# A Method for Optimizing Complex Graphical Interfaces for Fast and Correct Perception of System States

Marie-Christin Harre[(⊠)], Sebastian Feuerstack, and Bertram Wortelen

OFFIS - Institute for Information Technology, 26121 Oldenburg, Germany
`harre@offis.de`

**Abstract.** The amount of information a human has to process continuously increases. In this regard, successful human performance depends on the ability of a human to perceive a system state as quickly and accurately as possible - ideally with a single glance. This becomes even more important in case several tasks have to be performed in parallel. It was shown earlier that monitoring user interfaces with a limited amount of information can be optimized for fast and accurate perception by combining all information into one integrated visual form. But systems that consist of several parallel tasks, each involving a whole bunch of parameters cannot be condensed into one single visual form. We propose an improved method that supports optimizing entire user interfaces consisting of several parallel tasks for fast and accurate perception (Konect). We evaluated the method in 6 workshops for that a total of 12 designers applied the method, which they learned by written instruction cards. Working in teams of two they were all able to design and optimize their designs first on a single task level (i.e. the original method) and thereafter on the global level (i.e. applying the new version). We evaluated their design outcomes thereafter in a laboratory experiment with 18 participants that were asked to distinguish critical and non-critical situations as fast and accurate as possible. Subjects were significantly faster ($p < 0.001$) and also significantly more accurate ($p < 0.001$) for those designs that were gained by the new version of Konect than those for the old one.

**Keywords:** Information visualization · Graphical interfaces
High amount of information in parallel · Systematic method

## 1 Introduction

The amount of information humans are expected to process on a day-by-day basis is steadily increasing. The more automation, the more the human is in charge of supervision. The more tasks are required to be handled in parallel, the more the human is required to divide attention between them. A user interface

that allows a human to perceive and process relevant information *at a single glance* would be the ideal one in such situations. In the area of automation, this relevant information often can be condensed to two basic questions: First, is something going wrong? And second: why? Model-based user interface engineering has implemented rigorous engineering processes for deriving user interfaces based on several models, such as task analysis and design, domain models [2] and derived intermediary models to display information in a structured and consistent way. For large or complex systems, task analysis typically ends up with a huge amount of identified task-relevant information, which often needs to be processed in parallel to record the system status and make appropriate decisions. For example, it is very important for an operator of a nuclear power plant, a power grid or on a ship bridge to capture a large amount of information efficiently and correctly. But even non-safety-critical systems, such as a booking system at a hotel reception or information dashboards, would benefit from being able to be observed efficiently with just a few glances. For example, more valuable time can be spent interacting with a customer than looking up and collecting that information in the user interface.

The research question that arouses in this context is: How does an engineering process that allows systematic derivation of **complex** graphical user interfaces look like that ensures a high amount of information to be perceived in parallel?

We propose a new version of Konect, a method for designing graphical user interfaces. Konect focuses on condensing information into visual forms (a glyph [36]), which are then optimized for fast and accurate perception [27]. Until now, Konect was applied for graphical interfaces with reduced complexity (not more than 11 information elements). In this contribution we applied the method to design **complex, multi-glyph interfaces** for that each glyph condenses all relevant information for processing one task. Konect comes with an instruction card, which is a five pages long written instruction that enables an interface designer with average design knowledge to apply the method without any prior coaching. We extended this card by an operationalization of three basic well known guidelines: Consistency, simplicity in shapes, and simplicity in colors. In an evaluation study, we show that the operationalization is facilitated by the Konect method and can be successfully used by designers. The resulting interface designs significantly reduce the perception time required to identify important information while at the same time improve perception accuracy.

The paper is structured as follows: First, we describe related works on methods for the design of visual monitoring interfaces that we considered for our work. Then we briefly motivate a use case about vessel performance monitoring in Sect. 3 that we use to introduce the method in Sect. 4 and later detail as it is also our use case for the evaluation. For the evaluation of the impact of the proposed optimization guidelines we hosted six design workshops with two designers per workshop, who first collaboratively applied the original method followed by the new parts of the revised method (Sect. 5.2). Thereafter, we performed a laboratory experiment with 18 participants to measure reaction times on the created original designs in comparison to the ones where the new Konect

version was applied (Sect. 5.3). Results are reported in Sect. 6 and discussed in Sect. 7. Supplementary material (Instruction card, workshop material, raw data and data processing scripts) for this study can be found online[1].

## 2   Related Work

In the context of this work, related work relevant for the design and optimization of human machine interfaces (HMIs) for monitoring can be subdivided in three areas. First, *efficient visual form representation and perception* is a topic that is relevant in data visualization. Second, *engineering processes* like e.g. model-based user interface engineering have been proposed that enable traceable links between design objectives, several intermediate models and the visual forms composing an HMI (e.g. by mappings). Finally, *methods and heuristics for interface composition* have been proposed (e.g. by guidelines) to ensure that single visual forms do not interfere with each other and as a negative consequence decrease the time and accuracy needed to detect anomalies.

In the scope of *efficient visual form representation*, Ware [36] provided a collection of key principles on how to create efficient information visualizations for single data elements. Similar to Ware, Meirelles [23] provides best practices for effective information visualizations. Cleveland and McGill [9] conducted experiments on how accurate individual forms (e.g. color, length) visualizing quantitative data are perceived by the human visual system. They offered guidelines for the visualization of quantitative data, which have been later extended by Mackinlay et al. [21] for nominal and ordinal data types. While Mackinlay et al. and Cleveland and McGill are situated in the data chart visualization domain, our work is focusing on user interface design. In this domain, Shneiderman et al. [31] present rules and key principles for user interface design and Tidwell [32] gives an overview about good design patterns applicable in the user interface design domain. Compared to our work these works have the main emphasis in interaction design while we have a focus on the visualization of monitoring information. Tufte [34] describes design strategies for presenting information about motion, process, mechanism, cause and effect. Gruhn et al. [17] provided best practices for visualizing information quite specifically for monitoring HMIs in the safety-critical domain (e.g. usage of trends, reduction of text). These works offer a profound background about best practices and hints for efficient visual forms for data visualization. However, they do not offer a systematic step-by-step approach in which the knowledge can be applied and therefore the systematic consolidation and application of this knowledge can be a challenge.

Regarding *methods and engineering processes*, Card et al. presented the "visualization pipeline", a sequence of steps and characteristics of a mapping functions for mapping information elements to visual forms [7], but do not state specific guidelines or mappings. Combined versions (step-by-step approach and specification of efficient mappings) exist in the domain of model-based user interface engineering [24]. Approaches in this domain involve the systematic derivation

---

[1] https://hcd.offis.de/wordpress/?p=399.

starting from user scenarios, via task analysis to the modeling of abstract, modality independent to concrete and final, executable user interface presentations [6]. This is based on predefined widgets (e.g. via a toolkit) and thus neglect creative visual forms and are mainly applicable for form based interfaces as toolkit widgets are limited. The idea of Post-WIMP widgets [10] addresses this limitation and arguing to open up the design space for new visual forms that consider new modes of interaction going beyond the desktop. [11] picks up on this topic and proposes COMETS, that implement plasticity and can adapt themselves to various contexts-of-use. Beside these general approaches, some methods have been focused on addressing specific challenges. For instance, Zhang et al. [45] presented a theoretical framework for the creation of relational information displays. Their framework describes properties and structures that focus on the expressiveness in designing visualizations. The Konect method is also situated in this context. It implements a step-by-step process to derive visual forms for specific information elements while considering the tasks and perceptual skills of an operator with regard to fast and accurate perception of visual forms. These works either focus on single visual forms or do not consider interference of visual forms, which might lead to an efficiency decrease with regard to the time and accuracy to perceive e.g. critical states.

Related work regarding *combination of visual forms* has been published by Woods [43]. Woods presented the concept of a "visual momentum", in which he presented guidelines to integrate data across successive displays so that the display system design can support an effective distribution of user attention. This has been later reexamined by Bennett and Flach [3]. The core concept is a consistent usage of display techniques. Wickens et al. [40] presented the "Proximity Compatibility Principle", which demands that information relevant for a common task or mental operation should be rendered close together in the perceptual space. Treisman [33] suggests that when perceiving a stimulus, features are "registered early, automatically, and in parallel, while objects are identified separately" and identified a preattentive stage in that different parts of the brain automatically gather information about basic features like e.g. colors, shapes and can be perceived in milliseconds. Wolfe et al. [42] conducted studies about what supports fast detection of single visual forms (called targets) in a range of other visual forms (called distractors). The authors revealed that feature search becomes less efficient as the target-distractor difference declines and as distractor inhomogenity increases.

## 3   Fuel-Efficient Operation of Maritime Vessels

The use case to which the method was applied for validation purposes is situated in the maritime domain: Economic vessel navigation is an open issue in the maritime domain. According to studies published by the European Federation for Transport and Environment, the fuel consumption of ships actually increased by 10% between 1990 an 2013 [13]. In 2016 the average design efficiency of ships was even worse compared to 2015 [14]. Much of a ship's emissions arise

close to the coast line and probably affect air quality negatively [35]. To further identify the causes of fuel consumption of ships, Ando et al. [1] presented a break down analysis: A high amount of the consumption is caused by the ships speed profile, the distances and the trim. In order to optimize the total energy consumption and reduce emissions, some ships offer user interfaces for vessel performance monitoring in order to be able to make right decisions to save fuel (e.g. use optimal speed, trim etc.). Currently such user interfaces are difficult to use in critical situations as docking: In berthing scenarios information about the distance to the quay wall, the current, position of tug boats or state of the thrusters is of higher importance for the task at hand as this directly affects safety of the ship. Information for the secondary task - save fuel and reduce emissions - has lower priority. Hence, as long as it is quite difficult for the operator to monitor these values, it will not be used in situations near land and port. Therefore, the idea was to design a user interface depicting all relevant information for fuel-efficient operation that can be monitored with minor effort while the operator performs the primary task (like e.g. berthing or maneuvering the ship).

## 4   Konect: Designing Visual Interfaces for Efficient Monitoring

The Konect method has been proposed to design glyphs that condense several information elements into integrated visual forms. The method consists of four basic process steps (Task analysis, Idea Box Creation, Design Creation and Global Design Adjustment with Composing Guidelines) like depicted by Fig. 1. All steps are performed by following a written instruction sheet. The design phase requires pen and paper as it involves free sketching. The steps are described in the upcoming subsections. For this, we describe the concept of the step and its realization in the vessel performance monitoring use case.

The first three steps focus on the visual form generation, while the fourth step focuses on glyph composition and optimization.
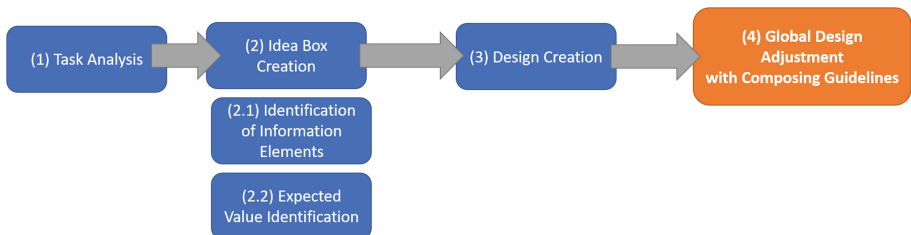


**Fig. 1.** Model based process for composing graphical interfaces for fast and correct perception.

### 4.1    Task Analysis

The first step is performing a task analysis to figure out for which tasks an operator has to acquire or monitor information [27]. The result of the task analysis is a task model. Various notations for task models exist like GOMS [8], ConcurTaskTrees (CTT) [29] or the HAMSTERS notation [15], which also considers temporal task relations like introduced by the CTT notation. Figure 2 shows parts of a task model created with HAMSTERS for crew members on ship bridges. The illustration was modified here to highlight the key elements of the task model required for the Konect method. Omitted elements of the task model are indicated by dots (...). The full HAMSTER model is included in the supplementary material (c.f. footnote at the end of the introduction). Optimally, the task model is not created just for the Konect method. Task modeling techniques, like HAMSTERS, can be used for the development and analysis of various aspects of the Human-Machine interaction [12].
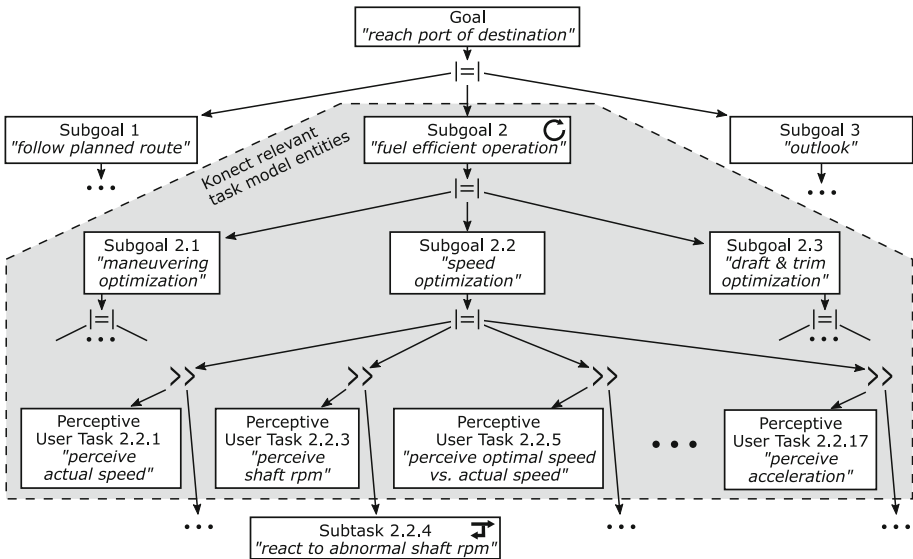


**Fig. 2.** Part of a task model for officers on ship bridges. Due to space restrictions many parts are omitted, indicated by dots (...). Aspects that are relevant for the Konect method are within the gray box. Graphical notation based on HAMSTERS notation. |=| indicates order independent execution of subtasks. >> indicates sequential execution of subtasks.

The main goal of the crew in our use case is to reach their port of destination. This task is depicted at the top of Fig. 2. To accomplish it several subgoals have to be achieved, like following the planned route, looking out for traffic, performing maneuvers and ensuring fuel-efficient ship operation. We focus on fuel-efficient operation as a use case, because a lot of information needs to be

observed in order to assess the ship state with respect to fuel-efficient operation. The task of assessing the state of complex systems can often be decomposed into observing the states of several subsystems. In Fig. 2 the three subsystems associated with maneuvering, speed control and draft& trim control have an influence on fuel-efficient operation and need to be monitored. This is depicted by the three subgoals 2.1–2.3. Typically there is no predefined order in observing the subsystems (indicated by he |=| symbol). To assess the state of a subsystem several information elements must be perceived. In the HAMSTERS notation this is expressed by a set of perceptive user tasks (e.g., Tasks 2.2.1 and 2.2.3 in Fig. 2 for subgoal 2.2). We assume that the order of these tasks is also independent. Perceiving an information element might result in a subsequent reaction, if the value of the information is outside of the desired range. This can also be included in the task model as an optional task (see subtasks 2.2.4). However, the Konect method focuses only on optimizing monitoring tasks, where the user has to observe large amount of information elements. It does not consider interactive aspects. The aspects of the task model relevant for the Konect method are included in the gray area in Fig. 2. Other methods like e.g. dialogue modeling have been proposed to support an interactive navigation between several user interface [22], which is out of scope for this work. An overview on model-based user interface design approaches can be found for example in [24]. The task analysis is important to derive the information that need to be presented by the HMI for the human operator to successfully conduct the tasks. Also the task-based grouping of the information is an outcome since in Konect, each task is represented by one glyph that spatially groups together all relevant information to perform one task.

## 4.2   Idea Box

The second step is filling the idea box. The idea box aims at fostering creativity of the designer by opening up a design space for one glyph and showing possibilities of visualizing the information at hand with different visual attributes. For each basic task of the task analysis (c.f. subgoals 2.1–2.3) one idea box is setup and filled with all information relevant for performing the respective task. As depicted in Table 1, the idea box consists of five columns:

**Importance.** The first column rates the importance of the information. The idea box is sorted according to the importance. Listing the important information before less important ones to ensure that the most efficient visual form is used for the most important information. Research on visual attention has identified four main factors influencing visual attention: The Saliency, the Effort, the Expectancy, and the Value of information (SEEV-Model). Whereas the former two factors describe bottom-up effects, the latter two are knowledge-driven (Expected-Value model [39]) and could be understood as the gold standard stating how an operator should optimally distribute her/his visual attention [41]. Konect applies the Expected Value model to derive the importance ranking.

**Table 1.** Extract of the idea box for the vessel performance use case.

| Importance | Information | Insight | Efficiency ranking | | | | Combination |
|---|---|---|---|---|---|---|---|
| 1.6 | Actual speed | Perceive quantitative value fast | Length | (1) | Volume | (3) | Symmetry |
| | | | | | | | Figure and ground |
| | | | Slope | (2) | Color hue | (4) | Spatial proximity |
| 0.8 | Actual speed vs. optimal speed | Perceive if value is ok fast | Color hue | (1) | Slope | (4) | Connectedness |
| | | | | | | | Continuity |
| | | | | | | | Closure |
| | | | Shape | (2) | Volume | (5) | Relative size |
| | | | Length | (3) | | | Similarity |
| 0.2 | Optimal speed | Perceive quantitative value | Position | (1) | Volume | (6) | |
| | | | Length | (2) | Density | (7) | |
| | | | Angle | (3) | Saturation | (8) | |
| | | | Slope | (4) | Color hue | (9) | |
| | | | Area | (5) | | | |
| . . . | . . . | . . . | . . . | | | | |

**Information.** All information relevant for a task and identified by the prior task analysis is listed as a separate row.

**Insight.** In the third column the insight the human operator should gain when looking at the information is described (e.g. perceive if value is ok fast). Insights are chosen from a predefined set of insights of the Konect method (all possible insights are listed in Table 2). In this regard, one has to keep in mind that one information element can appear twice in the idea box e.g. a speed value with insight to perceive it as quantitative value and speed with insight to perceive if it is ok or not. The insight is obtained and specified within the task analysis of the prior step, e.g. by talking to domain experts as part of interviews.

**Efficiency Ranking.** Based on the insight, the forth column of the idea box lists the most appropriate visual forms (e.g. color, length) according to an efficiency ranking. The most efficient attribute with regard to time and correctness of the percept is at the top of the list (ranked with (1)). Details about the efficiency ranking can be seen in Table 2.

**Combination.** Konect aims at generating glyphs – integrated visual forms. For this, all information elements listed in one idea box should be combined in one glyph. The last column lists possibilities to combine visual forms e.g. symmetry or proximity. This is strongly based on the Gestalt laws [38].

**Table 2.** Table listing predefined insights, the mapping to the visual efficiency ranking and the scientific reason for this mapping.

| Insight | Visual efficiency ranking | Discussed by |
|---|---|---|
| perceive **quantitative value**<br>perceive **summary (min, max, avg,%)** | position (l), length (2), angle (3), slope (4), area (5), volume (6), density (7), color saturation (8), color hue(9) | [7, 18] |
| perceive **quantitative value (fast)** | length (1), slope (2), volume (3), color hue (4) | [7, 16, 18, 28, 31] |
| perceive if **value is ok** | position (l), color hue (2), texture (3), connection (4), containment (5), density (6), color saturation (7), shape (8), length (9), angle (10), slope (11), area (12) | [7, 18] |
| perceive if **value is ok (fast)**<br>**find**<br>detect **anomaly**<br>perceive **outliers/exceptions** | color hue (1), shape (2), length (3), slope (4), volume (5) | [7, 16, 18, 28, 31] |
| perceive **category/mode**<br>perceive **certainty** | position (1), density (2), color saturation (3), color hue (4), texture (5), connection (6), containment (7), length (8), angle (9), slope (10), area (11), volume (12) | [7, 18] |
| perceive **category/mode (fast)** | color hue (1), length (2), slope (3) | [7, 16, 18, 28, 31] |
| perceive **pattern**<br>perceive **relationships**<br>perceive **trade-offs**<br>**compare**<br>perceive **clusters (groups, similarities)**<br>perceive **paths** | edges/depth/orientation at multiple scales (1), size/location (2), categorical relation (3), coordinate relation (4) | [16, 17, 32, 38] |

## 4.3   Design Creation

The idea boxes created in the previous step for each task are used in the design creation phase. With guidance offered by the box, the designer starts working with the idea box row-wise from top to bottom. For each row of the idea box, the designer chooses the most efficient visual form (e.g. color) for an information. In this step, the designer has to design an exact *instantiation* of this form. Exact instantiation means that for instance an exact color coding is chosen. This can be for instance that a critical or abnormal state is colored in red (e.g. actual speed does not comply with plan speed) while a normal state is colored in green

or a neutral color. Another example is the exact instantiation of a *length*. This might be for instance a horizontal or vertical bar.

### 4.4   Visual Form Composition

Complex interfaces supporting several parallel tasks and therefore consisting of several glyphs need to be carefully designed. Interference between glyphs can have an impact on the overall operator's perception accuracy and reaction time. Looking at related work (e.g. [17–19, 23, 26, 31, 32, 40, 43]) we identified three guidelines as most promising to have an impact on perception performance and accuracy:

**Consistency.** Use the same visual attribute for the same kind of insight for similar important information elements.

**Simplicity in Shapes.** Choose simple shapes and visual forms, choose non-accidental visual forms with regard to orientation.

**Simplicity in Colors.** Reduce colors for elements that do not carry any information besides from structuring the interface.

We are aware that there exist even more guidelines in related work. Based on the basis of human perceptual skills, we decided to use exactly these three guidelines. Details about the reasons are given in the upcoming subsections. Furthermore, the guidelines in related work are presented as abstract hints. We are not aware of a systematic procedure on how to apply these guidelines systematically. But this becomes especially important when dealing with a high amount of information. Thus we operationalized these statements to be used within Konect as follows:

**Consistency.** All information elements (i.e. rows) of all idea boxes are sorted according to their insights and the chosen visual property for each information is marked (e.g. color or length in the efficiency ranking column). Then these markings are checked to ensure that the same kind of insight (e.g. detect anomaly) is reflected by the same visual property (e.g. color). Furthermore, the same manifestation of the visual attribute needs to be chosen for the same kind of information (e.g. if red is an indicator for a failure, red should be used for this kind of information throughout the entire HMI design). The consistency is necessary for ensuring that pre-attentive perception of critical information is even possible at a global level. Ware et al. [37] stated that pop-out effects depend on the relationship of a visual search target to other targets that surround it. The strongest effect occurs when a single target-object differs from all other objects and where all other objects are identical or very similar to each other [37]. This means they have to be consistent to each other to ensure that an abnormal state directly pops-out and is easily distinguishable form normal state. This aspect is further emphasized based on findings of Wickens et al. [40]. Wickens presented the Proximity Compatibility Principle for display design. This specifies that displays relevant to a common task or mental operation should be rendered

close together in perceptual space which means that visual elements representing similar states should be encoded close together e.g. by using similar colors.

**Simplicity in Shapes.** To ensure simplicity, the metric for estimating the "goodness" of visual shapes [44] must be applied. One indicator for visual form simplicity is it's orientation. If a form is perceived as "non-accidental" with regard to orientation (thus it looks the same regardless of the viewer's perspective) it can be understood as a simple form [4]. For calculating the form simplicity, all visual shapes on the entire glyph designs are rotated and mirrored. Then the number of different appearances of the same form represents an indicator for a shapes goodness. Prior work has shown that the simpler a form is, the easier it can be recognized [36]. More complex shapes lead to higher visual clutter causing a decreased recognition performance due to "occlusion, greater difficulty at both segmenting a scene and performing visual search" [30]. Since this metric figured out to be initially difficult to comprehend, Konect comes with an illustrating example on the instruction card, which is depicted in Fig. 3.



| 0° | 90° | 180° | 270° | Horizontal reflect | Vertical reflect | Diagonal reflect | Diagonal reflect | Count |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | 1 |
| | | | | | | | | 2 |
| | | | | | | | | 4 |
| | | | | | | | | 8 |

**Fig. 3.** Instruction card for simplicity in shapes.

**Simplicity in Colors.** Beside simplicity in shapes also simplicity in colors should be achieved in the user interface design. Therefore, each element on the design is systematically looked at and it is checked against the table whether a color appearing on the design encodes information or if it rather represents *visual sugar*. If the latter is the case, the color should be taken out by being replaced by a neutral color (e.g. grey). This step is especially important to maintain the pop-out effects produced by pre-attentive visible visual forms on a global level. The use of colors should be reduced following the works of Kosslyn et al. that proposed the principle of *Informative changes*. This principle indicates that large changes across properties of a display that do not carry information should be avoided [18, 20].

## 5    Konect Application for the Vessel Performance Monitoring Use Case

While earlier work applied the method for simple one-glyph user interfaces in the automotive domain [27,28], we applied it in a more complex setting with three parallel tasks in the maritime domain for vessel performance monitoring. For the first steps of the method (the task analysis and idea box generation) we performed domain expert (i.e. ship master) interviews and a literature study. The second part (the design and application of the guidelines) was performed as a workshop with designers. We have chosen this separation to support the evaluation. First, we were interested to start with an identical design space for all designers and second, we were not able to recruit sufficient designers with maritime expertise. Therefore, we provided the designers with the task analysis and initial ordered idea boxes for the three tasks and explained them the role of each task and information at the beginning of the workshop by an instruction sheet. We decided to use a textual description to avoid differences between the information, the different experts got in the workshops and to control the effect a potential bias caused by further oral instructions.

### 5.1    Task Analysis and Idea Box Generation

We performed the task analysis based on information of two different data sources: Regulatory information and an interview with a mariner. For the former, we reviewed the Resolution MEPC.282(70) [25] of the Marine Environment Protection Committee (MEPC) of the International Maritime Organization (IMO). Appendix 10 of the MEPC document deals with the development of a ship energy efficiency management plan. In its Sect. 5 it lists aspects that should be considered for fuel-efficient operation. We used this list as a first input for the task analysis for fuel-efficient operation. For the latter, we interviewed a mariner about tasks and information that are relevant for fuel-efficient operation of vessels. As the use case deals with monitoring of fuel-efficiency on board, we selected only those tasks that involve ship personnel during voyages. We ended up with a hierarchical task tree of 48 perceptive user tasks for the *fuel-efficient operation* goal, for that each relates to one source of information. At the topmost level, under the goal of a fuel-efficient ship operation, we identified three main subgoals:

**Speed Optimization**: To save fuel during voyage the speed, acceleration and braking should be minded. Each ship has an optimal speed at which the fuel consumption is minimal. To estimate if the ship is driving at optimal speed the actual speed has to be known and has to be compared to the optimal speed. Furthermore, the ship has to reach the harbor in the right time for unloading cargo to avoid waiting times or problems with the place to berth. This is often called "just in time speed". This speed has also to be minded and compared to the actual speed. Finally, the acceleration is important for saving fuel. A more smooth/constant acceleration reduces fuel consumption. This can be seen with

the acceleration trend and the shaft rpm trend – an indicator for the rotation speed of the propeller.

**Maneuvering Optimization**: Aspects influencing the maneuvering can also lead to higher fuel consumption e.g. headwind or high waves can lead to a higher resistance and thus to higher fuel consumption. Thus the relative wind direction and speed as well as relative wave direction and speed is relevant. Furthermore, the maneuvering should be quite smooth to avoid high fuel consumption (similar to acceleration trend). Thus the rudder angle trend is important and the autopilot state, as the autopilot can ensure smooth adaption of the ships heading according to a set course.

**Draft and Trim Optimization**: This includes the monitoring of trim (horizontal position of the ship), draft (the part of the ship below water), and the depth under keel (distance of the ship to the ground). For this one has to know that there exists an optimal trim and an optimal draft for each ship for having a minimal possible fuel consumption. This differs from ship to ship. The depth under keel becomes important at a certain value at which the low depth can increase fuel consumption. Thus sometimes the information that the depth under keel is above or below this value is more important than the actual depth under keel.

To generate the idea boxes, we ordered the task relevant information based in the Expected-Value model. Feuerstack et al. [16] showed that the more experts are being ask to determine the Expected-Value model the more the effect of individual errors can be reduced. The authors proposed the HEE software tool to systematically collect the data and to average the data collected. We applied the tool and identified the information gained by the task analysis [25] on current state of the art Vessel Performance Monitoring interface (from Kongsberg and Trelleborg) and asked three experts (a ship master and two Human Factors experts that were aware of the IMO regulations) to determine the relevance of each information for each of the operator tasks and the expectancy of each information with the tool. Expectancy in the context of the Expected Value-Model states how often one expects to perceive new information from a given source (e.g. a presented sensor value).

### 5.2 Workshop

We organized six workshops with each composed of two Human Factors Engineers or HMI Designers (5 female, 7 male). We captured prior experience of the workshop participants via a questionnaire at the beginning in which the participants estimated their knowledge and competencies according to a Likert-Scale ranging from 1 (never heard of) to 5 (often applied). The participants stated that they have often done an HMI Design (4.6) by applying task analysis techniques (4.08) and information visualization (4.17). They knew the concept of preattentive perception before but did not apply it (3.25). The workshop participants collaboratively applied first the original Konect method ending up with one design of the set depicted by Fig. 4. Afterwards all workshop groups applied

the guidelines (4th process step) to end up with one of the optimized designs shown in Fig. 5. We chose groups of two experts to ease our observation of the discussion of the experts in order to get further insights into the reasons for choosing visual forms, the overall HMI Design and their way of working with the method. Each workshop lasted about 3 hours. Within the workshop, we had the following procedure: As starting point, the designers read an instruction card of the Konect method and the use case. The experts received already filled idea boxes and were asked to focus on the third step of the Konect method – the free sketching phase. In this part each pair of experts created a design solution for the maritime use case. No time limit was set for the design creation. Instead, we asked the participants to indicate that they are finished with their design as soon as they are satisfied with their design solution. On average, the designers required around 2 hours to create the design based on the original Konect method. Their design results of the first part of the workshop are shown in Fig. 4.
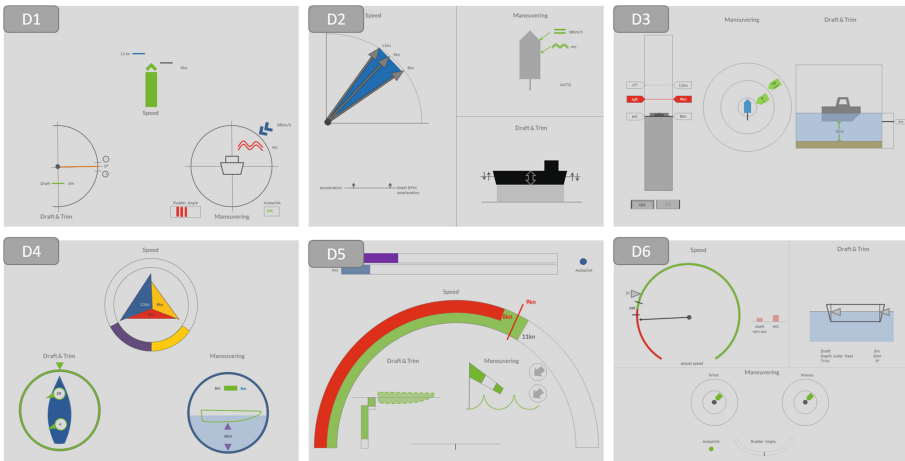


**Fig. 4.** Designs for the vessel performance use case after step (3) - design creation.

After none of the designers had any additional idea how to further improve the design results, the designers got a second sheet of instructions (they were not aware of a second step beforehand) with the operationalization of the three guidelines as described in the previous section. They were then asked to refine their designs according to these guidelines. The designers needed around one hour to implement the changes to their designs. The results are shown in Fig. 5. We observed the designer teams during their design creations. For applying the guidelines some exemplary changes they implemented were:

**Consistency.** An example based on the designs in Fig. 4 is for instance the depiction of wind and waves in design D1. In D1 wind and waves are depicted in the right corner at the bottom. The waves are depicted as a wave form which

is always red, and the wind with the arrows that are always blue. This has been changed after applying the consistency guideline. Figure 5 shows the designs after application of the composition guidelines. The designers of D1 recognized that green is used as color to indicate an uncritical state throughout the whole HMI and thus applied this color coding also to the wind and wave visualization. Thus a red indicator is used to create a pop-out effect in G-D1 in Fig. 5 while in D1 in Fig. 4 red has been used as indicator for waves, causing the pop-out effect for "red" to be reduced throughout the whole display.
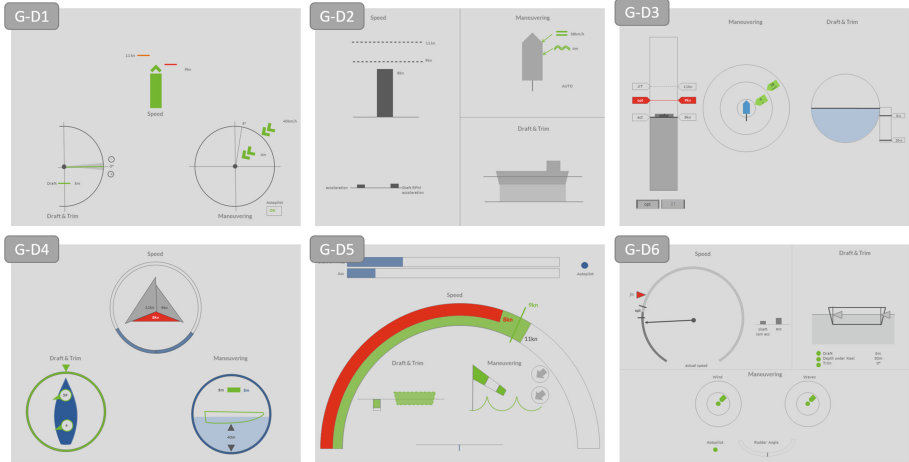


**Fig. 5.** Designs for the vessel performance use case after step (4) - global design adjustment with composing guidelines.

**Simplicity in Shapes.** For applying this guideline, the experts rotated and mirrored all shapes of all visual forms used in the HMI design. In case rotation or mirroring of a shape changes its appearance a simpler shape has to be found (e.g. see Fig. 3: a bar is simpler than a half-circle and should therefore be preferred). This can be seen for instance in D2 in Fig. 4: The experts used a circular chart with arrows in the upper left corner to visualize the speed. After applying the guidelines this has been changed to a simple gray bar chart (see G-D2 in Fig. 5).

**Simplicity in Colors.** For instance in design D4, the experts make extensive use of different colors. All colors that did not carry any information have been reduced, which resulted in the fact that the critical red information in G-D4 in Fig. 5 stands out more saliently.

### 5.3   Study

With the designs resulted from the workshop, we conducted a laboratory study in front of the computer to estimate the effect of the extended step 4. The

participants and the procedure are described in the upcoming subsections. With the experiment, we wanted to investigate in how far step 4 (as shown in Fig. 1 improved the efficiency of the design solution. For this, we had the following hypotheses:

H1. Designs after step 4 are faster to perceive compared to designs created with the original version of the method.
H2. Designs after step 4 have a higher accuracy of the perceived state compared to designs created with the original version of the method.
H0. There is no difference between designs after step 4 and designs created with the original version of the method.

**Participants.** Participants were recruited via a notice on the electronic bulletin board of the university. All participants gave prior written informed consent and were economically rewarded with 15 EUR at the end of the experiment. The study included 18 participants (14 women), whose ages ranged from 20 to 29 (mean $=24.5$, SD $=2.7$). We choose students as participants as the method focuses on human perceptual skills in general (fast and accurate perception of visual cues). A trained maritime professional would have included mental model knowledge that we intentionally excluded as the use case was just one example and Konect should be seen as domain-independent method.

**Procedure.** At first, the participants were given a description of all designs which they read carefully. After this participants were shown a series of figures. Each figure showed one of the 12 designs (6 initial designs + 6 designs after step 4) in one of five different situations that are either critical or non-critical ($S1$: non-critical; $S2$: critical ship speed; $S3$: critical trim state; $S4$: critical (low) depth under keel; $S5$: critical wind condition). The task of participants was to estimate as fast and accurate as possible, whether the shown situation is critical or not. After pressing a key the image disappeared immediately and participants had to type their response. We measured the correctness of the response and the reaction time between showing the image and hitting the key. The sequence of designs and situations was randomized. However, we wanted to avoid having the initial design and the designs after step 4 of the same design concept (e.g. D1 and G-D1) in one sequence. Therefore we divided the experiment into two blocks (A and B) for each participant. Each A-block contained three initial designs and three composing guidelines designs from different design concepts (e.g., D1, D2, G-D3, G-D4, D5, G-D6). Block B contained the six remaining designs (e.g. G-D1, G-D2, D3, D4, G-D5, D6). To eliminate order effects we randomly assigned designs to the blocks and balanced it across participants. For each design each of the four critical situations (S2-S5) were shown once and the non-critical situation (S1) twice. This was repeated three times resulting in a sequence of $6\times(4+2)\times3 = 108$ stimuli shown in each block. To eliminate training effects blocks were iterated two times with pauses between them: A - B - A - B. Only data from the second iteration of block A and B was analyzed. We excluded

5 measurements, considering reaction times $>10\,\mathrm{s}$ as a failure and $<300\,\mathrm{ms}$ as a premature key press, because perception, decision and motor response are unlikely to happen within $300\,\mathrm{ms}$.

## 6   Results

We expected that the guideline application decreases reaction times and improves accuracy. However, we assumed that other factors also have an influence on the reaction times and accuracy. To take several factors into account, we used a linear mixed model to analyze the data.

We included the usage of the guidelines and the stimulus index as fixed effects in the model, with stimulus index $k$ indicating the $k$-th situation shown to the participant. When measuring reaction times in laboratory settings a training effect can often be observed, as did [27] when they tested HMIs created with the Konect method. We expected this effect in our experiment and tried to reduce it, by only analyzing data from the second iteration of block A and B. However, a small effect can still be expected. Thus we included the stimulus index as fixed effect. However, we did not expect to observe a significant training effect in the accuracy of responses. Due to individual differences, we expected that reaction times differ for each participants. Thus we added participant as a random factor. Similarly, the six pairs of designers created different design concepts. We expected to record different mean reaction times accuracy values for each design concept, but still expected an improvement for the designs after guidelines were applied. Thus we added the design concept as a random factor. This resulted in the following models (R-typical notation):

$$\text{response time} \sim \text{guideline} + \text{stimulus} + (1|\text{participant}) + (1|\text{design})$$
$$\text{accuracy} \sim \text{guideline} + (1|\text{participant}) + (1|\text{design})$$

Based on likelihood ratio tests we observed a significant effect of the guideline usage both on the reaction times ($\chi^2(1) = 182.17, p < 0.001$), lowering it by about $172\,\mathrm{ms} \pm 12.6\,\mathrm{ms}$ (standard errors), and on the accuracy ($\chi^2(1) = 43.41, p < 0.001$), improving accuracy by about $4.6\,\% \pm 0.7\,\%$ (standard errors) from $92.3\,\%$ correct responses to $96.9\,\%$. Thus, we were able to accept our hypotheses (H1) and (H2) and reject the null hypotheses (H0).

When exploring the data we realized that adding the designs as a random effect is not sufficient to describe the influence of the design on the reaction times and accuracy. The amount of improvement on accuracy and reaction times when applying the guidelines differs for each design. This sounds plausible. It is likely, that the initial designs of some design teams already adhered to the guidelines to some extent. Therefore the room for improvement after applying the guidelines differs for each design team. Adding random slopes for guidelines to the random effect term of our models (changing (1|design) to (guideline|design)) resulted in a significantly better model fit for reaction times ($\chi^2(2) = 195.19, p < 0.001$) as well as accuracy ($\chi^2(2) = 85.38, p < 0.001$).
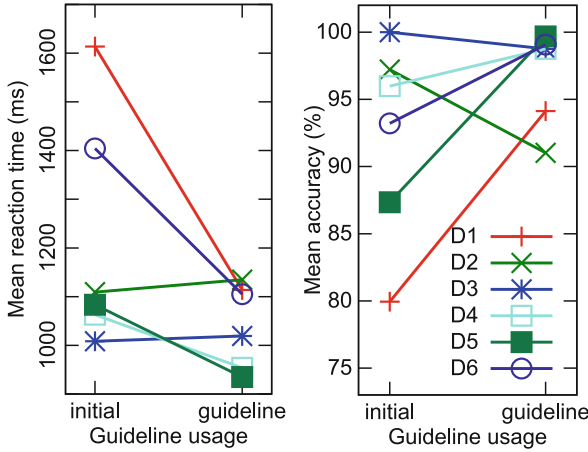
**Fig. 6.** Effect of guidelines on reaction times and accuracy differs for the different HMI designs.

Figure 6 illustrates this aspect. It shows how reaction times and accuracy changes with guideline usage in step 4 for each design. Those designs, where the initial design version did not perform well, showed the strongest improvement after step 4.

In Fig. 6 it can be seen that the effect size of the guidelines on decreased reaction time and increased accuracy differs between the designs: In case of D1 and D6 a high increase of performance with regard to reaction times and accuracy has been reached while for D4 and D5 a medium effect was visible and for design D2 and D3 nearly no improvement has been made. We assume that this is caused by quality differences in the initial designs: In D2 and D3, which implement a consistent and reduced usage of colors and simple shapes we measured already a fast reaction time and high accuracy. In case of D1 and D6, the designers applied colors inconsistently and used them also for structuring the user interface. In this case, the application of the guidelines strongly improved subjects performance. We assume that the positive effect of the Konect design principles on reaction times and accuracy of individual glyphs is negatively effected by inconsistent usage of colors and shapes across multiple glyphs in case the original method is applied to a higher amount of information. We could show that applying the guidelines improved reaction times and accuracy significantly. When we compare the designs without guideline application and the ones with guidelines application, we could observe that most design teams only changed few aspects of their design. According to our assumption, the reaction times to all visual forms should improve and not only to the ones that were changed. If this is the case we should observe that the reaction times for all situations (S1–S5) for the same design concept should change roughly by the same amount. An indication of this effect can be seen in Fig. 7. It shows the change of reaction times separately for each situation in each design. It can be seen that the range

of changes within a design is considerably smaller than the range of changes over all designs. This supports our assumption. However, given our dataset, this is only an indication. To reasonably test this a study design with more than just five situations and even more complex use cases is needed.

In Fig. 7 it can be seen that all reaction times have been improved or almost remained the same – except of design D2 in case of assessing a critical speed. Having a deeper look at the changes done after guidelines application in D2 reveals the reason for this effect: the color coding for assessing a critical speed has been totally removed instead of adapted to a more consistent form. Thus the workshop participants overlooked that they reduced the amount of information shown on the user interface. For future application of the method this problem might be solved by adding an additional review step at the end in which a verification that all information in the idea box are visualized in the design solution is conducted. We are currently working on this step and a tool support to resolve this problem.
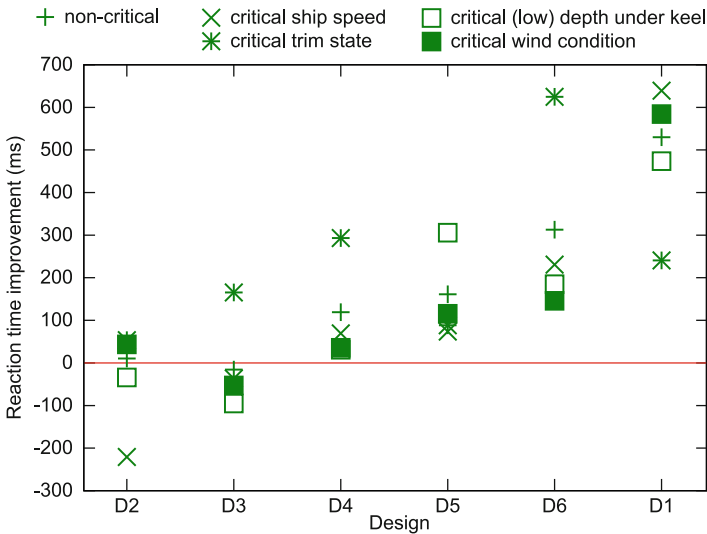


**Fig. 7.** Change of reaction times for all 5 situations in each design.

## 7    Discussion

One might argue that the guidelines alone might be good enough to achieve the objective of designing user interfaces that are more efficiently perceived by users without having their operationalization as part of the Konect method: Bornoe et al. [5] tested the improvement process of existing designs with design cards with novice users. The design cards offered hints as they contained principles, an explanation of effect and for what purpose the principle is used. The

authors found that the cards diversified the range of system aspects that novices considered and supported ideation. Nevertheless, the cards did not compensate the limited design experience of novices, as the participants had problems implementing them in the design process. With the exact systemization of steps to conduct to implement the guidelines as it is offered by Konect, this problem is addressed and did not appear in the prior described workshops.

The Konect method addresses a very specifc class of user interfaces: Visual user interfaces for monitoring in supervisory control situations with medium to high criticality to timely responses. The method can in principal be used in other contexts, e.g., for designing visualisations for infographics or games. However, fast and accurate perception might not be the optimality criterion in such contexts. Intuitive understanding of visualisations might be more important for infographics, while the joy of use might be a key factor for gaming visualisations. We like to emphasize, that a user of the Konect method should always be aware of the context of use and the criterions that apply to the specific context.

## 8    Conclusion

We presented an extension to the Konect method, which is used to create glyphs (integrated visual forms) optimized for fast and accurate perception. The method itself opens up a huge design space (since it fosters free sketching), while at the same time is still able to optimize the results for perception. The extension introduces operationalizations of three design principles (consistency, simplicity in shapes and colors). While the original method was not well suited for designing large user interfaces, we showed that the extended method can be applied to design complex user interfaces with multiple glyphs. By workshops with a total of 12 designers working in pairs of two, all of them were able to end up with designs that were created solely by following the written instructions of the method. In a subsequent laboratory study, we could show that 18 participants were on average significantly faster and more accurate in distinguishing between critical and non critical situations visualized by those designs that were created by the revised version than for those designed by the original one.

## References

1. Ando, H.: Performance monitoring and analysis for operational improvements. In: International Conference on Ship Efficiency (2011)
2. Balme, L., Demeure, A., Barralon, N., Coutaz, J., Calvary, G.: CAMELEON-RT: a software architecture reference model for distributed, migratable, and plastic user interfaces. In: Markopoulos, P., Eggen, B., Aarts, E., Crowley, J.L. (eds.) EUSAI 2004. LNCS, vol. 3295, pp. 291–302. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30473-9_28
3. Bennett, K.B., Flach, J.M.: Visual momentum redux. Int. J. Hum. Comput. Stud. **70**(6), 399–414 (2012)
4. Biederman, I.: Recognition-by-components: a theory of human image understanding. Psychol. Rev. **94**(2), 115 (1987)

5. Bornoe, N., Bruun, A., Stage, J.: Facilitating redesign with design cards: experiences with novice designers. In: OZCHI (2016)
6. Calvary, G., Coutaz, J., Thevenin, D., Limbourg, Q., Bouillon, L., Vanderdonckt, J.: A unifying reference framework for multi-target user interfaces. Interact. Comput. **15**(3), 289–308 (2003). https://doi.org/10.1016/S0953-5438(03)00010-9
7. Card, S.K., Mackinlay, J.D., Shneiderman, B.: Readings in Information Visualization: Using Vision to Think. Interactive Technologies Series. Morgan Kaufmann Publishers, Burlington (1999)
8. Card, S.K., Moran, T.P., Newell, A.: The Psychology of Human-Computer Interaction. Erlbaum, Hillsdale (1983)
9. Cleveland, W.S., McGill, R.: Graphical perception: theory, experimentation, and application to the development of graphical methods. J. Am. Stat. Assoc. **79**(387), 531–554 (1984)
10. van Dam, A.: Post-wimp user interfaces. Commun. ACM **40**(2), 63–67 (1997). https://doi.org/10.1145/253671.253708
11. Demeure, A., Calvary, G., Coninx, K.: COMET(s), a software architecture style and an interactors toolkit for plastic user interfaces. In: Graham, T.C.N., Palanque, P. (eds.) DSV-IS 2008. LNCS, vol. 5136, pp. 225–237. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-70569-7_21
12. Diaper, D., Stanton, N. (eds.): The Handbook of Task Analysis for Human-Computer Interaction. CRC Press, Boca Raton (2004)
13. Faber, J., 't Hoen, M.: Historical trends in ship design efficiency. Technical. report (2015). www.transportenvironment.org/sites/te/files/publications/
14. Faber, J., 't Hoen, M.: Estimated index values of ships 2009–2016: analysis of the design efficiency of ships that have entered the fleet since 2009. Technical report 17.7L97.69, CE Delft (2017). https://www.cedelft.eu/en/publications/download/2345
15. Fahssi, R., Martinie, C., Palanque, P.: Embedding explicit representation of cyber-physical elements in task models. In: IEEE International Conference on Systems, Man and Cybernetics (SMC 2016) (2016)
16. Feuerstack, S., Wortelen, B.: The human efficiency evaluator: a tool to predict and explore monitoring behaviour. Kogn. Syst. **2017**(1) (2017)
17. Gruhn, P.: Human machine interface (HMI) design: The good, the bad, and the ugly (and what makes them so) (2011)
18. Hegarty, M.: The cognitive science of visual-spatial displays: implications for design. TopiCS **3**(3), 446–474 (2011)
19. ISO - International Organization for Standardization: ISO 9241: Ergonomics of human system interaction (2006)
20. Kosslyn, S.M.: Graph Design for the Eye and Mind. Oxford University Press, Oxford (2006)
21. Mackinlay, J.: Automating the design of graphical presentations of relational information. ACM Trans. Graph. TOG **5**(2), 110–141 (1986)
22. Mbaki, E., Vanderdonckt, J., Guerrero, J., Winckler, M.: Multi-level dialog modeling in highly interactive web interfaces. In: ICWE 2008 Workshops, 7th International Workshop on Web-Oriented Software Technologies - IWWOST 2008 (2008)
23. Meirelles, I.: Design for Information: An Introduction to the Histories, Theories, and Best Practices Behind Effective Information Visualizations. EBSCO ebook Academic Collection. Rockport Publishers, Beverly (2013)
24. Meixner, G., Paternó, F., Vanderdonckt, J.: Past, present, and future of model-based user interface development. i-com J. Interact. Media **10**(3), 2–11 (2011)

25. Marine Environment Protection Committee: 2016 guidelines for the development of a ship energy efficiency managment plan (SEEMP). Resolution MEPC.282(70), International Maritime Organization (IMO), 28 October 2016. www.imo.org (search: "MEPC.282(70)"), annex 10 to the report of the Marine Environement Protection Committee on its 70th Session

26. Nielsen, J.: Usability Engineering. Interactive technologies. Morgan Kaufmann, Burlington (1994)

27. Ostendorp, M.C., Feuerstack, S., Friedrichs, T., Lüdtke, A.: Engineering automotive HMIs that are optimized for correct and fast perception. In: Proceedings of the 8th ACM SIGCHI Symposium on Engineering Interactive Computing Systems, EICS 2016, pp. 293–298. ACM, New York (2016)

28. Ostendorp, M.C., Friedrichs, T., Lüdtke, A.: Supporting supervisory control of safety-critical systems with psychologically well-founded information visualizations. In: Proceedings of the 9th Nordic Conference on Human-Computer Interaction, NordiCHI 2016, pp. 11:1–11:10. ACM, New York (2016)

29. Paternò, F.: ConcurTaskTrees: an engineered notation for task models. In: Diaper, D., Stanton, N. (eds.) The Handbook of Task Analysis for Human-Computer Interaction. Lawrence Erlbaum Associates, Mahwah (2004)

30. Rosenholtz, R., Li, Y., Nakano, L.: Measuring visual clutter. J. Vis. **7**(2), 17–17 (2007)

31. Shneiderman, B., Plaisant, C., Cohen, M., Elmqvist, N., Jacobs, S., Diakopoulos, N.: Designing the User Interface: Strategies for Effective Human-Computer Interaction. Pearson Education Limited, London (2016)

32. Tidwell, J.: Designing Interfaces: Patterns for Effective Interaction Design. O'Reilly Media, Sebastopol (2010)

33. Treisman, A.: Preattentive processing in vision. Comput. Vis. Graph. Image Process. **31**(2), 156–177 (1985)

34. Tufte, E.R.: Envisioning Information. Graphics Press, Cheshire (1990)

35. Viana, M., Hammingh, P., Colette, A., Querol, X., Degraeuwe, B., de Vlieger, I., van Aardenne, J.: Impact of maritime transport emissions on coastal air quality in europe. Atmos. Environ. **90**, 96–105 (2014)

36. Ware, C.: Information Visualization: Perception for Design. Interactive Technologies. Elsevier Science, Amsterdam (2004)

37. Ware, C.: Visual Thinking: For Design. Morgan Kaufmann Series in Interactive Technologies. Morgan Kaufmann, Amsterdam (2008)

38. Wertheimer, M.: Laws of organization in perceptual forms. In: Ellis, W. (ed.) A Source Book of Gestalt Psychology. Kegan Paul, Trench, Trubner & Company (1938)

39. Wickens, C.D., Helleberg, J., Goh, J., Xu, X., Horrey, W.J.: Pilot task management: testing an attentional expected value model of visual scanning. Technical report ARL-01-14/NASA-01-7, University of Illinois, Aviation Research Lab, Savoy, IL, November 2001

40. Wickens, C.D., Hollands, J.G., Banbury, S., Parasuraman, R.: Engineering Psychology & Human Performance. Taylor & Francis, Boca Raton (2015)

41. Wickens, C.D., McCarley, J.S.: Applied Attention Theory. CRC Press, Boca Raton (2007)

42. Wolfe, J.M.: Guided search 2.0 a revised model of visual search. Psychon. Bull. Rev. **1**(2), 202–238 (1994). https://doi.org/10.3758/BF03200774

43. Woods, D.D.: Visual momentum: a concept to improve the cognitive coupling of person and computer. Int. J. Man Mach. Stud. **21**(3), 229–244 (1984)

44. Garner, W.R., Clement, D.E.: Goodness of pattern and pattern uncertainty. J. Verbal Learn. Verbal Behav. **2**, 446–452 (1963)
45. Zhang, J.: A representational analysis of relational information displays. Int. J. Hum. Comput. Stud. **45**(1), 59–74 (1996)