Visualizing Workload and Emotion Data in Air Traffic Control: A Tool to Support the Supervisors Awareness of a Complex Situation

Linda Pfeiffer*, Tabea Sims‡, and Paul Rosenthal§

*Institute of Data Science, German Aerospace Center DLR, Jena, Germany
Email: Linda.Pfeiffer@dlr.de

‡ Faculty of Computer Science, Chemnitz University of Technology, Chemnitz, Germany
Email: Tabea.Sims@informatik.tu-chemnitz.de

§Institute of Computer Science, University of Rostock, Rostock, Germany
Email: research@paul-rosenthal.de

***Abstract*—A supervisor, working in an area control center in air traffic control has direct impact on the performance of the air traffic controllers. Thus, he/she has to be well informed about the situation within the center. In order to support the supervisors work, researchers try to measure the air traffic controllers' cognitive workload and stress. Within this paper we present a visual tool to support the supervisors awareness about this data. We outline the whole design process including preliminary studies, the iterative process leading to a final prototype, and its evaluation. We propose a design divided into three views, to serve the supervisor's information needs. Using cumulated color values to represent time series, seems to be a promising approach for getting a quick overview on the whole situation.**

Keywords–air traffic control supervisor; visualization; workload data; emotion data; decision making.

I. Introduction

Usually known for having one of the most stressful jobs or causing huge delays when being on a strike, air traffic controllers provide a safe, orderly, and fluent handling of the air traffic. In this regard, tower controllers coordinate departing and landing aircraft on airports. On the other hand, airborne traffic is constantly monitored, managed, and sustained by area center controllers. In order to overview the whole air space, it is divided into sectors, where each one is overseen by two controllers. In fact, this job can be very demanding, so someone is needed to keep an overview of what is going on across sectors and prevent them and the assigned controllers from getting overwhelmed.

This task goes to the air traffic controllers' supervisors, who administer air traffic on a bigger scale, mostly by supporting and directing controllers. Since the air traffic controller's job is taxing and his/her emotions may have a big impact on his/her condition, their supervisors want to know about the controllers' mental states. This way, they are able to balance out their workload and to offer better support to them.

At ACHI 2017 [1], we already presented our analysis of the supervisors' work and decision processes as well as the resulting requirements for a workload and emotion data display. In this paper, we present an extended analysis together with the whole development process and our final visualization.

In Sections II and III we introduce the related research project and the application area. The design process and methods are presented in Section IV, followed by a description of our preliminary studies and the resulting design requirements in Sections V and VI. We present the final design in Section VIII and conclude in Section X.

II. StayCentered project

The main goal in air traffic control is to assure safe, orderly and fluent handling of the air traffic. This is a highly demanding task. Thus, the project "StayCentered - Methodenbasis eines Assistenzsystems für Fluglotsen (MACeLot)" at the Chemnitz University of Technology aims for giving support to air traffic controllers in stressful situations. The resulting system should be capable of identifying the emotional and cognitive state of the air traffic controllers. In addition, it should be able to simulate future states in relation to projected air traffic some hours in advance [2].

Galvanic skin response, facial action coding, body posture, vocal properties, eye movements, and pupil dilation are recorded. These measurements are used to infer emotion valence, arousal level, and cognitive load. In this work, we will not discuss limitations or uncertainty of these measurements or used algorithms but postulate that such data can be gathered as exact as needed. The assistance of the air traffic controllers should be realized by self-adapting interfaces [3] and by providing the information to their supervisors. Such a visualization of the controller's emotional and cognitive state may support the supervisors decision upon the opening of a sector, in order to reduce the controller's workload.

III. The Air Traffic Control Supervisor

Before going into detail of our work, we want to give a short description of what an air traffic control supervisor does and what tools and general environment he is provided with. As a superior and shift leader of air traffic controllers, the main work of a supervisor is to manage assignments and shift-structures of controllers, while regulating their workload by handling air traffic flow across sectors.

The regulation of traffic flow is done by splitting or merging sectors, permitting or forbidding special maneuvers (like, e.g., skydiving or air force trainings) and in extreme cases even regulating sectors by setting a maximum number of allowed planes and rejecting any exceeding traffic, which usually leads to delays. However, the amount of traffic that can be handled is limited by the number of present controllers and the supervisor's job is to assign them in such a way that simultaneously no employee is overexerted and air traffic can flow undisturbed.

Beside the obvious impact of the supervisors' work to air traffic safety by organizing shifts and regulating greater traffic flows, their work is crucial to air traffic safety in several non-obvious aspects. In general, a good supervisor-employee relationship has direct impact on the employees' (air traffic controllers) work motivation [4]. The supervisor's emotional intelligence on both, a personal and a group level, improves collaboration and communication, as well as the ability for emotional contagion increases efficiency and decreases the conflict potential in teams [5][6]. An investigation by Broach et al. [7] showed a correlation between the supervisor-controller-ratio and the number of errors made by air traffic controllers. The latter suggests that a sufficient number of supervisors should be present. Supervisors in the German air traffic control centers usually do not work alone. During a day shift there are two to three supervisors present. They divide their work by region, where each one can operate independently but still help each other if necessary.

The actual working place is located within the area air traffic control center, often in the middle of the operations room or a little elevated, providing a good overview. Other non-controller positions can be found alongside. These include technical surveillance, data assistants, flight data agents/operators, flow management position, and the technical supervisor.

Each supervisor's workplace comes with a computer, equipped with two monitors, serving as their main working tool. Among other work specific software, the most important one is a shift management program, where active sectors are scheduled, air traffic controller's shifts are organized, and controllers can be notified by publishing the current plan on a separate screen. In addition supervisors have several information systems, specific to each center, as well as communication devices, such as an email program, a land-line telephone, and a direct-dial telephone.

IV. Design Process

We followed an iterative design approach, which has been proven as good practice in the research fields of human-computer interaction and visualization [8][9][10]. Throughout the whole design process, we tightly cooperated with air traffic supervisors from the German air traffic service provider DFS - Deutsche Flugsicherung GmbH.

The first phase of the our design process consisted of some preliminary studies. We did observations in the area control center, we interviewed seven supervisors, and did a review of related work. Our main objectives in this phase were: getting to know the users, their work, and their decision-making process in order to define the requirements to our visualization. This phase was followed by the iterative process of designing, prototyping, evaluating, adapting the requirements, and redesigning. The prototype's fidelity grew with every iteration. We started by presenting simple sketches to the supervisors and ended up with a technical interactive prototype.

V. Preliminary Studies

Main goal of the preliminary studies was the identification of design requirements, such that the intended visualization serves the supervisors needs best. This implies, first establishing an understanding about the supervisors' tasks and their way of decision making. Furthermore, we wanted to find about the qualities of stress and emotion data needed by the supervisors and we wanted to learn from the use of the current interfaces and from its positive and problematic impact on the supervisors' work.

A. Methodology

We did two full-day observations at the end of September 2016 of the supervisors' working place at the area control center in Munich, Germany. The researchers had the chance to ask clarifying questions during the observation and collected the data by handwritten notes. Furthermore, we interviewed seven supervisors, each interview lasting 18 - 51 minutes, on their decision-making process. This included their information needs, the role of the air traffic controller's workload and emotional situation in their decision-making process, as well as the data's level of detail favored by the supervisors.

During the interview, we invited the supervisors, to sketch their decision process and the considered information. This was done to make them reflect their decisions in a structured way and for not forgetting anything. The data was audio recorded during the interviews and transliterated. For analysis purposes, the data was coded and categorized. We did not use a standardized coding scheme, because of the exploratory nature of the research questions.

B. Supervisor's Tasks

Our first step for designing a well suited stress and emotion visualization, was to understand the supervisor's tasks. The supervisor's task area includes tasks concerning ongoing operations and tasks beyond. Beyond ongoing operations, the supervisors mainly fulfill tasks in human-resources management. Each supervisor is responsible for 15 to 20 air traffic controllers. Additionally, they may have optional special tasks, like the participation in research projects or committees. As a visualization of workload and emotion data is most beneficial in ongoing operations we concentrated on this part of the supervisor's work.

The principal task of a supervisor is to keep ongoing operations fluently going. This means, doing everything, such that the circumstances allow air traffic controllers to handle air traffic safely, orderly, and fluently. The supervisors called themselves well-paid secretaries, in order to express that they are responsible for every concern in the operation room. The principal task can be divided into sub tasks. At the one hand there are somehow formalized tasks and at the other hand more informal tasks. Even the formalized tasks are rarely provided with clear instructions, allowing for a multitude of alternatives. The only task with clear instruction is the documentation of events, which happened during the supervisors shift. This task is not very favored by the supervisors. According to them, they are spending too much time documenting insignificant events. This may be a task, which is very suitable for automation.

Each day, there are two briefings scheduled, wherein the supervisor updates the controllers. Most of their time the supervisors spend on planning the day's shift schedule and solving occurring bottlenecks and problems. We observed two ways of planning: some supervisors prefer planning of the whole shift and changing the schedule if necessary. Others avoid this strategy, because of the numerous changes and prefer a piecewise planning. A piece has often a size of about two hours. On the other hand, there is the solving of bottlenecks and problems, which may put the safety of aircraft in risk. This is a very creative task, because of the numerous possible solutions. Sometimes supervisors even consult their colleagues and the air traffic controllers on this behalf.

The informal tasks are rather some kind of good practice and skills. Their implementation depends on the individual supervisor. Over the entire shift, even if the supervisor seems to relax a while, he is observing the current situation at the operation control room and looking for abnormal situations. Each abnormality may induce safety problems. A controller speaking to the technician, holding his interaction device in his hand, may be an indicator for a malfunction of equipment. Especially, if a controller switches his status on the status display to a warning level, the supervisor will go to the controllers working position to assess the situation. In order to assure the controller's ability to work under pressure, the supervisor is trying to determine their daily emotional state and basic stress level. Thus, he can consider this information in the shift schedule or, in extreme cases, advise a controller to rest.

C. Collaboration Aspects

Generally, the supervisor tries to prevent controllers from stress by using formal means, like splitting up a sector, by regulating the number of aircraft that are allowed to enter the sector, or by using more informal means like warning the controllers of a short high traffic peak. Usually the supervisor complies with flat hierarchies. This has practical effects, like asking the controllers for their opinion about suggested solutions or by considering the controllers' wishes in the shift schedule. They are also trying to support the air traffic controllers on their issues, even if they are not in their field of responsibility, e.g., they check for the location of a meeting.

Altogether, a supervisor needs interpersonal skills, he should be sensible to individual communication patterns. It is a well known issue in leadership studies that the political skills of a supervisor may have a positive impact on the team performance [11][12]. The ability to identify the others' needs by observation and to attune to divers social situations, allows for better communication and improvements in supporting the controllers. However, the supervisor is not just concerned about the controllers issues. He/she also tries to support his colleague. He/she stands in for his colleague during breaks and reminds of important tasks. This is appreciated by the other supervisor.

D. Decision Challenges

There are some typical decisions a supervisor has to face in his daily work. Besides the decision of the briefing topics, the most critical decisions are made in the tasks of planning the day's shift schedule as well as preventing and solving bottlenecks and problems. The planning task includes the decision on which controller has to work at which position. This decision is guided by several constraints. The solution should be safe as well as cost efficient. This means that safety rules need to be met, e.g., considering breaks, two controllers should be responsible for one sector, assuring that no controller is overstrained. Simultaneously, every controller should be busy, taking special tasks, trainings and so on into consideration.

Potential bottlenecks and problems may be caused by external demands or extraordinary circumstances. The supervisor has to decide whether to allow for external demands, like photo flights, gliding flight areas, or planned detonations. Other external demands are obligatory (e.g., activation of special air spaces or military trainings) and the supervisor has to decide on a suitable reaction to this. The decision on the reaction to extraordinary circumstances includes malfunction of the technical equipment or potential overloads in traffic quantity, which may result in splitting up a sector or a regulation of the number of aircraft. In addition, there may occur staff concerns, like illness or spontaneous meetings, which force the supervisor to find alternative solutions, like obtaining a spare controller. In addition, a multitude of other troubles may occur, e.g., fire alarms.

None of these decisions can be seen separated. Each decision on one variable of the system has impact on another and may result in new decisions to be made. For instance, the decision on splitting up a sector entails a change in the shift schedule, On the other hand, regulating the number of aircraft in one sector increases the number of aircraft in other sectors.

E. Problem Solving Variables

Based on their experience, the supervisors identify variables that are affecting the capacity of a sector or constraining their scope of action. Information upon these variables should be available to the supervisors. They should know about the standard sector plan that tells which sectors should be open. It is based on statistics of the past years and is the basis for the shift schedule. The available staff is a framing variable for the scope of action. This includes the air traffic controllers on duty as well as controllers, who are around but fulfilling other tasks (paperwork, trainings, meetings, and so on). The latter may be consulted in the case of staffing shortage. However, usually the information is hardly available.

Also, the staff's condition is a factor to the capacity of a sector: their daily performance, fatigue, and their satisfaction. To keep satisfaction high, the supervisors try to assure that the controllers are facing varying demands, which they are sharing a position with someone with whom they accord, and that some of their preferences are be met. The information about alternative tasks, a controller has to do, is necessary in order to assure cost efficiency. However, it is often incomplete.

The main factors on a sector's capacity are the expected traffic load and the weather conditions. The weather forecast is needed two hours in advance. However, up to this day, weather predictions are not always reliable. The quantity of the expected traffic is also automatically predicted, by considering the aircrafts' flight plans. In consequence, a two-hour forecast is highly unreliable. Still, with each minute this estimation is getting more precise. In contrast to the quantity, predicting the traffic quality is even challenging for an expert. Handling 15 aircraft flying straight in a line may be much more less demanding than seven aircraft climbing and descending with different headings.

Extraordinary circumstances as safety issues in the area control center (e.g., fire alarms) and technical concerns (malfunction of equipment or the use of backup systems) may reduce the capacity of a sector extremely. Also, visual clutter on the radar screen, coming from a lot of aircraft below or above the sector, is limiting the sector's capacity. A variable that is consulted rather unconscious is the own constitution. It has some effect on the consideration of external demands or controllers' wishes. When a supervisor has not the full overview of the current situation, he is not willing to generate any additional workload to his controllers or himself.

This listing of variables is an attempt to get a structured view on the variables needed by the supervisors and is not complete, as every situation is unique and may require other information. Altogether the supervisors are facing complex problems [13][14]. They have to address many variables that are interrelated, the time for decision making is limited, and some events occur unexpectedly (illness, external demands, emergencies, etc.). They have to outweigh different goals (safety, cost efficiency, controller's satisfaction) and the information needed is incomplete or sometimes not reliable. They make decisions for the future, based on current data, personal heuristics, and unreliable predictions. The heuristics they use for problem solving are based on their experience. They are able to anticipate the effect of the variables on the sector's capacity and they know how to weigh the influence of a variable to a specific type of problems.

F. Workload and Emotions

As we are interested in designing a stress and emotion visualization, we wanted to have a further look into the role of stress, workload, and emotion data. As stated above, the supervisors already take the controller's basic stress level and severe emotional states into consideration. By now, they have to look for this information during conversations with the controllers.

The information on the controllers workload, stress, and boredom is considered as useful in two terms: A prediction of the workload in a specific sector, is seen as alternative approach to current traffic quantity predictions and may, thus, help with the planning of the shift schedule. Moreover, an information upon the controller's former and current stress level may help in assigning suitable tasks to the controller. The relevance of detecting stress seems to be much more important than boredom. During periods of boredom, controllers lean back and start chatting. Thus, the supervisors stated that they can easily observe boredom. In contrast, stress is sometimes not even recognized by the controllers themselves.

The use of emotion data is seen much more controversial than the workload data. On the one hand, the emotion data may be useful when the controller's emotional state in extreme situations hinders him from doing his job. On the other hand, they refuse using this data. This fact arises from the expected professionalism, from the fear of treating other controllers unfair, when someone is pretending to be in a bad mood, and from concerns about privacy. Both, showing the data linked to an individual controller and showing it linked to a sector, may be useful. Sector-related data is similar to current traffic quantity predictions, whereas individual data may help by assigning each controller a suitable task. There are ethical concerns about showing the person-related data.

G. Supporting Tools

As already mentioned in Section III, the supervisors' working places offer a multitude of tools and information systems to support their decision making. The three main tools are the shift management program, an overview of the planned controllers, and a notepad. This physical notepad is an important tool, since it allows for quick note taking and, thus, remembering important tasks and lines of thought. This is necessary, since the supervisor's thoughts are often interrupted by incoming demands and information.

These main tools are complemented by a multitude of information and communication systems, where information can be retrieved and is pushed through. Conspicuous about the interfaces was the importance of clear arrangement. Consistency of representations and data between systems and tools are as important as unambiguous interaction strategies [15][16][17]. For example, using different time zones or coding same meanings differently would lead to misunderstandings and inefficient occupation of cognitive resources. In fact, the supervisors had to transfer data from one tool to another manually, which took time and cognitive resources. Furthermore, several similar interaction devices, each of them belonging to another system, were confounded, slowing the progress down.

VI. Problem Definition and Design Requirements

As we had seen in our preliminary studies, air traffic control supervisors are dealing with complex problems in their daily work [1]. These problems cannot be automatized due to the high complexity and are best suited to be solved by humans, supported by the computer [18]. In the current problem solving process, workload and emotion data are already considered by the supervisors. Emotion data is currently accessed by observation and conversation with the controllers. Workload data is currently accessed by observation and by a prediction of the number of aircraft that will enter the corresponding airspace.

This assessment could be simplified by the StayCentered system by measuring cognitive load and emotion data automatically. The model measures an index for cognitive load, one for arousal, and one for emotional valence. This mental state is measured for every single air traffic controller. From this data, a combined value for each sector is generated as well as a prediction into the future. Thus, there is a triple of scalar values (workload, arousal, valence) available for each time step.

The time frame, shown by the visualization, can not be infinite. During the interviews the supervisors stated that they need about two hours of the past values and ten minutes up to four hours of the predicted values. As the observed planning period has a length of about two hours, we decided to set the upper limit of the predicted values there.

R1 In order to consider the controllers state during scheduling, the workload data needs to be accessible in an overview over a period of time ranging from two hours in the past to two hours in the future.

The most important information to the supervisors is whether a controller is able to fulfill his task. As seen in the preliminary study, the use of emotion data is seen dubiously. So we decided to restrict emotion data to the data that indicates the controllers ability to work. First, we thought about restricting emotion data just to extreme situations. Later, we decided just to show the arousal level.

The performance in solving a medium to complicated task, e.g., the task of an air traffic controller, is known to be highest at an intermediate arousal level, while it is descending at high and low arousal levels. An overview of this Yerkes-Dodson law was presented by Teigen [19]. Thus, the arousal values are a good indicator for the controller's ability to work, while indicating just minimal information on the emotional state.

R2 Emotion data should be restricted to the arousal value or to only extreme situations, in order to address ethical issues.

R3 Extreme situations should be visible at a glance. Extreme stress, very high boredom, as well as extreme negative emotions may hinder the air traffic controllers' work, so that intervention of the supervisor may be appropriate.

Thus, the data is reduced to a pair of scalar values per air traffic controller and sector, evolving in time. Additionally, there exists an assignment of controllers to the working positions. The supervisors' tasks cover looking for bottlenecks, concerning the workload in the open sectors, as well as looking for a suitable controller for taking over a position. Consequently, the workload data is needed in both notional categories.

R4 The representation of the data should support the supervisors notional categories. The data should be available related to the individual controller as well as to the working positions.

Beside the requirements, resulting from our preliminary studies, there are also design principles arising from the human cognitive and perceptual abilities.

R5 The visualization must concentrate on a minimal set of primitives to produce an expressive and effective visualization [20] with minimal disturbance of the work flow. All important features should be easily identifiable and all visual elements should have an important meaning. Color should only be used when really needed to highlight very important features and taking into account human visual perception [21].

VII. Related Work

In addition to a long tradition on visualizing time-oriented data, an overview was given by Aigner at el. [22], there exists also some work that is specialized on emotion data. Most of the work deals with emotion data deriving from the context of social media, including work suitable for presenting a single emotion or emotion data of multiple persons, as well as work related to time.

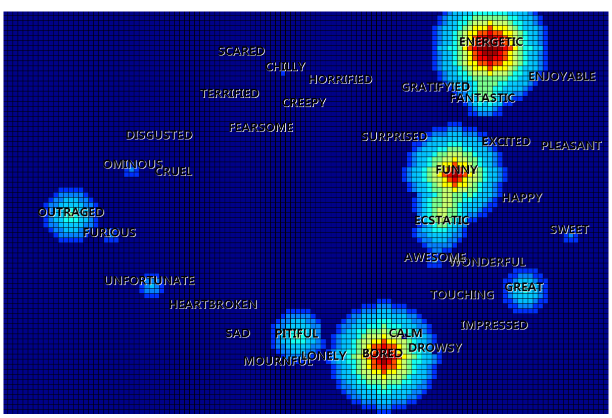
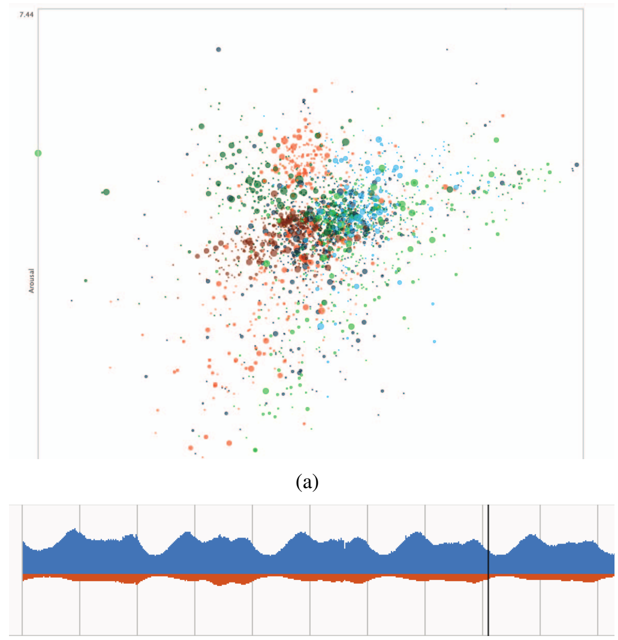


Figure 1. A visualization of the value and amount of emotional connections to a movie as a heat map on a valence-arousal coordinate system [23].

For example, Ha et al. [23] visualize sentiments connected to movies with their focus on easy recognition of clusters and intricate network structure. The visualization in Figure 1 is a detail view for one node within that structure, showing emotions connected to a single movie as a heat map on the valence-arousal coordinate system. Additionally, some points in the coordinate system are labeled with the common name of that emotion.



(a)

(b)

Figure 2. A sentiment visualization, concerning a specific topic on twitter [24]. (a) The emotion is shown as arousal and dominance on the axes, the valence is color coded, and its amount is shown as the size of each dot. (b) The number of positive and negative posts is shown over time.

Steed et al. [24] constructed a similar view, shown in Figure 2(a), within their visual application to dynamically analyze twitter sentiment. Their coordinate system uses arousal and dominance as axes while showing valence in color (orange for negative and blue for positive). However, this again is just an additional display, next to a geographical depiction of tweets, while the main view is a visualization of the amount of tweets (divided in binary valence) over time, depicted in Figure 2(b). This view is designed interactively to select time intervals for further inspection in the other visualizations.

Working on the same problem, of analyzing twitter sentiments over time, Wang et al. [25] propose a solution, integrating valence, arousal, and time into a single visualization, as shown in Figure 3. Each ring in the circle represents a different time step (designed to resemble the view in a tunnel), while the amplitude within the ring is defined by the valence-arousal coordinate system. Since the curve is additionally color coded by valence and arousal, the rings could also be displayed as straight lines. This might be a little less disorienting for some users. Thus, the amount of tweets, currently shown by the bar on the left, could also be aligned.

A completely different application is that of Cernea et al. [26], who designed an emotion visualization on touch displays. The design gives users direct feedback in the color and shape of the selection highlighting. However, they also created a separate view, showing the different emotions of the

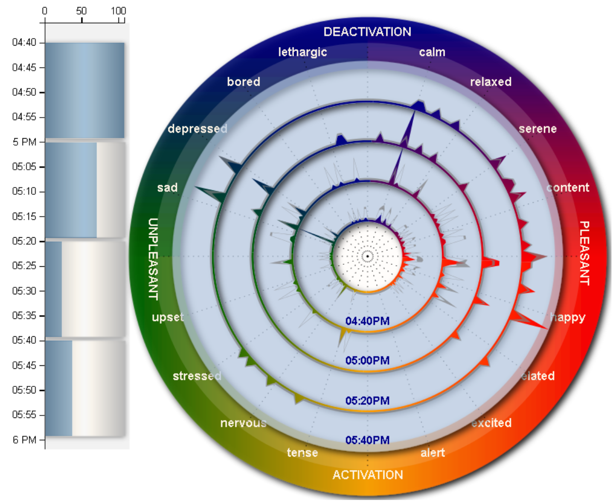


Figure 3. Another twitter-sentiment visualization [25]. Each ring is a step in time, while the curve depicts the valence and arousal.

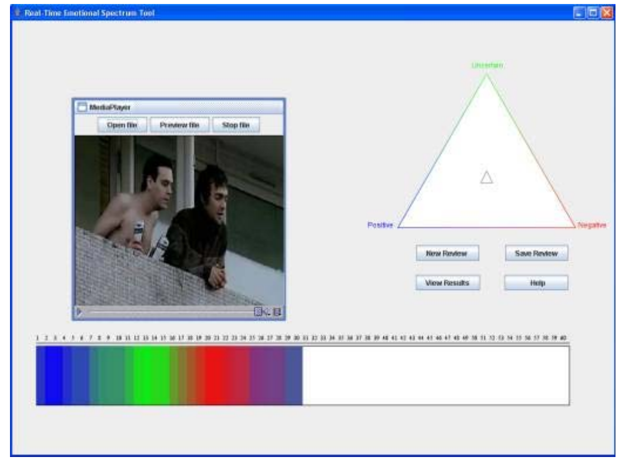


Figure 5. An emotion input interface concerning a video [27]. The bottom bar shows collected emotions on a time line, while the color shows the value of each step.

touch events over time, in order to let users reflect and compare themselves with other users (Figure 4). They display time on the horizontal axis while valence is shown in direction as well as size of the bars and arousal is color coded (blue refers to low arousal and red represents high arousal).

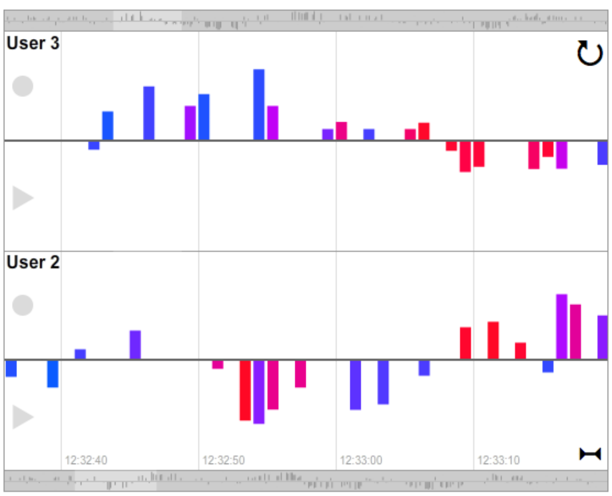


Figure 4. An emotion visualization over time for individual users by Cernea et al. [26].

The approach of Ohene-Djan et al. [27], illustrated in Figure 5, was developed to directly track user emotions while they are watching or listening to some kind of media playing. Thus, the media player box on the left. At the triangle on the right-hand side, the recipient is supposed to report current emotions by accordingly placing his mouse within that triangle. Thus, valence and uncertainty are measured. The input gets tracked and displayed in the bottom bar as visual feedback of the opinion development. Time is the horizontal axis in the bar while the color gets computed as the distance to each corner of the triangle, mapped on the opacity of one of the base colors in RGB. It is also possible to render the results of multiple users on top of each other, showing the mean opinion of all of them.

Yet another approach is to use some kind of emoticon, like the manikins in Figure 6, designed for visual feedback, e.g., in questionnaires. Sonderegger et al. [28] even found in a user study, comparing different pictographs, that the ones shown in Figure 6 could be further enhanced by using an animated heart as arousal indicator instead of the rather abstract shape depicted here. This might be the most intuitive way for emotion visualization. However, research shows that the bigger the set size and complexity of the icons, the harder they are to identify, even more so when they are rather similar [29][30].

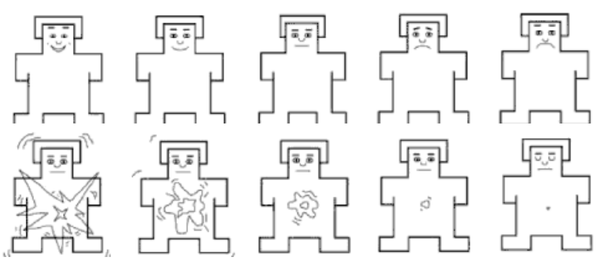


Figure 6. An emotion visualization in manikins [28]. The facial expression shows valence while the shape on the chest depicts arousal.

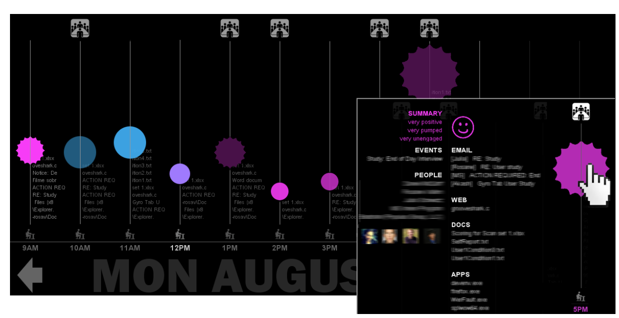


Figure 7. A visualization of emotions during a work day for personal reflection [31].

Finally, there is also the work of McDuff et al. [31], who create a very detailed visualization of emotions, connected to work. Their goal is to use the device for personal reflection over longer periods of time. The result is shown in Figure 7. They encode emotions in the bubbles with color showing valence (pink is positive, purple neutral, and blue negative), shape showing arousal (calm is round and aroused is pointy), and opacity showing engagement. Work related information is indicated by the height and size of the bubbles (desktop activity), by little icons at the top and bottom of each bar (meetings), and text (further work related information).

TABLE I. OVERVIEW OF ENCODINGS OF DIFFERENT VARIABLES IN THE DISCUSSED EMOTION VISUALIZATIONS.

Figure	Valence	Arousal	Time	No. of Persons
1	horizontal axis	vertical axis	-	multiple
2(a)	color	vertical axis	-	multiple
2(b)	direction + color	-	horizontal axis	multiple
3	horizontal amplitude + color	vertical amplitude + color	rings (tunnel view)	multiple
4	size + direction	color	horizontal axis	single
5 (Triangle)	horizontal axis	-	-	single
5 (Bar)	color	-	h. axis	single
6	facial expression	shape	-	single
7	color	shape	horizontal axis	single

The summary in Table I gives an overview of the just described visualizations and how different variables are encoded.

VIII. Final Design

In this section, we describe the design decisions we made concerning our workload and emotion visualization for air traffic control supervisors. As we have seen in our preliminary studies, air traffic control supervisors consider a lot of information for making their decisions. This information is presented on different screens and media. Hence, it is important to design the new visualization in a way, which it is minimally disturbing but clearly communicating critical situations (R3). Consequently, we decided to split the visualization in an overview and a detail view. The overview is visible all the time and details will be available on demand.

Figure 8. Color scale used to indicate a value too low (blue) or to high (red) for concentrated work.

The main purpose of the overview view is to give a quick summary on the situation in the center and to clearly communicate critical situations (R3). Since color is best suited for being perceived in peripheral vision [32], we decided to color-code the workload and arousal levels. The chosen color scale (Figure 8) is diverging, ranging from blue (underload) over light gray (comfort zone) to red (overload), as proposed by ColorBrewer [33]. The comfort zone as well as the other elements are colored in light gray in order to highlight the most important information of the visualization, which are the extreme workload and emotion situations (R3, R5).

As pointed out by the experts, the overview should be kept as simple as possible (minimal set of primitives). Meanwhile should it not just indicate current critical situations but also future critical situations. Showing a time line for each sector is a quite complicated shape not suitable for a quick overview. Thus, we decided to include future critical situations into the color coding, by making the light gray fully transparent and adding an additional opacity value that is decreasing with time. Finally, we summed up all the future color values to the current color.

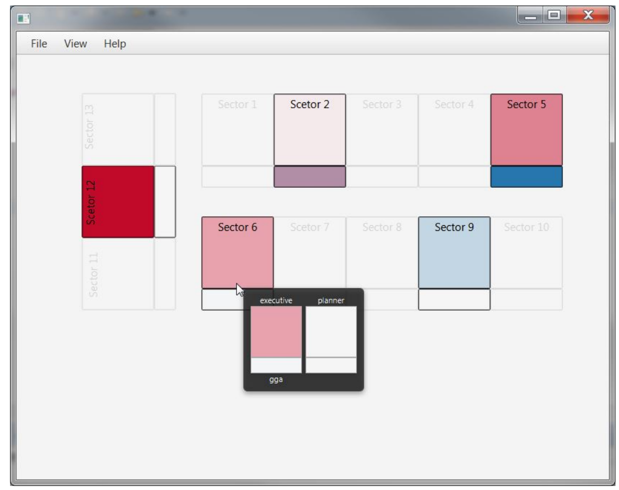


Figure 9. The center view provides an immediate overview of the situation in the center.

In order to create a visualization that matches the supervisors notional categories, we decided to put these color values on a map, showing the working positions in the center. On this map, illustrated in Figure 9, the supervisor sees at a glance the sector values. In addition, a mouse-over window shows the data of the air traffic controllers at the working positions. The colored symbols have the advantage that a smaller version of them can easily be added to a list (Figure 10). By adding the smaller symbols to the existing air traffic controller list in the supervisors planning tool, design requirement R4 is met. The overview data is now available related to the sectors on the map and related to the controllers in the list view.

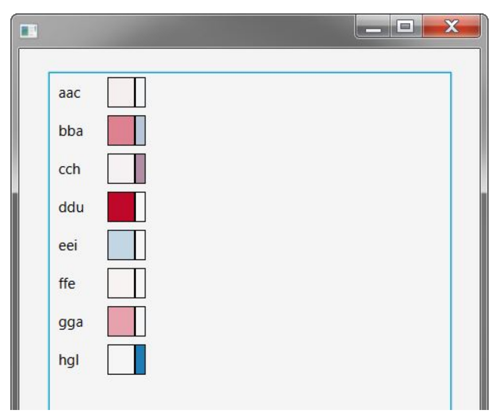


Figure 10. The symbols, known from the center view, can easily be integrated into a list of controllers (e.g. to the current planning tools).

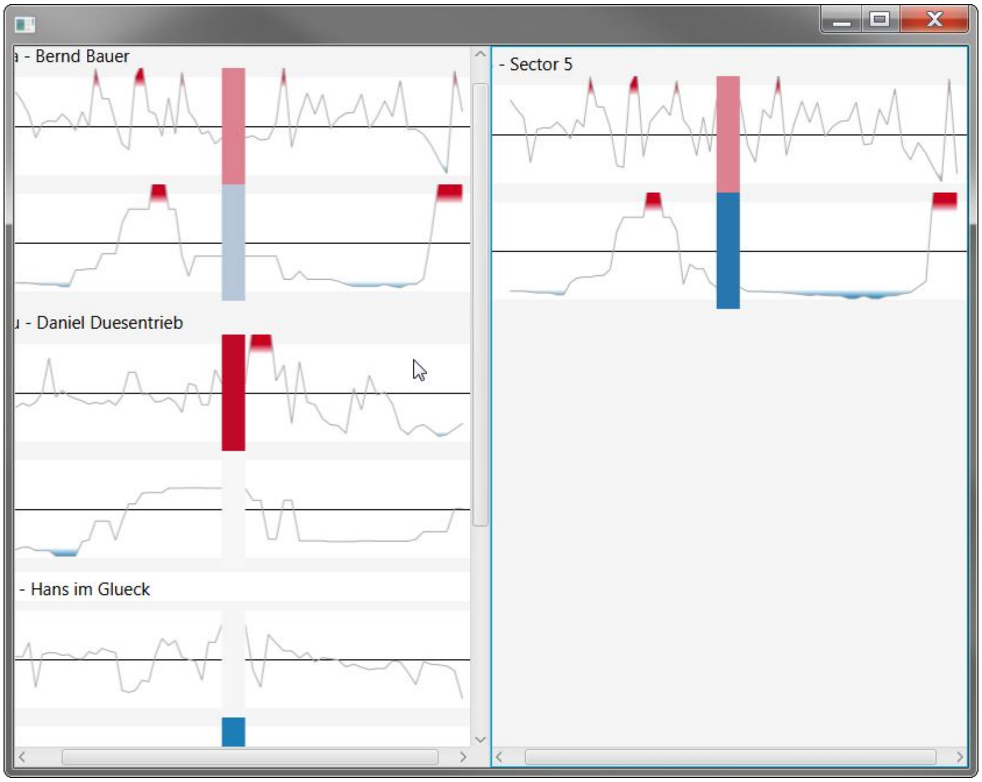


Figure 11. Line diagrams, supported by color highlighting in the critical areas allow for a more detailed situation analysis.

The overview views show the supervisor whether there is any critical situation that may require their attention. For a more detailed analysis of the situation, the supervisor can choose some controllers or sectors and show them in the detail view (Figure 11).

The detail view is showing the workload and arousal data for a period ranging from two hours in the past to two hours in the future (R1). For each chosen data set, there are two simple line diagrams shown. These allow the supervisors to follow the development of the values and to analyze them. This representation is enhanced by colored areas in order to highlight the extreme situations of under- and overload.

IX. Evaluation

Throughout the design process, we worked closely together with the air traffic control supervisors and adapted our design iteratively by considering the supervisors' feedback. Our final evaluation study touches upon two points. First, we wanted to ensure that the chosen design of the future stress values as color values is similarly perceived as usual visualizations of time series. Second, we asked the participating supervisors to assess the whole design, including center view, list view, and detail view.

For reasons of participant availability, we decided to perform the evaluation as an online experiment and online survey. To ensure that the supervisors were undisturbed during the experiment and that they are using similar hardware, they had to confirm that they were at a silent place, that they switched their phones off, and that their screen has at least a minimum resolution. In order to adapt the physical size of the visual stimulus, we adapted the size of the images to the size of a credit card as proposed by Woods et al. [34].

Due to the high specialization, it was only reasonable to include real air traffic control supervisors, which were hard to recruit. However, we were able to obtain six air traffic control supervisors, participating in the study. All of them were male. Their average age was 47.67 (SD = 3.93), they had on average 9.5 years (SD = 4.55) of experience as an air traffic control supervisor, and on average 20.83 years (SD = 5) of experience as an air traffic controller. Due to the small number of participants, quantitative results of the study should be interpreted cautious.

The following part of the section describes the two parts of the evaluation study, by outlining their methods and results, respectively. The section concludes with an overall discussion of the results.

A. *Comparison of Visualizations for Future Stress Values*

Within this part of the study, we compared different visualizations for the future stress levels in order to ensure that the interpretation of the color values does not differ significantly from other visualizations of time series. Such visualizations are more familiar to the supervisors, including the line graph, used in the detail view. The five visualizations that were compared are shown in Figure 12. The supervisors obtained an introduction into each visualization before they did the tasks with the respective visualization type. The ordering of the visualization types was randomized, in order to prevent sequence effects.

Supervisors need to assess the situation of a single air traffic controller reliably. In addition, they need to get a quick overview of the situation of multiple controllers. Consequently, these parts should get investigated in experiments. Furthermore, the user's satisfaction and perceived usability have an

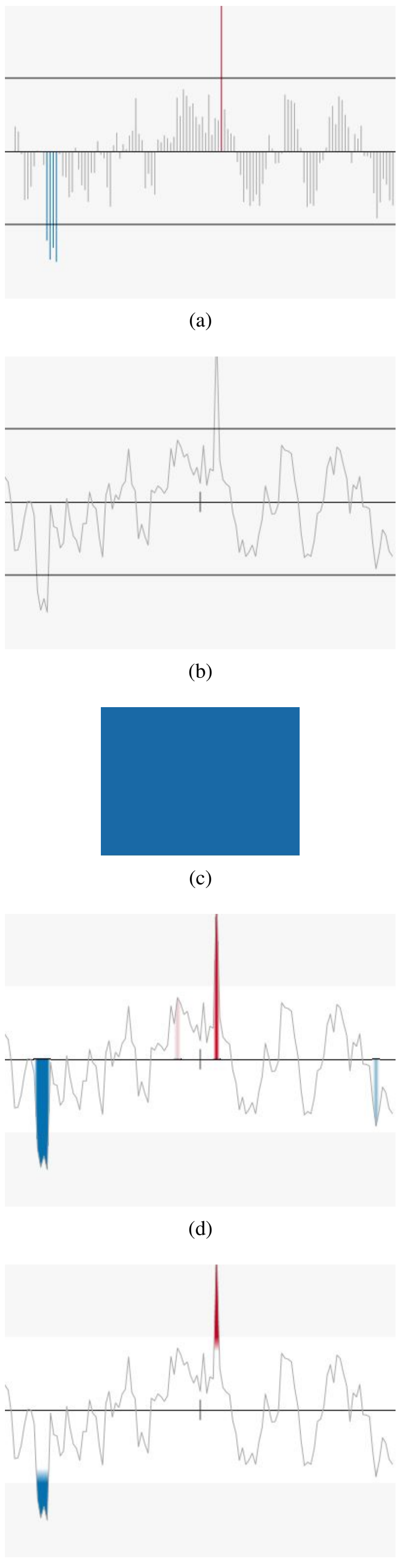


Figure 12. The visualization methods compared in the evaluation study: (a) A bar chart with colored bars when the stress threshold is exceeded. (b) A simple line graph. (c) The color value. (d) A line graph with fully colored areas till the base line when the stress threshold is exceeded. (e) A line graph with colored areas when the stress threshold is exceeded.

impact on the success of a visualization. In our experiment, we compared the perception and interpretation of single graphs and asked for the supervisors' subjective ratings.

1) Method: In order to judge the interpretation of a single time series (future values of a single air traffic controller) within a short time frame, we presented six scenarios to the supervisors. The scenarios were developed in cooperation with supervisor instructors. Each scenario was presented to the participants for a period of one second. Time of stimulus presentation was chosen to be longer than time to perceive and identify objects in pictures (300 ms according to Thorpe et al. [35]) but still limited in order to simulate a short glance on the plot in busy times. Afterwards, the supervisors were asked to assess whether they would want to inspect the situation further, whether they would instantly react to the situation, and how they would rate the severity of the situation on a 5-point-Likert-scale. Each scenario was rated by each supervisor in each visualization and they were presented in a randomized order.

The comparison part was concluded by asking participants to rank the five visualizations according to personal preference and clarity aspects.

2) Results: In order to compare the ratings across scenarios, we computed for each scenario the mean rating independent of visualization type. Relative ratings were formed through the difference between the mean rating of the scenario and the individual value. The box plots of the supervisors' relative ratings for the two reaction options are illustrated in Figure 13.

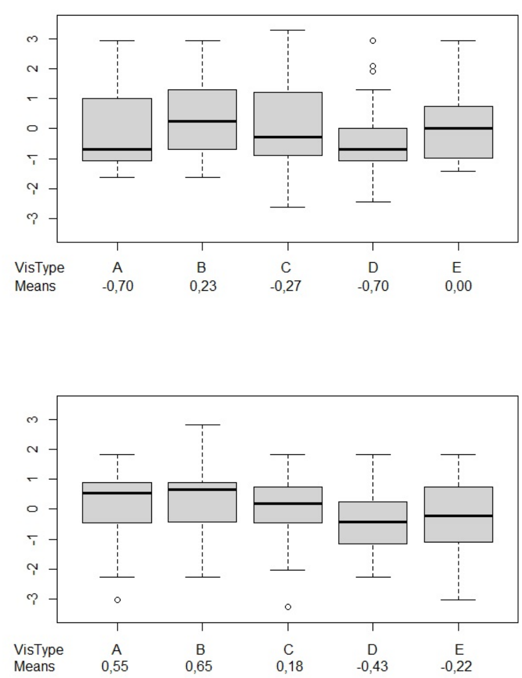


Figure 13. Box plots of the relative ratings of the supervisors for the option of further inspecting the situation (top plots) and for the option of instantaneous reaction to the situation (bottom plots) with respect to the five different visualization types. A relative rating of zero would indicate an optimal visualization.

As data did not meet the assumptions of parametric tests, we used the non-parametric Friedman test to compare adapted ratings between visualization types. The Friedman test about the related ratings for further inspection of the situation showed a slight significance ($\chi^2(4) = 9.6384, p = 0.04698 < 0.05$). Post-hoc tests were used with Bonferroni correction applied. The critical difference 37.66031 ($\alpha = 0.05$, corrected for the number of tests) was not exceeded in any case.

The relative ratings about instantaneous reaction on the situation differed significantly ($\chi^2(4) = 11.697, p = 0.01975 < 0.05$). But post-hoc tests did not show any significance.

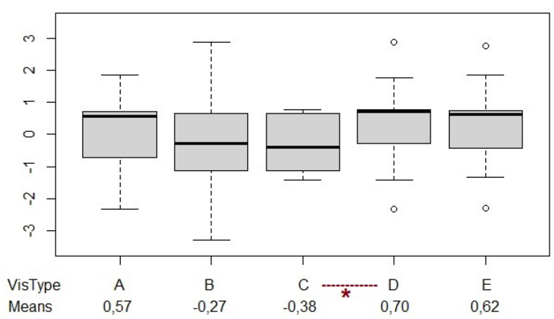


Figure 14. Box plots of the relative severity ratings (* - is significant) with respect to the five different visualization types.

Relative ratings about the severity of a situation were highly significant in Friedman test ($\chi^2(4) = 15.231, p = 0.004245 < 0.01$). Box plots of the respective relative ratings are depicted in Figure 14. Post-hoc tests only revealed a significant difference between visualization types C and D (difference = 38.0 > critical difference = 37.66031). In the condition using the color value (type C) supervisors systematically rated the situation as less severe than in the condition with visualization type D.

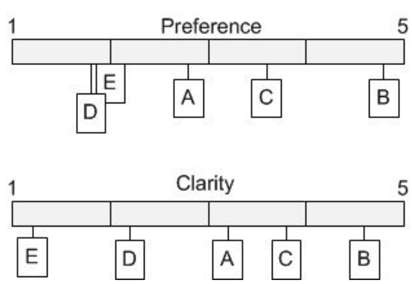


Figure 15. Mean ranks for supervisors' personal preference and perceived clarity with respect to visualization type.

When it comes to the ranking of the visualization types according to the personal preference, supervisors preferred visualization types D and E (mean ranks 1.8 and 2.0). They also considered these two as most clearly structured (mean ranks 2.2 and 1.2). In both rankings the color value (type C) was at the second lowest rank (mean ranks 3.6 and 3.8). The respective mean ranks are illustrated in Figure 15.

B. Assessment of the Overall Design

The second part of the study was an assessment of the proposed overall design, including all views presented in Section VIII. The supervisors rated the design according to several aspects and they had the chance to comment on the design.

1) Method: First, there was a description of the design, the supervisors had to read. Then, the rating was performed on a 5-point Likert scale, ranging from 1= "I completely agree to the statement" to 5 = "I completely disagree". The statements, presented in a randomized order, were the following:

- "I am able to identify sever situations quickly."
- "The design is easy to understand."
- "The design is well adapted to the requirements of my
- "The design is complicated."
- "I would like to work with the design."
- "The design contains too few information."
- "The design contains too much information."
- "The design is clearly structured." work."

The questionnaire ended by asking the supervisors for the advantages and disadvantages of the design. The answers were given in free text.

2) Results: The answers to the ratings are depicted in Figure 16. They were mainly neutral.

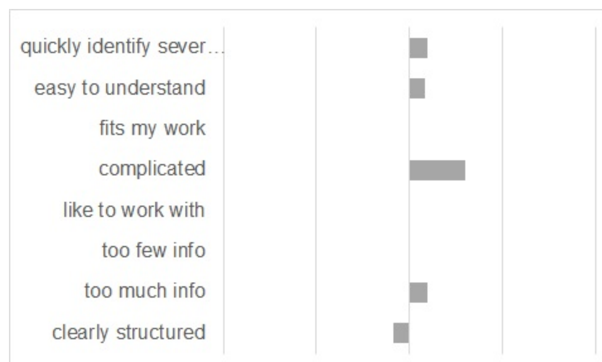


Figure 16. Resulting average ratings of the overall design on a 5-point Likert scale.

The comments revealed that most of the supervisors preferred the center view for getting an instantaneous overview of the entire situation and it was perceived as clearly structured. One of the supervisors perceived the ordering in working positions as useless, he preferred a list ordered by sector distribution. Opinions about the detail view were split. One half pointed out that it shows exactly the information needed. In contrary, the other half mentioned that there is too much information shown. Thus, they perceived the detail view as less intuitive and less clearly structured. There was only one comment about the list view, which described it as a good tool for doing the planning for the next hours. One supervisor perceived the usability of the overall design as complex.

C. Discussion

The impact of using the color value as an indicator for the future stress situation remains unclear. There was a slight underestimation of the severity of the stress of one controller compared to visualization type D. A slight adaption of the computation of the color value may counterbalance this effect.

This seems adequate, since the interpretation of the values did not differ significantly in any other case.

The comments indicate that the center view is well suited for a quick overview and, thus, the use of color values does not clutter the display. This indicates that the use of color values for visualizing future stress situations of multiple air traffic controllers may have advantages which outweigh the bad rankings for preference and clarity related to single plots. A further experiment about the perception and interpretation of multiple plots would give more insights.

The good ranks, that were achieved by visualization type F, confirm the use of this visualization type in the detail view. The divided opinion about the amount of information in the detail view confirms the choice to show this view on demand.

X. Conclusion and Future Work

Within this paper, we presented a workload and emotion data visualization for air traffic control supervisors and the related design process. Based on a detailed analysis of the work and decision processes, we developed the design requirements. Continuing with an iterative design process, including steady feedback from the users, we ended with a design prototype. We have evaluated the chosen plot types for presenting future stress levels and collected feedback about the prototype.

In the first part we asked the air traffic control supervisors to have a glance at single graphs and to assess the perceived situation according to severity, their wish to further investigate the situation, or to instantaneously react on it. The results revealed few difference between interpretation of the chosen color value and interpretation of standard plots for time series presentation, like line graphs or bar charts. Although the ratings on the overall design did not show clear positive nor negative tendencies, comments indicated that the design is well understood by the supervisors and that the center view is perceived as clearly structured.

In order to conclude on the use of the cumulated color values for time series in this context there is further investigation needed. The perception and interpretation of color values for future stress situations of multiple air traffic controllers, especially a whole center, needs further investigation. For this case, a study with recorded real-world data may reveal further issues.

Altogether this piece of work taught us once again the importance of a detailed examination of the application area and a tight feedback loop to the users. The use of a single color value for time series data in a context, where plenty of those time series have to be interpreted according to their severity within a glance, seems to be promising. As in the single plot condition there was marginal difference between the visualization types. According to the comments, the color values seemed to contribute to a clear structure in the design. However, further investigation is needed also in this regard. Additionally, the project showed that besides the work of an air traffic controller, there is only little attention payed to the work of air traffic supervisors, making this application area worth for further investigation.

Acknowledgments

This work was accomplished as part of the project "Stay-Centered - Methodenbasis eines Assistenzsystems für Centerlotsen (MACeLot)" at Chemnitz University of Technology. It was supported and partially funded by the German Federal Ministry of Education and Research.

We thank the German air navigation service provider DFS Deutsche Flugsicherung GmbH and the participating supervisors for their support.

References

[1] L. Pfeiffer, T. Sims, and P. Rosenthal, "Visualizing workload and emotion data in air traffic control - an approach informed by the supervisors decision making process," in Proceedings of the International Conference on Advances in Computer-Human Interactions, 2017, pp. 81–87.

[2] J. Buxbaum, N. H. Müller, P. Ohler, L. Pfeiffer, P. Rosenthal, and G. Valtin, "Emotion-sensitive automation of air traffic control - adapting air traffic control automation to user emotions," International Transportation, vol. 68, no. 1, 2016, pp. 36–39.

[3] L. Pfeiffer, G. Valtin, N. H. Müller, and P. Rosenthal, "The mental organization of air traffic and its implications to an emotion sensitive assistance system," International Journal on Advances in Life Sciences, vol. 8, 2016, pp. 164–174.

[4] M. Sollitto, M. M. Martin, S. Dusic, K. E. Gibbons, and A. Wagenhouser, "Assessing the supervisor-subordinate relationship involving part-time employees," International Journal of Business Communication, vol. 53, no. 1, 2016, pp. 74–96.

[5] V. U. Druskat and S. B. Wolff, "Building the emotional intelligence of groups," Harvard Business Review, vol. 79, no. 3, 2001, pp. 80–91.

[6] S. G. Barsade and A. P. Knight, "Group affect," Annual Review of Organizational Psychology and Organizational Behavior, vol. 2, no. 1, 2015, pp. 21–46.

[7] D. Broach and C. S. Dollar, "Relationship of employee attitudes and supervisor-controller ratio to en route operational error rates," US Department of Transportation, Federal Aviation Administration, Office of Aerospace Medicine, Tech. Rep., 2002.

[8] A. Dix, J. E. Finlay, G. D. Abowd, and R. Beale, Human-Computer Interaction (3rd Edition). Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2004.

[9] O. Kulyk, R. Kosara, J. Urquiza, and I. Wassink, "Human-centered aspects," in Human-centered Visualization Environments: GI-Dagstuhl Research Seminar, Revised Lectures, A. Kerren, A. Ebert, and J. Meyer, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 13–75.

[10] M. Sedlmair, M. Meyer, and T. Munzner, "Design study methodology: Reflections from the trenches and the stacks," IEEE Transactions on Visualization and Computer Graphics, vol. 18, no. 12, 2012, pp. 2431–2440.

[11] J. Liu, W. Wang, and K. Cao, "Leader political skill and team performance: A moderated mediation model," Nankai Business Review International, vol. 2, no. 1, 2011, pp. 5–22.

[12] R. L. Brouer, C.-Y. Chiu, and L. Wang, "Political skill dimensions and transformational leadership in china," Journal of Managerial Psychology, vol. 31, no. 6, 2016, pp. 1040–1056.

[13] J. Funke, "Complex problem solving: A case for complex cognition?" Cognitive Processing, vol. 11, no. 2, 2010, pp. 133–142.

[14] J. Quesada, W. Kintsch, and E. Gomez, "Complex problem-solving: A field in search of a definition?" Theoretical Issues in Ergonomics Science, vol. 6, no. 1, 2005, pp. 5–33.

[15] J. Nielsen, Coordinating User Interfaces for Consistency. Morgan Kaufmann, 1989.

[16] C. Ware, Information Visualization: Perception for Design. Morgan Kaufmann, 2013.

[17] W. Lidwell, K. Holden, and J. Butler, Universal Principles of Design, Revised and Updated: 125 Ways to Enhance Usability, Influence Perception, Increase Appeal, Make Better Design Decisions, and Teach Through Design. Rockport Publishers, 2010.

[18] P. Rosenthal, L. Pfeiffer, N. H. Müller, and G. Valtin, The Long Way to Intuitive Visual Analysis of Air Traffic Control Data, 1st ed. Routledge, 2016, ch. 11, pp. 138–148.

[19] K. H. Teigen, "Yerkes-dodson: A law for all seasons," Theory & Psychology, vol. 4, no. 4, 1994, pp. 525–547.

[20] A. Sears and J. A. Jacko, Eds., The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications. Lawrence Erlbaum Assoc Inc, 2007.

[21] E. Tufte, Envisioning Information. Graphics Press, 1990.

[22] W. Aigner, S. Miksch, H. Schumann, and C. Tominski, Visualization of Time-oriented Data. London: Springer, 2011.

[23] H. Ha, G.-n. Kim, W. Hwang, H. Choi, and K. Lee, "CosMovis: Analyzing semantic network of sentiment words in movie reviews," in Proceedings of the IEEE Symposium on Large Data Analysis and Visualization (LDAV). IEEE, 2014, pp. 113–114.

[24] C. A. Steed, M. Drouhard, J. Beaver, J. Pyle, and P. L. Bogen, "Matisse: A visual analytics system for exploring emotion trends in social media text streams," in Proceedings of the IEEE International Conference on Big Data (Big Data). IEEE, 2015, pp. 807–814.

[25] F. Y. Wang, A. Sallaberry, K. Klein, M. Takatsuka, and M. Roche, "SentiCompass: Interactive visualization for exploring and comparing the sentiments of time-varying twitter data," in Proceedings of the IEEE Pacific Visualization Symposium (PacificVis). IEEE, 2015, pp. 129–133.

[26] D. Cernea, C. Weber, A. Ebert, and A. Kerren, "Emotion-prints: Interaction-driven emotion visualization on multi-touch interfaces," in Proceedings of SPIE, Visualization and Data Analysis, D. L. Kao, M. C. Hao, M. A. Livingston, and T. Wischgoll, Eds., 2015, pp. 93 970A–93 970A–15.

[27] J. Ohene-Djan, A. Sammon, and R. Shipsey, "Colour spectrum's of opinion: An information visualisation interface for representing degrees of emotion in real time," in Proceedings of the International Conference on Information Visualisation. IEEE, 2006, pp. 80–88.

[28] A. Sonderegger, K. Heyden, A. Chavaillaz, and J. Sauer, "AniSAM & AniAvatar: Animated visualizations of affective states," in Proceedings of the CHI Conference on Human Factors in Computing Systems. New York, NY, USA: ACM, 2016, pp. 4828–4837.

[29] M. D. Byrne, "Using icons to find documents: Simplicity is critical," in Proceedings of the INTERACT and CHI Conference on Human Factors in Computing Systems. ACM, 1993, pp. 446–453.

[30] L. Nowell, R. Schulman, and D. Hix, "Graphical encoding for information visualization: An empirical study," in Proceedings of the IEEE Symposium on Information Visualization (InfoVis). IEEE, 2002, pp. 43–50.

[31] D. McDuff, A. Karlson, A. Kapoor, A. Roseway, and M. Czerwinski, "AffectAura: An intelligent system for emotional memory," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2012, pp. 849–858.

[32] R. Rosenholtz, "Capabilities and limitations of peripheral vision," Annual Review of Vision Science, vol. 2, no. 1, 2016, pp. 437–457.

[33] M. Harrower and C. A. Brewer, "Colorbrewer.org: An online tool for selecting colour schemes for maps," The Cartographic Journal, vol. 40, no. 1, 2003, pp. 27–37.

[34] A. T. Woods, C. Velasco, C. A. Levitan, X. Wan, and C. Spence, "Conducting perception research over the internet: A tutorial review," PeerJ, vol. 3, 2015, p. e1058.

[35] S. Thorpe, D. Fize, and C. Marlot, "Speed of processing in the human visual system," Nature, vol. 381, 07 1996, pp. 520–522.