absorbed more and more superficially, which leads to less knowledge formation. There are various models about theories of attention that explain the allocation of capacity to the various human work tasks.

A distinction is made between the sequential models (time based) and the theory of limited resources (capacity based). Wickens' multiple resources model is more often used in the field of HMI. It states, that different resources exist along three dimensions for information processing spanning a matrix (or cube): Encoding, modalities, and feedback. To use the model, Wickens makes two assumptions:

1. For each cell of the matrix, there are available resources that can be accessed independently.
2. Each of these resources is limited. For example, according to this theory, it is easier to make linguistic inputs and respond to auditory stimuli while driving a car than it is to make parallel visual observations and complete motor inputs. [14]

With these models including the resource conflict view, only assumptions can be made about the mental demands of an HMI technology. In order to make valid statements, investigations must be carried out with the explicit technology and its design - for example, by recording the performance in a parallel second task, measuring psychophysiological characteristics such as ECG, respiration, EEG, pupil change, eyelid blink or the subjective stress assessment.

At this point it should be added that the available capacity for information processing in the operator does not only depend on the resource or modality usage. The emotions of the user have a great influence. The so-called motivational-emotional subsystem also determines the available resources. When people are in a positive mood, they tend to process information intuitively and holistically, while negative emotions tend to process information analytically and sequentially. The capacity required for information processing is significantly lower when people are in a positive mood than when they are in a negative mood. Creative solutions to problems are also less likely to be found when people are in a bad mood. People react emotionally to the design of user interfaces. But what are the general ways of sharing information (modalities)?

2.2 Summary of modalities

State variables of the environment or within the body are determined with the help of the receptors of the sensory channels, also called modalities. Temperature, sound waves or even light waves are among the external "stimuli" that enter perception through sensory channels. Objective sensory physiology investigates the mechanisms and processes in the human body. However, the subjective impressions of the users are of interest for the design. Chen and Chuang, for example, created a systematics of tactile vocabulary that can be used to describe surfaces [15]. The qualities of each channel depend on the technological design. For example, there are different ways to simulate click sensations on trackpads.

In the following section, only a general overview of the possibilities for perception but also for input of information will be given. According to Klinke, a distinction is made between the visual system, the auditory system, the vestibular system, the somatosensory system and the chemical senses (smell and taste) [16]. The classification varies according to the authors - primarily important for HMI are the visual system, the auditory system and the "haptic" sense. Table 2 and Figure 3 summarize the research on the classification of sensory perception.

Modality	Visual	Auditory	Haptic
Response time	200-400 ms	100-150 ms	80-150 ms
Directionality (near or far)	Narrow field of sharp vision Overlooking of important information possible	With a clear signal, the perception is not bound to a focus	Perception is tied to physical contact
Complexity of the Information transmitted	Very large number of easily distinguishable symbols suitable for graphical and textual information	limited number of easily distinguishable symbols unsuitable for graphical information	Only information with low complexity can be transmitted Only modality for the acquisition of mechanical object properties
Selectivity	User can select from total information user decides on time of perception high sensitivity for movement/change	difficult to select consciously Perception only possible on-line	It depends on the body parts involved.

Table 2: Summary of the key characteristics of the most important feedback channels

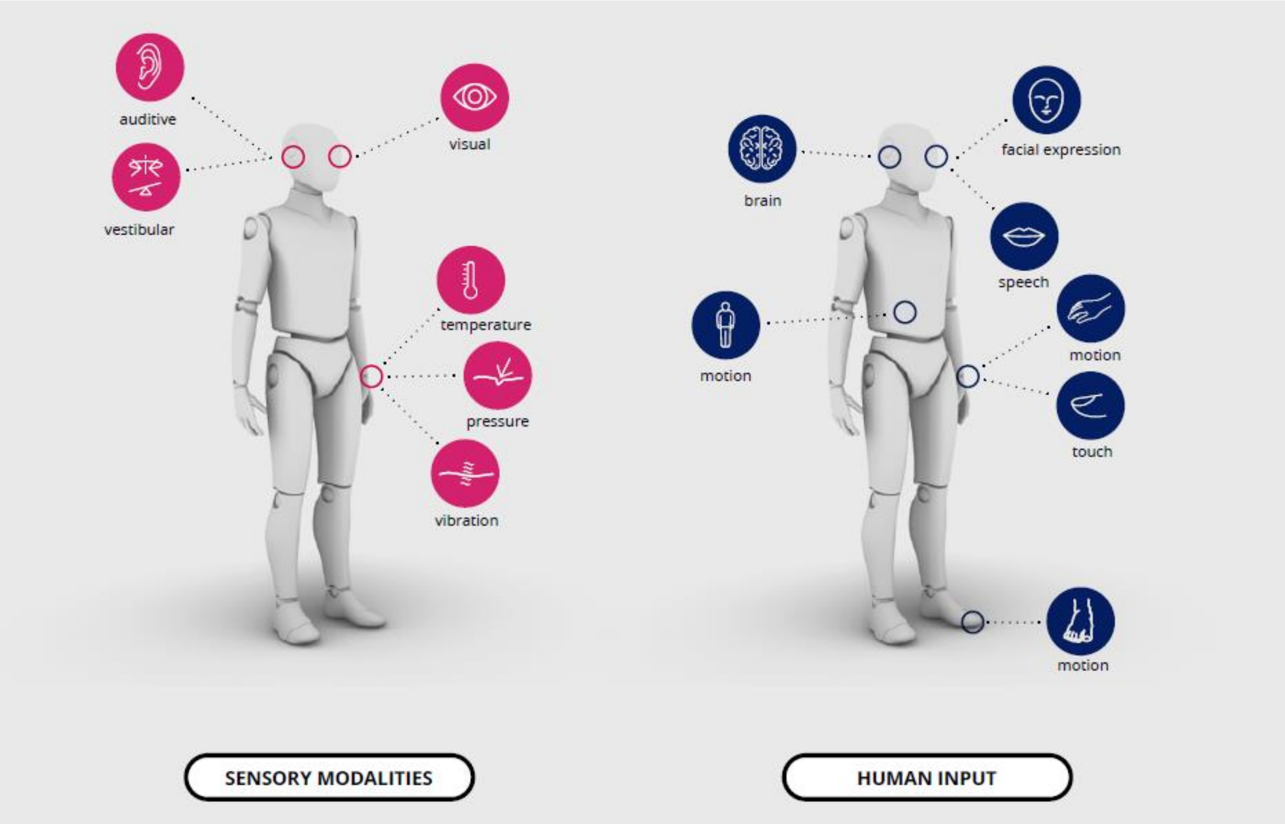

Figure 3: Human sensory modalities and possibilities for input

3 Analysing criteria for designing multimodal interactions

Criteria based on literature are listed by authors in Table 3. Furthermore, own criteria, which deduce from section 1 and 2, were added. The collected criteria still differ quite strongly in their degree of abstraction and their thematic assignment. In order to provide a better overview, the criteria were divided into topics. These topics were defined after a comparison of all collected criteria as described below.

A topic includes technological "System properties" - in the sense of: Does the specific technology have more or less potential to enable adaptive interaction? What is the level of technical complexity compared to other technology categories? However, "Operating properties" can only be well assessed when there is a concrete idea of the overall interaction process. Properties of "interaction aesthetics" are interesting on a design level and usually concern more detailed design. Other topics include the meta-level, context, concrete ergonomics properties, and feedback properties. The catalogue is quite open, which means that more and more differentiated areas can certainly be defined.

Source	Criteria	Topic
DIN EN ISO 9241-420	Cleaning, reliability, environmental interactions, light irradiation, electromagnetism, dust, vibration, temperature, humidity, user, work tasks, environmental work equipment (hardware, software and materials), physical and social environment	System in relation to the environment / context
	posture of the user, position of the device, weight, actuation force, ambidextrous use, body parts involved	Ergonomics
	operating effort, predictability, effectiveness, efficiency	Operating properties
	reliability, charging function, energy consumption, portability	System properties
DIN EN ISO 9241-10 (2011)	task appropriateness, self-descriptiveness, conformity to expectations conformity, error tolerance, controllability, adaptability in customized usage, learning conduciveness	Operating properties
Lenz et al. 2014 [9]	connectivity, input modalities, tasking, output modalities, versatility, external connections, body parts involved, combination/number of touchpoints, number of participants, including objects	Meta
	adaptability, robustness, dependency, feedback, freedom of interaction, initiative, sequence, presence	System properties
Thüring & Mahlke 2007 [8][10]	learnability, effectiveness, controllability	Operating properties
Löwgren 2009 [10]	pliability, rhythm, dramaturgical structure, fluency	Interaction Aesthetics
Janlert & Stoltermann 2018 [13]	richness, interactability, pace/time	Interaction Aesthetics
	agency, awareness level, attention demand, freedom of action, predictability, receptivity, precision	System properties
	complexity, control, functionality	Operating properties
Damian 2017 [12]	level of Detail, scope, duration, prominence	Feedback

Csikszentmihalyi (2008) [11]	Feedback, immersion level in activity, attention level, interference, effortlessness, control	Operating properties
Own Analysis	Mobility, customizability, flexibility, extensibility, robustness, complexity (technical), functional reliability, adaptability, privacy	System properties
	Intuitiveness/naturalness, complexity for the user, learning effort, adaptability, attention, empathy, user fatigue, user motivation, process assistance	Operating properties
	Multiple Users, Multiple Platforms	Context
Own Research	Near/far, conscious, subconscious, locative, multimodal	Feedback
	level of automation	Operating properties

Table 3: Full list of criteria for designing multimodal interactions

The list illustrates the numerous possibilities of an evaluation of interactions. Moreover, the assessment is highly context-dependent and cannot be generalized. Hence, a framework can only take individual aspects into account. During the following elaboration, specific criteria were selected and tested for suitability and effectiveness in expert panels.

4 Framework Concepts

4.1 Drafts and approaches of the design tool

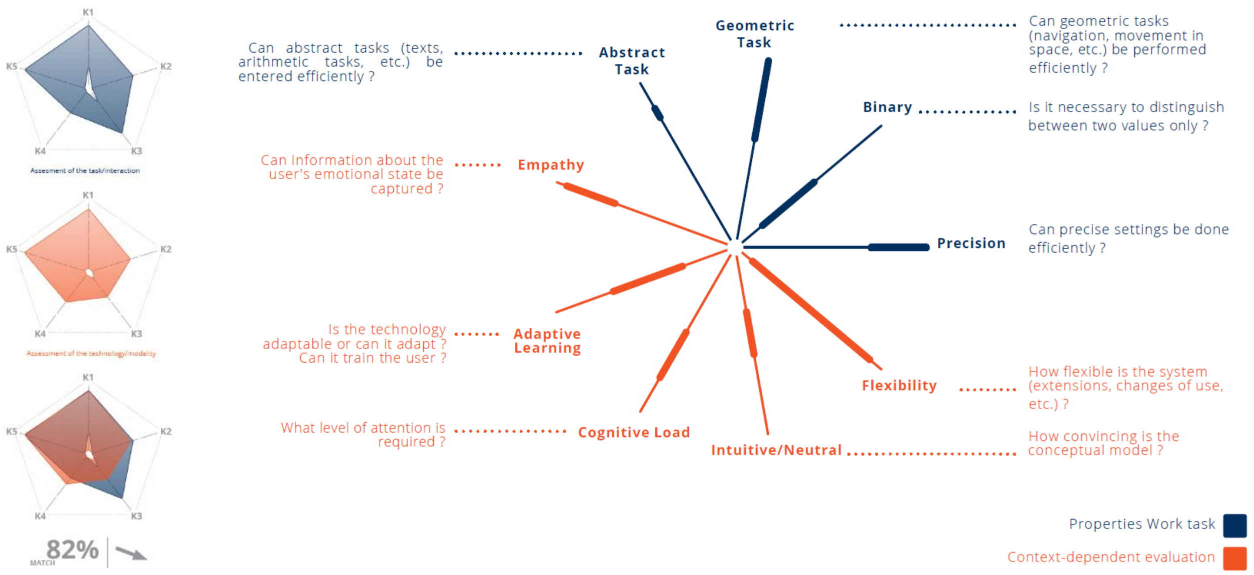

Figure 4: Concept of Evaluation Charts

Three concept ideas were described: Evaluation Charts, Modality Matrices, and Information Characterization. The concept of Evaluation Charts synthesizes the criteria resulting from the analysis into guiding questions, which not only help to describe the task and interaction (context), but at the same time set the requirements for the technology and modality. The evaluation is done by superimposing radar charts and their match. Figure 4 shows the draft of possible questions. The modality matrix simply lists all possible modalities to the essential information and functions in the identified use cases of D4.1. In contrast, the concept of information characterization outlined a generally valid description of the possible signals (information), which allows a more general evaluation and assignment of modality and information. It quickly became apparent that the individual approaches alone would be too complex in making or not produce manageable results. The Digital Configurator and the Digital Taxonomy combine the ideas into a modular system, with the help of which the developer can build himself complex interaction concepts and get feedback on the suitability of the modality choice (see Figure 5).

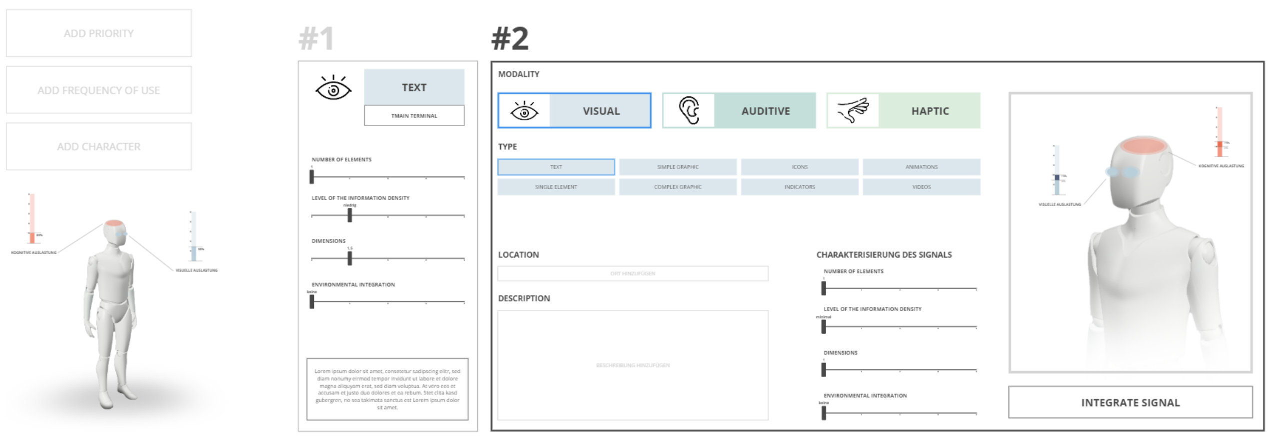

Figure 5: Draft of the Digital Configurator as a modular tool for describing and evaluating multimodal interactions

The design of the actual digital tools was transferred into a physical work materials and tested in a workshop with two HMI experts. For this purpose, concrete design tasks representing typical HMI developments were defined, which were to be processed with the help of the tool.

With the results of the workshop the weaknesses and possibilities of the developed draft became clear. Using the sheets shows that an accompanying modality overview is missing, i.e. the possibility to consult the modalities and to choose a suitable one. The existing work material used was a great help, but in some places it led to an information overload - the test persons were not able to efficiently process the amount of information in a result-oriented manner. In the debriefing of the workshop, some key points were collected, which can be applied to the entire process chain, regardless of the individual modality. Ultimately, the question was asked whether there was a possibility of a digital connection, how the design could be raised to a digital level and whether this would offer advantages. These four major areas triggered the elaboration and further development of the work.

4.2 Final draft of the tool components

Modality Cards: There is a need to make individual modalities present to the designer. Whether it is just to create a handy overview or to actually develop a tool to collaboratively design in a workspace. The card content was designed based on the existing information and the needs of the test workshop held. The front side of the card should carry the sensory level (visual, auditory or haptic) with corresponding modality (for example text). A description and examples should complement the front side, whereby the latter as a visualization gives a first idea of what the modality represents and provides a visual differentiation of the cards. On the reverse side, further information is provided: The representing medium and its requirements, advantages as well as disadvantages complete the card.

The advantages of these cards are seen in the simplicity and speed of use. The modality cards are self-explanatory and uncomplicated. For their use no technology is necessary, one can use the cards immediately. Particularly the caring adapted character of the cards, thus the situation-dependent use of the set is considered as very positive.

It must be noted that the application reference to a concrete context is missing. This can be seen as negative in particular or even helps the designers to get a more detached view. There is potential for expansion in the further concretization: Which dimensions can be adapted in each modality and can the context be limited to working machines.

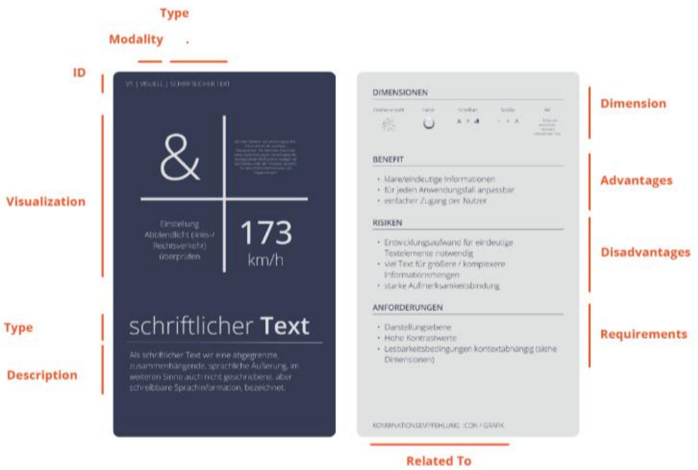

Figure 6: Layout and design of a Modality Card

Characterization: The last stage of development included characterization of information with associated signal specification. The workshop showed that the processing of the signal causes the context to be lost and the developer becomes engrossed in filling in a sheet and loses the context in the process. An extensive shortening with selective revision was thus necessary.

The division into information character and signal specification was kept. The former was changed in that the interaction vocabulary was thinned out and thus only six character pairs can be described. The signal specification received a reworking of the modality visualization and a reflection of the signal with advantages and disadvantages. The signal character was therefore not considered further, because the definition of generally valid parameters is very difficult and in the end it did not help the developer to generate a better solution by trying to perform an abstract evaluation.

The advantage of the clear and systematic division of information and its underlying individual signals remains and strengthens an analytical procedure. The documentability, the simple generation of different variants by modification of single signals and the character assessment of the information open new perspectives and speak for the presented design.

A disadvantage is that with every system the creativity is limited - how beneficial or harmful a system is for the development process is not to be examined further here. Furthermore, the problem remains of having to deal with many pieces of paper and sheets of paper, the designer loses himself in the documentation of his ideas. The creative process is also slowed down by the characterization; the rapid gathering of ideas cannot be taken into account.

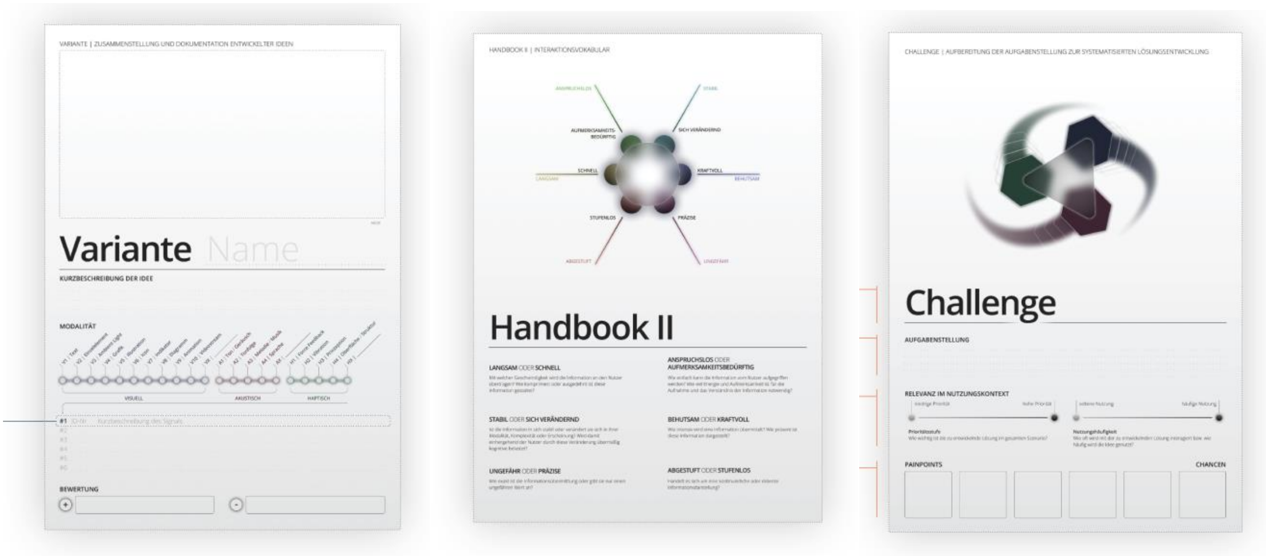

Figure 7: Final solution sheet, handbook and task definition template of the design tool

Handbook: As accompanying information all generally valid aspects were summarized, which developed during the testing and can be applied to all modalities. They are basic questions that can be asked by each participant in order to build up his solution proposals in a multimodal way. In addition, an illustration of the modality systematics is intended to help maintain an overview. Six thematic blocks have been identified:

- ***Main Modality and Secondary Modality:*** *Which modality is the main carrier of the information and which modalities can be integrated in a supporting way? How can information be enriched with an additional modality? Is there a clear hierarchy or are all modalities equal? Does the chosen combination lead to uncertainties for the user?*
- ***Purpose:*** *What is the underlying causality? What is conveyed by a concrete modality and why? What is the purpose of the concrete information? What importance can be assigned to the information with the corresponding modality? Is the purpose completely fulfilled or are aspects left unsatisfied?*
- ***Opportunities and risks:*** *Where is the greatest potential? Are there possibilities that are not (yet) used in comparable contexts? How can the user experience be further enhanced compared to competitors? Which problems can occur that affect the user experience? Where do you see*

technological limitations and how can you work around them? Where is the risk of misinterpretation of the user's interaction and what are the worst-case consequences? Location of information mapping. Where is the information mapped? Which locations are particularly suitable for building on existing structures? Can established technologies be used for other purposes? Where can new possibilities be created? But where can information be omitted and the provision of information be streamlined?

- ***Advantages and disadvantages:*** *What are the advantages of the chosen modality? How can a recognizable added value be generated? Where can problems occur and how can they be solved?*
- ***Links to other information:*** *How can the available information be linked to existing structures? How can already existing technologies be used multimodality and their potential be fully exploited?*

This variant is supported by the presence of all important information, which, when bundled, provides the developers with a guideline for multimodal design. In parallel, the creative process itself is not hindered. Thought-provoking impulses are given, which the designer can take into account at his own discretion.
The restraint, which is considered positive, means at the same time that there is a danger of ignoring the accompanying information. In addition, there is no clear methodological framework, a direct application reference is missing and the information can be integrated less well into the work in brain-storming, it merely accompanies this process.

Task definition. The fact that the precision and clarity of the task definition is essential in the development process and should always be the first step in the result-oriented development of solutions can be regarded as an axiom in large parts of the developer community. Therefore, the decision that a clear definition of expectations is also necessary in brainstorming sessions is all the clearer, so that an implementation in the present tool is more than appropriate.

The goal was to develop a methodical sheet which takes up the given task or problem, documents it in a way that is understandable for all participants and forms the basis for what is to be developed and why. An elaboration of such a tool requires an enormous psychological and work-scientific background, which can be recognizable at most in beginnings, since a founded elaboration would have exceeded the time frame and the available resources by far. Nevertheless, an attempt was made to design a proposal for the outlined problem.

The relevance of the task in the context of use was evaluated as very important, since it decides which solutions in the area of HMI development come to the fore. The relevance is made up of the priority factor, i.e. the importance of the solution to be developed in the context, and the frequency of use factor, i.e. the number of interactions per time interval. Relevance plays an important role because it can be used to tailor interaction elements to their usage. For example, it could be assumed that an electronic window regulator in a car would have a different appearance or functionality if opening and closing a window were similarly highly frequented as pressing the accelerator pedal.

A short assessment of the context can also be found on the sheet. Here it should be briefly stated what problems and what possibilities the solution of the task would mean. Which pain points does the user have and can these be eliminated? What potential and opportunity does the development offer for the use of the product?

Solution sheets: One of the biggest problems with workshops in design is the final documentation and summary of the ideas and thoughts that have arisen. It often happens that a summary is not worked out immediately after the meeting, so that a part of the content is unfortunately lost again without being considered. At the same time, not all ideas are interpreted in the same way by all participants, so misunderstandings arise if someone only documents the idea but did not develop it himself. Furthermore, considerable additional work is required if this content has to be digitized subsequently, if designers can no longer recognize the handwriting of their colleagues, or if logics are presented in an unclean and therefore incomprehensible manner.

To counteract these problems in the present tool, a way was sought to record the output of a workshop efficiently and directly. In the end, solution templates were created that should be able to do just that: A simple and quick documentation of the ideas by summarizing the concepts briefly and concisely by the developer or team in a uniform way according to a fixed pattern.

The solution sheet itself follows the basic layout of the other cards for consistency. The header identifies the sheet. A general graphical representation has been omitted in order to create space for a drawing layer. In this, the idea is to be explained in a sketchy manner. Below follows the title of the card, which can be provided with an ID. In the second half of the sheet an evaluation and classification of the idea is carried out. First, the developed idea should be briefly described in words. The used modalities are marked in the overview, below with an ID and the corresponding signal behaviour is explained. Thereby not every signal shall be programmable to the last detail, rather only an understanding shall be created. Finally, the developers should briefly discuss the variant and note both positive and negative aspects of the variant. The underlying signal is briefly described.

In order to describe the solution details of a modality more precisely, i.e. to detail the previously described signals, additional detail cards are provided. On the front side a representation in the selected scenario as well as the ID number are noted, on the back side the initial state of the signal is visualized as well as further explanations to the signal behaviour are recorded.

Digital Layer: To address digital collaborative work and Internet-based workshops the tool needs to work purely digitally or to digitize the physically created results. The idea was to place all solution variants on a uniform layout. All participants receive an identical underlay with a personal colour and use it to work out their solution variants. The files are then exported appropriately and brought together in one place. If the developed variants are then placed transparently on top of each other on an underlay, the developed ideas can be visualized, discussed and, above all, brought together to a suitable extent. New variants are created by combining different ideas and proposals. In parallel, a heat-map of the solutions can be developed, showing how multimodal the designers think or whether all solutions concentrate on one area.

In this way, the solution variants can be displayed in context, mixed with each other and entire catalogue of variants can be generated. At the same time, documentation and further processing is very simple.
The defined context layout limits and influences the design, an extension of the context is excluded. When selecting an appropriate underlay, questions arise about the size, focal length or view - a separate development process for the underlay is therefore inevitable.

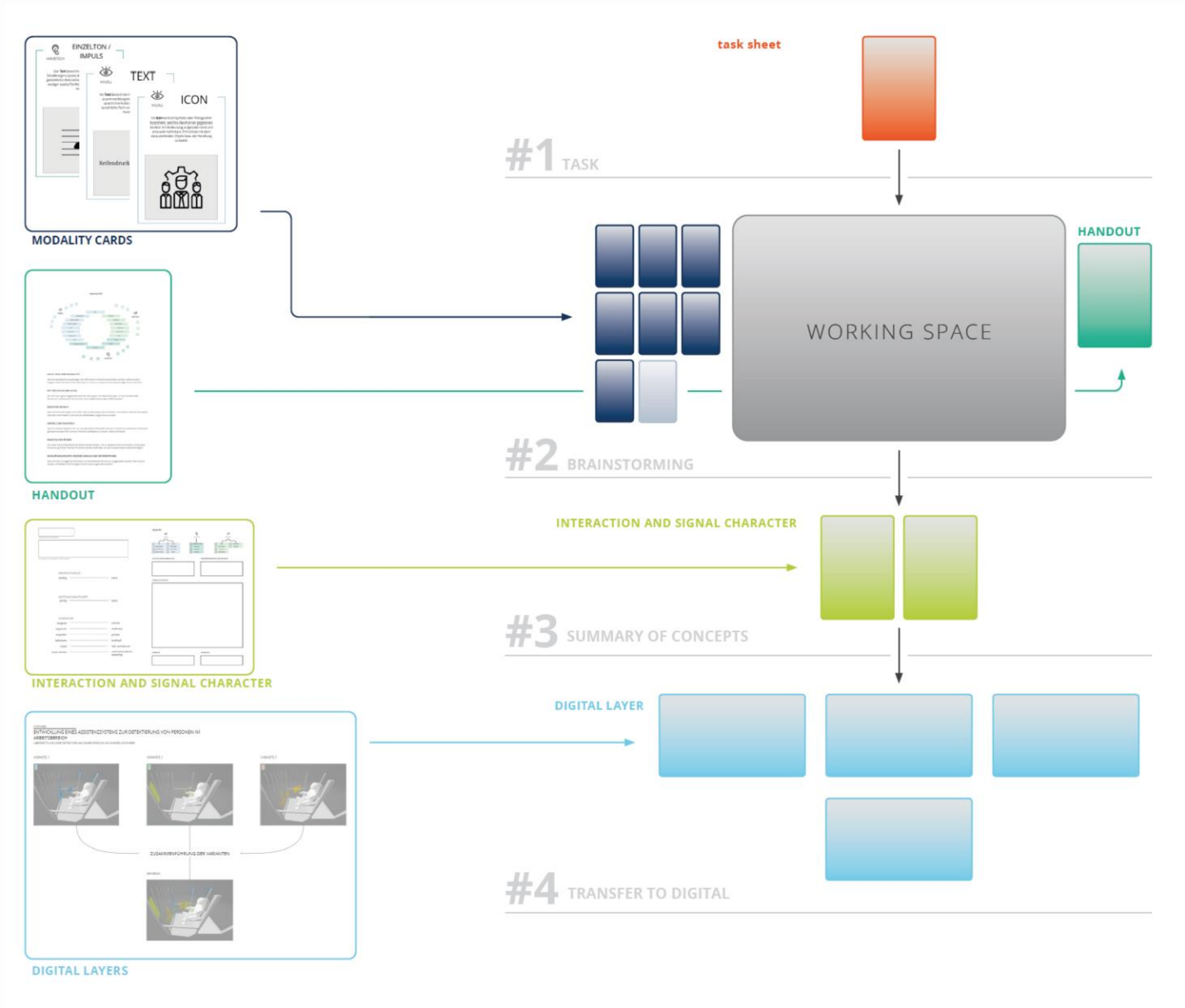

Figure 8: Synergy of the developed design tool components

Merging the components: Each of the described components plays its own part in the development of HMI or enriches the underlying process in its own way. This provides the unique opportunity to not weigh the ideas against each other as is usually the case, but to put them together as part of a larger whole. A development environment can be created that generates variants from the definition of the task, via brainstorming with accompanying information, which can ultimately be digitally assembled, compared and evaluated, or even offer the possibility of generating a final variant as a superposition solution.

Alternatively, a selection of one of the tool variants for detailed elaboration would also have been another possible procedure, but this would probably end up with a solution that is more difficult to use without an environment. The decision not to focus on one aspect of the solution, but rather to focus on a broad elaboration of all aspects of the solution at this point in the development process, has contributed significantly to the degree of practicality of the final product.

Inevitably, this decision meant a high amount of work, not only because each individual variant now had to be worked out, but also because they had to be synchronized with each other. This fact was willingly accepted due to the quality of the overall solution.

5 Results

5.1 Introducing the design toolkit

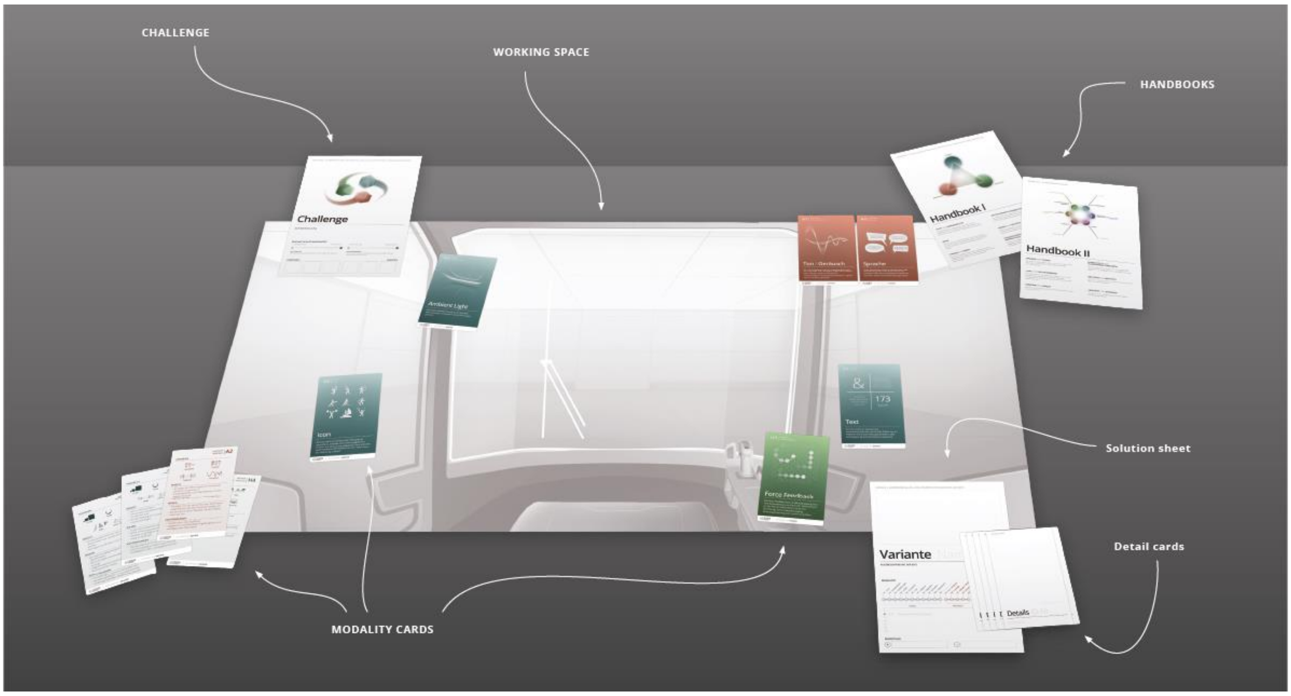

Figure 9: Developed toolkit for multimodal interaction design | Combining interaction challenge, modality cards, handbook I+II and solution sheets with detail cards

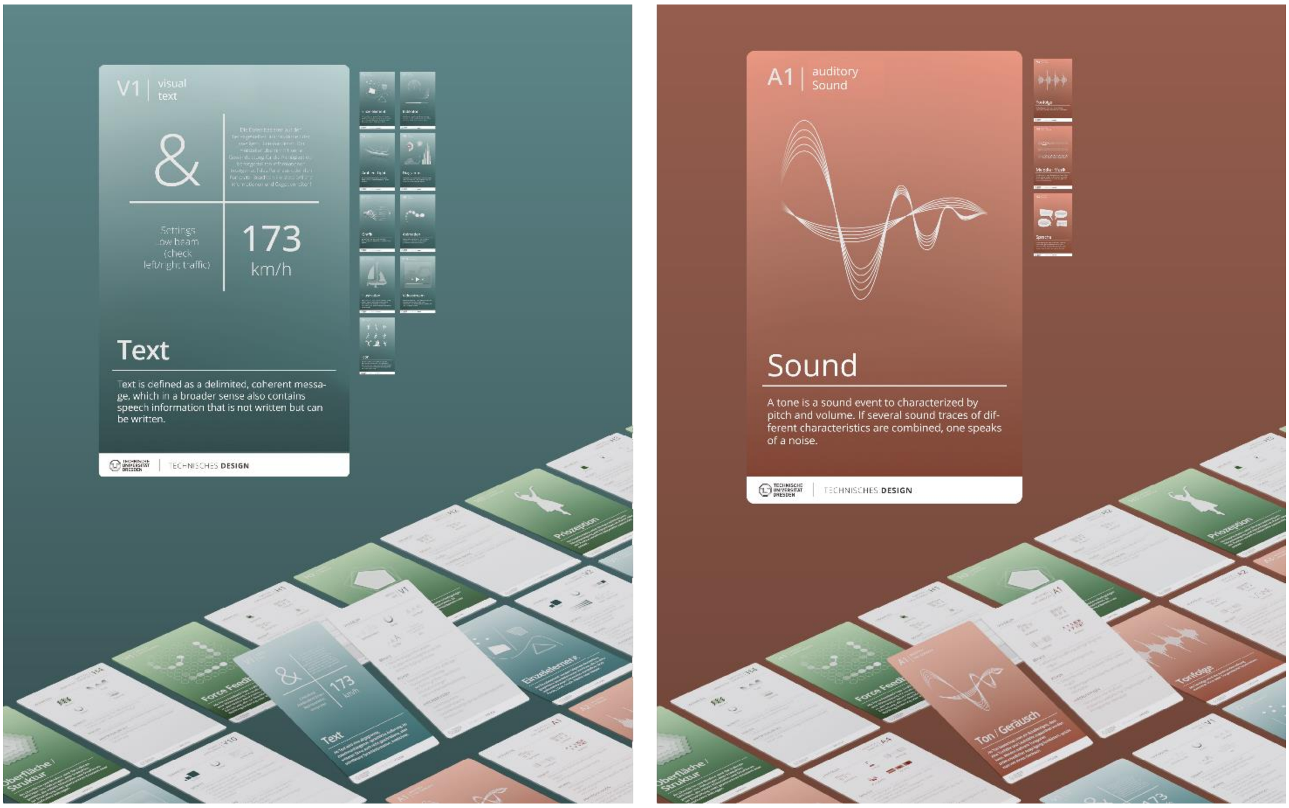

Figure 10: Result of the modality card set, the visual modality "text" and the auditory modality "sound"

The design toolkit provides a comprehensive methodology for developing multimodal interaction concepts. As an interplay of the challenge, modality cards, handbooks, solution sheets, and attached detail cards, the focus is on promoting multimodal thinking for HMI elements and their use for off-highway commercial vehicles and machines. By connecting these components, the purpose is to support developers’ innovative HMI solutions and further improve the user experience. It encourages the creative process while providing systematic guidance and subsequent evaluation of the ideas developed.
The set consists of ten visual modality cards, four auditory modality and four haptic modality cards. The consistent layout makes it possible to expand the set. The design is restrained aesthetically to provide a clear focus on the content aspects. The colour scheme is intended to support quick acquisition and assignment of the modalities.

Figure 11: Colour scheme of the modality cards

5.2 Brief description how to use the toolkit

The full toolkit consists of 5 components. To identify and evaluate the problem, the task definition (challenge) is used to analyse the context. The modality cards attached with the handbooks support the ideation process towards a plausible concept. Thereby, the handling of the materials is by no means fixed, rather, depending on the situation, it can be worked with in a different and very variable way:

Expansion of the mindset: *At the beginning of a workshop, the cards can be studied by all participants. A joint discussion of the modalities clarifies questions and raises awareness of their diversity. This is followed by the elaborative part of the workshop, in which the cards can be used as a cognitive support at certain points.*

Free brainstorming: *In known sequences, concepts are developed on the basis of the designers' ideas, without having previously dealt with the reality issue in greater detail. If unsatisfactory solutions are identified and no alternatives can be found, the cards can be used. From this, new ways of interaction can then be created with new, unthinking modalities.*

Extensions: *Iteratively, alternatives to already existing concepts can be conceived. It may be necessary to replace the existing modalities with new modalities in order to create space for the new. Subsequently, the resulting ideas are evaluated and reconciled with the existing solutions, so that an extension of the existing solution improves it.*

Systematics: *Another way of generating ideas can be a systematic approach. In this way, a solution can be developed for previously defined signal flows with each modality. Subsequently, one evaluates all developed thoughts, brings them together and thus creates a new interaction concept.*

Random: *The following approach can be much more playful: At the beginning of the workshop, all cards are shuffled. Then each participant draws a fixed number of cards. Now everyone has to work out a solution with the cards at hand, but may only use the drawn modalities. Subsequently, these ideas are evaluated, weighed against each other and a resulting preferred variant is developed.*

The possibilities of dealing with the toolkit are therefore very large and can be varied according to the purpose of use and the attitude of the participants. A wrong handling can be excluded, since already the occupation with the cards means an increase in value for the mental width of the developers. In any case, the variations should be documented and digitalized afterwards. The materials worked out are to help with it. Subsequent processing in digital form also helps to take the concept to the next level, but at the same time it provides an interface for making meaningful use of the results in the course of the process.

6 Discussion

6.1 General

Every theoretically conceived idea can sound so convincing on paper, but without corresponding practical tests, one can never assess how viable a concept actually is.

Therefore, it was decided to conduct a one-week test phase together with the design agency UXMA at their Dresden location. This agency was chosen because UXMA is a large international comparatively developer company with a lot of expertise and know-how - also in the field of HMI development. The employees involved should be provided with the full scope of the tool and be able to use the multimodality methodology after a 30-minute kick-off. The main interest of the study was to find out whether the modality cards cover all currently used modalities, details are missing for the respective modalities. Moreover, the test should identify interfering factors. In addition, the feedback from designers would be valuable, how they felt about the tool and its use, whether they adhered to one of the suggested ways of working, or how they adapted the tool for their use. In a questionnaire at the end of the week, the test persons were asked to classify their experiences. Ten questions are listed for a score ranging from zero to ten, with zero indicating complete negation and ten indicating complete agreement. Half questions ask about general aspects and half about specific points. One question, for example, is: "Do modality cards help in the creation of concepts at the human-machine interface and do they thereby generate added value for the development team, the product or the customer?”. During the test phase, UXMA also had an editable set of materials at its disposal on which changes could be noted, problems identified, or opportunities for improvement identified.

Unfortunately, the test phase could not be started at the time of the submission of this work due to scheduling difficulties on the part of the partner. But it is planned and should subsequently evaluate the usability of the resulting tool. A revision of the tool in the sense of completion with the hints of the lived practice at UXMA is expected.

What impact the design tool has on the development of the multimodal interaction remains uncertain. However, it offers a low barrier to entry into complex interaction design and promotes an explorative, creative process of idea development.

6.2 Limitations

Deliverable 4.2 does not provide concrete assignments of information to modalities, but provides recommendations for use of modalities based on their advantages and disadvantages. The concrete assignment and also evaluation can only take place in the ongoing development process. The tool supports this process without validation.

The implemented evaluation is not complete. The tool is aimed more at idea generation and checks suitability based on user-centred guidelines. Criteria from a technological (feasibility, complexity, robustness) or ecological (cost) point of view must be considered separately, but have an essential influence on the choice of modalities. This includes the evaluation of contextual factors. Appropriate expertise should be brought to the workshop by the developers, designers and users of the tool.

7 Further Work

In addition to the test phase with external design partners, the suitability of the toolkit should also be tested in smaller working groups in preparation for work package 6 *Integration, Demonstration, Validation and Verification*. Currently, there is a lack of experience as to how many participants and what scope the workshop formats should have and up to what complexity the toolkit supports efficient development. Non-designers should also participate in this test workshop. Their feedback is essential for its successful use, possible improvements and further developments. According to the design goals, the toolkit should not require special skills or mindsets, but its use and effectiveness in interdisciplinary development teams have not yet been tested. These workshops need to be planned, hosted, and evaluated.

In addition to internal test workshops at the TU Dresden, use-case specific workshops will be conducted with the project partners in an online format. To do this, the design toolkit has to be digitized. In retrospect, the physical approach was correct, but a stronger integration of digital connections would add further value to the result and the practical use in the project. At the next consortium meeting, a face-to-face workshop on ideation will be held, followed by rapid prototyping. The expected results are HMI ideas, which show concrete interactions via different modalities in an understandable way in the form of texts, images and first prototypes. If the challenge task focuses on isolated interaction flows, the resulting solutions must be transferred into a holistic operating concept. This illustrates the great influence of the task. Therefore, the preceding test workshops should also examine the right degree of sharpness and freedom of the task.

The given toolkit is recommended for use in early phases of development, brainstorming or design workshops, when the task is detailed enough to address concrete solutions. Another application is the revision of existing concepts or the iteration of a concept. Due to the wide range of modalities, it can be attempted to selectively adapt the information transfer multimodality and thus generate new solutions for data feedback. Furthermore, the technology choice and an overall evaluation of the developed ideas should be integrated into the development process. The listed criteria in sections 2 and 3 provide a comprehensive starting point.

The development of a generally valid analysis tool classifying operators' handling or cognitive load to recommend suitable information modality mapping seems too complex. Moreover, these recommendations would have to be checked for validity first in order to rely on them in the development process. Nevertheless, it would be a milestone in the development of human-machine interfaces. The testing and intervention within THEIA^{XR} could contribute to this.

References

[1] Norman, Donald A. (2013). The design of everyday things. Revised and expanded edition. New York: Basic Books.

[2] Goodwin, K. (2011). Designing for the Digital Age: How to Create Human-Centered Products and Services. John Wiley & Sons, *ISBN 1118079884, 9781118079881.*

[3] Reeves, L. M., Lai, J., Larson, J. A., Oviatt, S., Balaji, T. S., Buisine, S., Collings, P., Cohen, P., Kraal, B., Martin, J. C., McTear, M., Raman, T. V., Stanney, K. M., Su, H., & Wang, Q. Y. (2004). Guidelines for multimodal user interface design. Communications of the ACM, 47(1), 57-59. https://doi.org/10.1145/962081.962106

[4] Turk, M. (2013). Multimodal interaction:A review. Department of Computer Science, University of California, Santa Barbara. *Pattern Recognition Letters. Volume 36, 15 January 2014, Pages 189-195*.

[5] Diefenbach, S., Hassenzahl, M., Kloeckner, K., Nass, C., and Maier, A. Ein Interaktionsvokabular: Dimensionen zur Beschreibung der Ästhetik von Interaktion. In H. Brau, S. Diefenbach, K. Göring, M. Peissner and K. Petrovic, eds., Usability Professionals 2010. German Chapter der Usability Professionals' Association e.V, 2010, 27--32.

[6] Schmid, M.; Maier, T. (2017): Technisches Interface Design. Anforderungen, Bewertung und Gestaltung. Berlin: Springer Vieweg.

[7] Dumas, B., Lalanne, D., Oviatt, S. (2009). Multimodal Interfaces: A Survey of Principles, Models and Frameworks. In: Denis Lalanne und Jürg Kohlas (Hrsg.): Human Machine Interaction. Berlin, Heidelberg: Springer Berlin Heidelberg. S. 3–26.

[8] Thüring, M., Mahlke, S. (2007): Usabi-lity, Aesthetics and Emotions in Human–Techno-logy Interaction. In: International Journal of Psychology 42 (4). S. 253–264.

[9] Lenz, Eva; Diefenbach, S., Hassenzahl, M. (2014): Aesthetics of Interaction. A Literature Synthesis. In: Proceedings of the 8th Nordic Conference on Human-Computer Interaction Fun, Fast, Foundational. S. 628–637.

[10] Löwgren, J., Stolterman, E. (2004): Thought-ful Interaction Design. A Design Perspective on Information Technology. Cambridge: The MIT Press.

[11] Csikszentmihalyi, M. (2008). Flow. The Psychology of Optimal Experience. New York: HarperCollins eBooks.

[12] Damian, I. (2017). Social Augmentation Using Behavioural Feedback Loops. Dissertation. Universität Augsburg. Online verfügbar unter opus. bibliothek.uni-augsburg.de/opus4/frontdoor/index/index/docId/37750. Stand 11.05.2019.

[13] Janlert, Lars-Erik, Stolterman, E. (2017). Things that keep us busy. The elements of interaction. Cambridge: The MIT Press.

[14] Wickens, C., Derrick, W. L. (1984) A Multiple Processing Resource Explanation of the Subjective Dimensions of Operator Workload.

[15] Chen, Y.T., Chuang, M.C. (2014). The Study of Tactile Feeling and It's Expressing Vocabulary. In: International Journal of Industrial Ergonomics 44 (5). S. 675–84.

[16] Klinke, R.; Pape, H.; Silbernagl, S. (2005). Physiologie. ideal für die neue AO. 5. Auflage. Stuttgart: Georg Thieme Verlag.

ABBREVIATIONS / ACRONYMS

TTC	TTControl GmbH
TUD	Technische Universität Dresden
UC	Use Case
WP	Work Package
WPL	Work Package Leader
XR	Extended Reality
HMI	Human-Machine Interface
UUX	Usability and User Experience
RPM	Rotations per minute
KAL	Kalmar
PRIN	Prinoth AG
HdM	Hochschule der Medien