Int. J. Human-Computer Studies 95 (2016) 54–65

Contents lists available at ScienceDirect

Int. J. Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Continuous sonification enhances adequacy of interactions in peripheral process monitoring

Tobias Hildebrandt [a,*], Thomas Hermann [b], Stefanie Rinderle-Ma [a]

[a] Workflow Systems and Technology, Faculty of Computer Science, University of Vienna, Waehringerstr. 29, 1090 Vienna, Austria
[b] Ambient Intelligence Group, Cluster of Excellence - Cognitive Interaction Technology (CITEC), Bielefeld University, Germany

ARTICLE INFO

Article history:
Received 24 August 2015
Received in revised form
20 June 2016
Accepted 23 June 2016
Available online 1 July 2016

Keywords:
Sonification
Multi-modal display
Process monitoring
Quantitative study
Dual task
Industrial monitoring

ABSTRACT

As many users who are charged with process monitoring need to focus mainly on other work while performing monitoring as a secondary task, monitoring systems that purely rely on visual means are often not well suited for this purpose. Sonification, the presentation of data as (non-speech) sound, has proven in several studies that it can help in guiding the user's attention, especially in scenarios where process monitoring is performed in parallel with a different, main task. However, there are several aspects that have not been investigated in this area so far, for example if a continuous soundscape can guide the user's attention better than one that is based on auditory cues. We have developed a system that allows reproducible research to answer such questions. In this system, the participants' performance both for the main task (simulated by simple arithmetic problems) and for the secondary task (a simulation of a production process) can be measured in a more fine-grained manner than has been the case for existing research in this field. In a within-subject study ($n=18$), we compared three monitoring conditions – visual only, visual + auditory alerts and a condition combining the visual mode with continuous sonification of process events based on a forest soundscape. Participants showed significantly higher process monitoring performances in the continuous sonification condition, compared to the other two modes. The performance in the main task was at the same time not significantly affected by the continuous sonification.

1. Introduction

Business processes such as in manufacturing and logistics or in administration, but also technical processes like in robotics, are becoming increasingly complex while they are at the same time more and more automated, computerized and monitored in real-time (Malone et al., 2003).

This is true for processes in many domains, but especially for industrial productions, where a delayed delivery of raw materials can lead to a standstill in production and thus high loss of profit. On the one hand the increasing amount of data offer an enormous potential to better monitor and control processes. On the other hand it puts increasing pressure on monitoring personnel who need to observe processes.

The status quo in large-scale process monitoring is heavily focused on control centers where users observe production on multiple screens, using both video features as well as schematic overviews of process and machines/facilities, charts/graphs, textual descriptions and alerts (Sauer, 2004).

Especially in smaller- and medium-sized production companies, there are often no dedicated personnel charged with full-time monitoring, but instead engineers and supervisors need to primarily perform other tasks, yet monitor the process' status at the same time. However, especially in such peripheral or *serendipitous-peripheral monitoring* scenarios where the attention is focused on a primary task and other information is monitored indirectly at the same time, visual means are not well suited, as pointed out by Vickers (2011).

Meanwhile, maintenance experts have been using the auditory sense to identify or anticipate possible machine problems, a technique referred to as vibration analysis, for a long time. Crucial vibration properties are amplitude, frequency, phase and modulation (Renwick and Babson, 1985). Therefore, traditional production monitoring is still considered to be a holistic approach, covering the visual, auditory and even olfactory sense, even though automation has enhanced *manual* vibration analysis in the recent years (Hildebrandt et al., 2014b).

In modern production settings, sound is typically only used as a means to convey warnings and alerts, e.g., to convey an alarm situation when a machine broke down or a predefined threshold had been exceeded (Siemens, 2007). In a production scenario, this

* Corresponding author.
E-mail address: Tobias.Hildebrandt@univie.ac.at (T. Hildebrandt).

http://dx.doi.org/10.1016/j.ijhcs.2016.06.002
1071-5819/

could for instance be the case when the stock level of a resource has dropped below a critical level, or when a temperature sensor of a machine measures a critical temperature, indicating imminent machine failure (SAP SE, 2015).

However, this type of auditory display has several drawbacks: on the one hand, if rules that define alert triggering thresholds are defined too *conservative*, i.e., requiring strong evidence before issuing positive classifications, potentially critical situations such as machine failures might occur without issuing an alert. On the other hand, if the values are defined too *liberal*, i.e. risking high false positive rates, the resulting flood of (in many cases unnecessary) alerts and alarms might lead to an information overload of the user, or to the situation that the user stops to take the alerts as serious as they are. Furthermore, in many scenarios engineers are not able to define all states and values that might lead to a critical situation beforehand. Levels and values that might constitute a critical state are often complex to decide, as e.g., the question if a specific parameter value constitutes a critical situation or not often depends on the context, given by various other parameters. But even if all possibly critical situations are covered by alerts and alarms, in most cases operators might prefer to be informed even *before* a situation might become critical, thus enabling them to *anticipate*, *intervene* and *avoid* the problem. A constant awareness of states and values through an *auditory ambient information system* might enable such an anticipation of critical situations. Thus, we suggest to use the mentioned tradition of auditory monitoring as a leverage effect by supplementing state-of-the-art visual process monitoring with techniques from sonification.

Sonification is the systematic, reproducible and thus scientific method for representing data as (mostly non-speech) sound in an auditory display (Hermann, 2008). Well-known examples of sonifications are the Geiger counter for displaying radioactive radiation, or the auditory parking aid which conveys the distance to the vehicle or obstacle behind as pulse rate of a beep sound. Beyond these very basic and simple types, sonification researchers have developed a plethora of approaches to represent more complex data such as multivariate time series (e.g. EEG and ECG), or spatio-temporal data (e.g. images and well logs), and also general high-dimensional data distributions.

Sonification has several key advantages that makes it suitable especially for the application area of real-time process monitoring, like our ability to process audio faster than visuals or the fact that we easily habituate to static sound sources, yet that we are at the same time very sensitive to changes (Vickers, 2011).

For these reasons sonification promises a solution to the aforementioned challenges of state-of-the art process monitoring. However, there are several open questions when it comes to supporting users in monitoring as a secondary task that concern the sonification design and as well as how different types of sound-enhanced process monitoring affect attention and concentration in main- and secondary task, which we tackled with this paper. Our main research goals were (a) to find out if a continuous, soundscape-based sonification of individual production steps can support users better in monitoring as a secondary task than a purely-visual solution, or one that is based on auditory alerts. Other open research questions were (b) to what extent the three different conditions distract users from their main task, (c) how users rate the three different conditions concerning relevant aspects such as pleasingness, helpfulnesses, intrusiveness or exhaustiveness. Answering those research questions poses several challenges, such as simulating the potential users' main- and secondary task in such a way, that they are both cognitively demanding and thus binding the undivided attention, while at the same time allowing for an easy and reliable measurement of task performance in a fine-grained manner. As there are no standardized environments that fit these requirements, we have developed the SoProMon system (Sonification for Process monitoring), that is a hard-/software system for reproducible research in sonification for peripheral monitoring, particularly for the investigation of attention allocation in dual-task-settings. The system has already been presented in Hildebrandt et al. (2014a), and consists primarily of a main task console to bind the user's attention by presenting simple arithmetic problems and a simulated production process that requires different types of user interactions (see Section 4). Based on the SoProMon system, we conducted an extensive experiment in a within-subject design (n = 18), whose results contribute to answering the mentioned open research questions, and thus to advancing research in this area (see Section 2) in the following ways:

- To our best knowledge no quantitative experiments using soundscapes in dual-task settings have been conducted so far. The experiment that we conducted featured a sonification based on a forest soundscape design (see Hildebrandt et al., 2014a) to enable long-term listening without fatigue.
- Sonification designs in previously conducted experiments for peripheral process monitoring either base on auditory cues, or on continuous sonifications. The experiment described in this paper compares three conditions: visual only, visual + auditory cues and visual + continuous soundscape sonification.
- Furthermore, dual-task experiments that have been conducted in this area measure the performance in both tasks typically using either binary correctness measures and/or response times. In our experiment, we employ a more fine-grained performance measurement in which for each user interaction a continuous score is assigned that either measures correctness (main task) or adequacy (secondary, process monitoring task).
- In most quantitative experiments, the user's opinion on e.g. the different conditions and his/her understanding of those is not gathered, or if it is, not in a very fine-grained way. For the experiment we developed an extensive questionnaire that features a pre experiment-, three postcondition- and one post-experiment part.

As industrial production is an area, in which it is especially crucial to monitor processes in real-time and that can probably be intuitively understood also by domain novices, the secondary task of this experiment is based on a simulated production processes. However, as the experiment design aimed at quite fundamental questions of attention allocation, the results should be generic enough to be transferable to monitoring scenarios in other domains as well.[1] The details on the current state-of-the-art concerning research in sonification for (peripheral) process monitoring as well as on the open research issues that we tried to tackle with the experiment can be found in Section 2. The hypotheses derived from the literature which we tried to tackle with the experiment are described in Section 3, followed by an introduction into the SoProMon system (Section 4) and the methodology of our experiment (Section 5). Experimental results will be presented in Section 6 and discussed in Section 7, followed by overall conclusive considerations.

2. Related work

There is a substantial amount of research concerning applications of auditory process monitoring, spanning various areas such as industrial production processes, program execution or web

[1] Preliminary results from selected questionnaire items have been presented in an extended abstract (Hermann et al., 2015).

server behavior. A good summary is given in Vickers (2011). In the following we focus on the most relevant works here with respect to the SoProMon system.

Gaver et al. (1991) explore with their *ARKOLA Simulation* the production processes of a bottling plant in a multi-modal representation that combines visual and auditory means (Gaver et al., 1991). In Rauterberg and Styger (1994), sonification has been applied for the direct monitoring of an assembly line. The authors concluded that participants of a study who had visual as well as event-based auditory feedback felt more self-assured and socially accepted than in the visual condition. A popular application area of auditory process monitoring is computer program debugging, as e.g. investigated in Alty (1995). This so-called program auralization is relevant as it assumes a similar monitoring mode as in process monitoring; in debugging, however, the monitoring becomes typically the main task. An area where monitoring is more in the periphery is auditory monitoring of web servers and computer networks, such as the concepts implemented in the systems *Peep* (Gilfix and Couch, 2000) or *WebMelody* (Ballora et al., 2010). Sonification is also used frequently for computer security and -intrusion detection (e.g. see Gopinath, 2004 and Ballora et al., 2011). As in both fields the focus of attention is often elsewhere, they offers relevant ideas and approaches for our task at hand.

First steps towards sonification in business process monitoring have also been taken in our preliminary work (Hildebrandt et al., 2014b). As process monitoring is in many cases a peripheral task, a very important aspect in designing process monitoring systems for peripheral monitoring is attention allocation. Auditory monitoring systems should ideally, during normal operation, hardly be perceived actively at all. In cases that require the user's attention, such as exceptional or even potentially dangerous situations, the sonification should nonetheless be able to attract the user's full attention. This leads to a trade-off between *awareness* and *disturbance*. Generally, the more information a sonification conveys, the greater the risk of disturbance. This trade-off has been researched by Gaver et al. (1991), among others. In Vickers et al. (2014), the authors suggest to use a soundscape that is designed to achieve unobtrusiveness by relying on nature recordings. There is a wide selection of research that investigates how sonifications can guide the user's attention, such as by Seagull et al. (2001) or (Anderson and Sanderson, 2004), and on how to design sonifications for peripheral monitoring (Watson and Sanderson, 2007).

In summary, there is a substantial body of research on sonification for process monitoring in general. There exists furthermore research dealing specifically with peripheral process monitoring, often in dual-task scenarios, although there are some research gap in this area. Only for a few of those approaches, studies have been conducted to test their effectiveness. Those studies that compare a visual-only condition to a multi-modal condition that conveys sporadic auditory alerts or alarms conclude, that the performances in both tasks seem to be in most cases not significantly affected by the auditory signals (McClimens et al., 2011; Brock et al., 2010). In a few cases, both tasks are negatively affected (e.g. McClimens et al., 2004). Typically there are less head movements and attention switches measured in the multi-modal condition (Brock et al., 2002), something that has also been observed for continuous sonifications (Sanderson et al., 2004).

When comparing the performances in the main task – which in many studies is simulated by presenting arithmetic problems – in experiments that include continuous sonifications, the results are mixed: in some experiments less mistakes were made in the multi-modal condition compared to the visual-only condition (Watson et al., 2003; Poguntke and Ellis, 2008), while in other experiments the best main task performance was observed in the visual condition (Crawford et al., 2002; Watson and Sanderson, 2004). In tendency, the main task performance seems to be slightly negatively affected by sound (more so in multi-modal conditions than in auditory-only conditions), although these differences between the conditions are in the majority of studies not statistically significant. Even so, especially the results in the conditions that include sound are typically better when the respective condition is not the subject's first, but second or third condition of the experiment (e.g. Watson et al., 2003 or Poguntke and Ellis, 2008). Thus, the observed distraction by sound seems to be smaller when the participants are already used to and familiar with the two tasks themselves.

The performance in the secondary task (monitoring) is typically significantly higher in multi-modal conditions that feature continuous sonifications, compared to visual-only conditions (Watson et al., 2003; Crawford et al., 2002), although a few studies report the opposite result (e.g. Sanderson et al., 2004). Like for the main task, there seems to be a strong familiarization effect that especially benefits the multi-modal conditions (Watson et al., 2003), which may be an explanation for why the advantage of auditory conditions over the visual-only condition seems to be greater for domain experts than for amateurs (Crawford et al., 2002). In general, participants and especially domain experts, when asked for their opinion, state that they preferred the multi-modal conditions including continuous sonifications, as – among other reasons – they made them feel more in control (Poguntke and Ellis, 2008; Crawford et al., 2002).

3. Hypotheses

Based on the related research in Section 2, we define several hypotheses. As a baseline, we take the two most common modes of monitoring in current production scenarios (and of many other domains as well):

- $C_{\mathbf{vis}}$, in which the process status is conveyed using only visual means.
- $C_{\mathbf{sota}}$, that combines C_{vis} with auditory alarms.

These two are compared to

- $C_{\mathbf{son}}$, combining the other two conditions with a continuous sonification.

In general, we expect differences between C_{vis} and C_{sota}, but especially between C_{son} and the other two, basically in favor of the auditory information types. We expect the differences to manifest both in the users' perceptions (measured by their questionnaire responses and their comments) and the quantitative performance measured in the monitoring task.

We expect, in accordance with the literature, that the three mentioned conditions will have no significant effect on the performance of the main task (H1.1). Concerning process monitoring performance and behavior, we expect that additional auditory cues have no effect (H2.1). We do however expect continuous sonification to have a significantly positive effect on monitoring performance, compared to the other two conditions (H2.2), as this is what the sonification design mainly aims at. Concerning the questionnaire responses, we expect items associated with helpfulness, attention switching and performance increase of the respective mode of operation to be more favorably rated in C_{son} compared to C_{vis} and C_{sota} (H3.1).

We furthermore believe that users feel more self-assured and in control with C_{son} (H3.2). By contrast, we expect additional information conveyed aurally to increase exhaustion, and therefore C_{sota} and C_{son} to be more exhaustive than C_{vis} (H3.3). For each of these hypotheses, the null hypothesis (H_0) says that there are no significant differences between the groups, while the alternative hypotheses (H_1) assumes the opposite.

4. The SoProMon system

The Sonification for Process Monitoring (SoProMon) system[2] is a hard-/software system for reproducible research in sonification for peripheral monitoring, particularly for the investigation of attention allocation in dual-task-settings. The core software components are (a) a process simulation, (b) a visual monitoring system including graphical user interface elements (buttons) to intervene, (c) a sonification system that allows to plug-in different sonification types for multi-modal variants and (d) a main task console to bind the user's attention to a (different) focus. This is complemented by a set of service modules for logging all relevant data, including a video camera mounted atop the user to store head orientations, and calibration modules for user-adjustment of the sound levels.

The system is highly modular and flexible, and individual modules can be replaced by other customized code if required. For our first practical implementation we decided to create a setting where a user is seated in front of two monitors, oriented perpendicular to each other, one for the main task and one for the monitoring console, in order to be able to stimulate and observe attention shifts. Furthermore, the keyboard and the mouse were fixed to the table, so that they could not be moved (see Fig. 1). Yet other implementations, such as letting users move freely in the room and solve practical problems, (as opposed to computer tasks) are certainly also conceivable.

Concerning (a), the process simulation, we chose a rather life-like scenario in the realm of manufacturing: we represent the process as a graph of 6 production steps that partially run in parallel and require input of one or more previous production steps at times. The number 6 is arbitrary, yet chosen here to have enough complexity to be not trivial and low enough to first learn about processes of limited complexity. Even though, as already motivated before, the current implementation of SoProMon is based on the simulation of a production process, the system can be adapted to simulate different types of processes as well. We designed the simulation so that it requires several user interactions, in order to measure the performance of auditory monitoring in attention allocation and in interrupting the users during their main task. The required interactions are:

Supply: One machine requires the user to refill the resource input. To simulate a realistic environment, the machine contains a random factor that influences the time between an input resource has been taken and the resulting material has been produced.

Empty: One machine requires the user to clear the output buffer (i.e. initiate a delivery/transport of goods) to make space for new assembled goods to be buffered.

Maintain: Two machines can encounter conditions of malfunction/maintenance stops, which require active attention of the operator. The time distances between such situations are quite constant, but contain a random element as well. A click on the 'maintain' button resolves the problem.

After a click, the machines idle a given short time, to discourage users from performing unnecessary actions. This is in so far realistic, as real machines often require a short downtime when they are being refilled, and a longer one if they are being maintained or repaired.

[2] Website: http://cs.univie.ac.at/wst/research/projects/project/infproj/1063/.

As to (b) the visual monitoring, we depict a graph of the machine setup and flow of goods. While Fig. 2 shows the *normal* state of simulation, in which all machines are working, Fig. 3 shows a critical state, in which several machines are out of order. The visualization can be assumed to be checked very quickly, leaving any time for the interpretation and reaction to be accounted for the interpretation of sounds. As our main concern is the assessment of sonifications that complement an existing visual monitoring console, the visual part remains invariant in all experimental conditions and thus does not require a dedicated motivation or testing.

Concerning (d), the main task, we chose a task that is both cognitively demanding and thus binds the undivided attention, while at the same time allows the easy and reliable measurement of task performance. For this purpose, it is best if the task consists of a series of repeated smaller elementary tasks whose correctness can be computed. Ideally the main task can be interrupted to attend to the monitoring. Typical tasks in real-world scenarios are, depending on the user group, processing documents such as emails, planning/scheduling, or repairing machines. For the sake of easier evaluation we selected the *adding numbers* task, which is a mental arithmetic task of summing up two numbers (each smaller than 50). The result is to be entered into a text field using the computer keyboard. On hitting the return key the task, the result, and the timestamp are logged, and the next pair of random numbers is drawn and presented. The window is displayed in low font size on the screen perpendicular to the monitoring screen.

5. Methods

As already mentioned, the main goal of the experiment was to answer the previously discussed research challenges. Three conditions, as explained in Section 3, are being compared to each other: C_{vis}, C_{sota} (combining C_{vis} with auditory alarms which were conveyed when a machine stopped), and C_{son} (combining the other two conditions with a continuous soundscape sonification introduced in detail in Hildebrandt et al., 2014a).

C_{son} is true to the event-like nature of individual machine executions, meaning that every process step in the production yields a tiny sonic counterpart so that their superimposition creates a soundscape that reflects the overall activity. Assigning different sounds/timbres to different machines results in 6 voices that play simultaneously. The reason to not only sonify those machines that require interactions is, that the sonification of irregularities in the production for machines that precede machines for which interactions are required can help to anticipate problems that might occur for those machines at a later stage.

We chose sounds of the forest theme, namely small bird, woodpecker, water drop, bee, river splash and cracking twig. The sonification is designed to form a soundscape by selecting sounds that constitute the perception of a coherent setting (e.g. forest). While soundscape ecology would suggest an optimization process so that the bandwidth allocation reduces the risk of masking, we here just select sounds based on subjective choices to fit into the theme. When a machine has reached a critical level, its sound is repeated at a fast rate and high volume, until the problem is solved. However, beyond a mere display of individual executions we add information by using a mapping of machine-specific data to the acoustic shape of the sound events:

- We map the output buffer filling level to pitch, thus making use of the analogy of a filling jug.
- We map the input buffer criticality to increasingly louder noise to the initial/transient phase of the sound, thus enabling listeners to anticipate that input buffers run empty. The increase

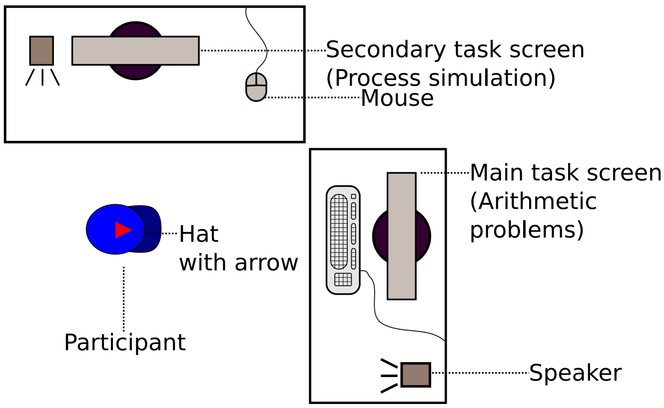

Fig. 1. Schematic overview of the SoProMon system setup (revised from Hildebrandt et al., 2014a).

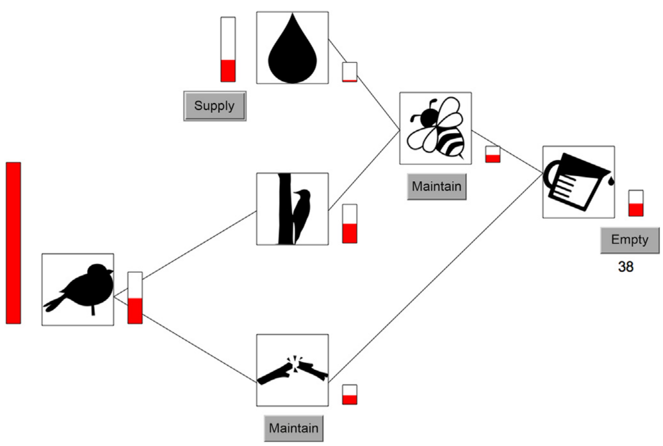

Fig. 2. SoProMon visualizations: filling levels are depicted in red, two machines include maintain buttons, the other buttons are for the management of buffers (supply/empty). (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

of noisiness becomes more and more discernible as buffers slowly run empty.

- We generally map the approaching of critical conditions to level, resulting in machine sounds gradually becoming louder and thus more salient over the 'normal' background soundscape, as the situation gets worse. Specifically, the sound level increases by 21 dB, starting when the buffer level is at 25% (machine with 'supply' button)/75% (machine with 'empty' button), or when the condition of a machine has reached 25% (machines that need maintenance).

For more details please refer to Hildebrandt et al. (2014a). Video examples for the SoProMon sonification (for both C_{sota} and C_{son}) used within the study are available on our website[3] and as supplementary material at the online version (http://dx.doi.org/10.1016/j.ijhcs.2016.06.002) of this paper. Spectrograms of the different machines and states are available as well.

In order to answer our research questions, different types of data were recorded during the experiment:

- Main task logs, containing the arithmetic problems and the entered solutions with timestamps.
- Process simulation logs, containing the timestamp, the performed type of interaction, and the current buffer levels and machine status'.
- Questionnaires and verbal comments.
- Audio and video recordings.

5.1. Study population

The study population had an age median of 26 ± 7.1 years. Table 1 shows the experiences of the participants in various relevant fields, as well as their visual and hearing impairments.

Table 2 shows the participants' opinions and estimations concerning the foundations of the experiment, such as the graphical user interface. Two participants criticized that they had to press a key in order to switch the focus of attention to the main task window. In recognition of this possible error source we corrected the calculation mistakes that were (supposedly) due to the first digit after an action being ignored in post-processing. Furthermore it was mentioned, that the mouse positioning favored left-handed users.

5.2. Experiment plan

Two pre-tests have been performed before the experiment to ensure the understandability of the system, the questionnaires, the two tasks and the provided instructions, the validity of the experiment, and the technical functionality of all components of the SoProMon system, especially the logging mechanisms. Several adjustments have been made after the pre-tests. In order to have three complete sets of condition sequences in this within-subjects design, 18 participants were recruited for the experiment itself. All 6 possible permutations of the three conditions were realized among the runs. Before signing an informed consent, the participants were give a written introduction into the goals and aims of the study, its duration and experiment procedure, as well as written instructions for the system. The experiment goal, namely to solve as many arithmetic problems as possible during the three experiment parts of 10 min each, while at the same time trying to avoid critical process states as well as possible was, shortly before the experiment, repeated verbally as well.

Before starting the process simulation with one of the three conditions (C_{vis}, C_{sota}, C_{son}), the participants were informed regarding how the process status and criticality was conveyed (e.g. no sound, sound in the case or errors, and permanent sounds). If the respective part of the experiment was the participant's first one, he or she had time to familiarize with the process simulation beforehand, and ask questions. If the participants' first condition was C_{son}, there was additionally a sound level calibration phase, during which the individual machine's volumes were adjusted until they were just loud enough to identify. Questionnaires were handed out before the experiment, after the 3 conditions, as well as after the experiment. In total, the experiment lasted around 65 ± 5 min for each participant.

5.3. Main task measurement

Different measures can be used to compare the performances. The number of calculations that the participants were able to execute in the given experiment time of 10 min, would be the most obvious choice. On the other hand, not only the number of calculations is relevant, but also their correctness. Therefore, the conditions are also compared concerning their mean *deviation*, which is calculated by averaging the deviation of each entered result from the correct result by

[3] https://pub.uni-bielefeld.de/publication/2904376.

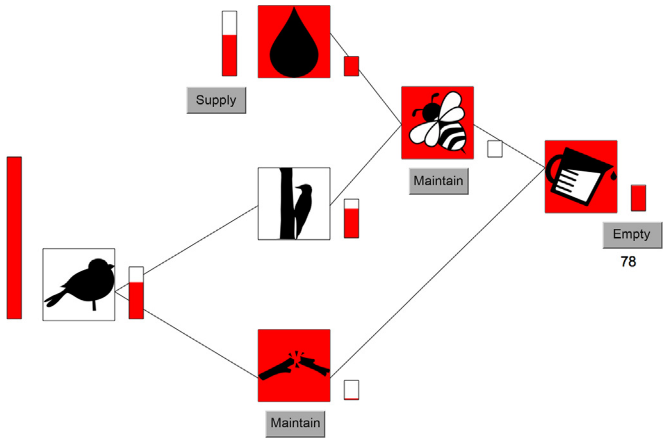

Fig. 3. Four out of the six machines have reached a critical state, respectively depicted by a red filling. The user can resolve these states by performing the respective actions. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Table 1
Experiences and impairments. Exp = Experience with.

Item	Not at all	Rather not	Rather	Fully applies
Hearing impairments	17	1	0	0
Visual impairments	13	2	3	0
Exp. process simulation	10	7	0	1
Exp. musical instrument	8	4	3	3
Exp. audio	6	7	4	1
Exp. sonification	8	4	5	1

Table 2
Opinions on foundations of experiment.

Item	*Median* $\pm$ *IQR*
Understood foundations of process simulation	10.0 ± 0.6
Understood interaction possibilities	10 ± 1.0
Felt overwhelmed by process simulation	1.0 ± 2.1
Felt overwhelmed by main task	1.0 ± 1.7
Visual representation was clear and understandable	8.0 ± 2.4
Understood GUI & process visualization	10.0 ± 1.4

$$\text{Deviation}^{\alpha} = \frac{1}{N^{\alpha}} \sum_{i=1}^{N^{\alpha}} \left| 1 - \frac{r_i^{\alpha}}{\hat{r}_i^{\alpha}} \right|,$$

where α is the participant number, N^{α} is the number of questions that participant α has replied, and r_i (resp. $\hat{r}_i$) are the given (resp. correct) results for the ith question. Furthermore, in order to compare the overall *performance* of the participants between the three conditions, we introduce the 'main task score', a variable that encompasses both the number of solved calculations and their correctness as

$$\text{Main task score}^{\alpha} = \frac{N^{\alpha} - \langle N \rangle}{\sigma(N)} - \frac{dev^{\alpha} - \langle dev \rangle}{\sigma(dev)},$$

with $\langle \cdot \rangle$ referring to the sample mean and $\sigma(\cdot)$ being the standard deviation of its variable.

5.4. *Process simulation measurement*

An obvious metric to measure the monitoring performance between the different conditions would be to compare how many clicks the users made in average for each condition. Furthermore of interest are the buffer values of the respective buffers at the time of the user's interaction with the simulation (e.g., the input buffer of a certain machine at the time of refilling it). A relatively high average buffer value can e.g. signify that the users do not trust that the respective mode of process monitoring conveys the need for interaction in time, leading the users to switching their attention to the process simulation in regular intervals, and performing interactions *just in case*. A low average buffer can, on the other hand, signify that the users rely on the respective conditions' ability to signal interaction needs. On the other hand, if e.g. an input buffer had already been completely depleted at the time of intervention, this may signify that the respective condition has failed to inform the users in time. In many cases, participants used double clicks for their interactions, while a single click would have been sufficient, a fact that was perhaps not communicated clearly enough to the participants. Therefore, if several clicks were performed directly one after another, only the first click was taken into account.

5.4.1. *Anticipation optimal rationality*

The concept of *anticipation optimal rationality* tries to be true to a real-life production scenario: interactions that come *too early*, and were therefore *unnecessary*, are here punished. This is realistic, as in most real-life production scenarios it is a goal to maximize production and to minimize downtimes as much as possible. As each interaction in the simulation entails a downtime (e.g. for maintaining a machine), it is logical to minimize such interactions. However, the anticipation optimality not only punishes interactions that are too early but also such that are *too late*. *Too late* in this case means, that a critical situation had, at the time of interaction, already occurred, or that it was so close, that it could not have been prevented assuming average reaction times. The rationale behind this is that downtimes due to critical states (such as a machine that has broken down) are to be avoided even more than 'planned downtimes' due to maintenance.

Thus, in anticipation optimal rationality R_A, we calculate the mean time that participants needed to shift the attention from the main task to the process simulation. This time span included finishing the calculation that they were working on, turning around to the process simulation, and performing the necessary action. If an interaction was performed when the respective machine had already stopped due to a critical state, or when e.g. the buffer value was already so low, that an interaction was likely to be performed too late, this interaction was evaluated with a score of 0. Interactions that have been performed exactly at a time at which it would have been, given average reaction times, possible to intervene just before a critical state would have been reached, were rated optimal (1.0). All interactions that occurred at a later point were mapped linearly between 1.0 and 0 from the optimal interaction point (1.0) to the point at which an interaction has been completely *unnecessary* (0), e.g. when an input buffer was still completely full (see Fig. 4).

5.5. *Preprocessing of questionnaire data*

In order to test the subjects' accuracy in answering and in order to be able to detect random answering, several items of the questionnaire dealt with the same subject, often in an opposing scale. If the given answers were too contradictory, the respective item pair for the specific participant has been marked as outlier/inconsistent data point and removed for the analysis. This way, out of the 1908 individual answers,[4] 13 of such pairs have been removed (e.g. "I was always in full control of the process simulation"

[4] Corresponds to 106 questionnaire items for 18 subjects.

vs. “I was overstrained by the process simulation”). As several analyses required complete sets of answers (such as the combination of different items), the removed answers have for these purposes been filled with the mean of the respective item.

In order to allow a more powerful analysis and more representative results, corresponding items have been combined into Likert scales (composite scores). The factors have been grouped by average. Subsequently, their consistency has been tested with Cronbach's alpha: only those factors with a reliable consistency outcome (i.e. ≥0.8) have been used for data analysis. If not stated otherwise, the scales range from 10 (fully applies) to 0 (does not apply at all). The complete questionnaire (translated from German) as well as tables with raw data and detailed analyses can be found as supplementary material at the online version of this paper.

5.6. Statistics

All results concerning the main task performance have the following format: mean ± standard deviation, as they are interval-scaled and normally distributed. For the other results (process monitoring performances and questionnaire results), the standard way of communicating results will be median ± IQR (interquartile range), due to the fact that those results have been tested to not be normally distributed, which is why mean values would have little significance. An exception is the discussion of the Likert scales toward the end of the questionnaire results, which have been aggregated out of different Likert items and tested for normality distribution.

For scores for which several data points for one experiment part existed (e.g. the buffer values when refilling a machine), for each of the 54 experiment parts (18 subjects × 3 conditions), a mean value has been calculated and compared between the three modes using Friedman tests for repeated measures, respectively repeated measures ANOVA (for the main task results). The reason aggregating using the mean, instead of the median, even though for some variables only a handful of data points may exist, was to avoid loosing the information of *outliers* and variance. This is due to the fact, that in process monitoring, outliers (e.g. standstills of machines) are crucial and to be avoided whenever possible, wherefore this information is important. Individual comparisons between two conditions were performed using Wilcoxon signed-rank tests for dependent samples, respectively dependent t-tests (main task results).

For all modes the false discovery rate has been adjusted with a Benjamini–Hochberg correction. All p-values have been divided by two, if a directed hypothesis has been defined beforehand. The same is true for presented correlations. The results of correlation tests, if not stated otherwise, combine the r value of Spearman's Rho with the p-value of Kendall's Tau, as it is more accurate for smaller samples.

6. Results

This section describes the most interesting and most significant experiment results. Further results are presented as supplementary material, structured into detailed main task results, non-significant process monitoring results and sound perception and mapping comprehension.

6.1. Main-task results

For the main task performance, no significant differences between the three conditions have been observed, neither in terms of the number of performed calculations, nor their correctness, nor the main task score (the combination of both).

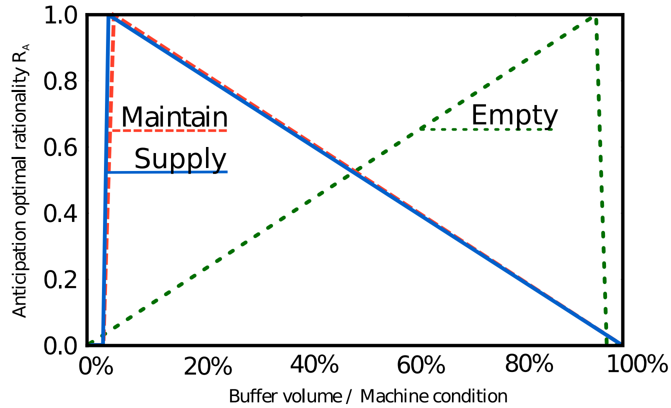

Fig. 4. The figure shows for each of the three interaction types – supply (blue, solid line), empty (green, dotted line) and maintain (red, dashed line) – how an interaction at a specific point in time, measured in terms of the current buffer volume or condition of a specific machine (displayed on the x-axis), would be rated concerning the concept of anticipation optimal rationality (displayed on the y-axis).

6.2. Process monitoring results

In total, 1367 interactions (supply, empty, maintain) were performed, out of these 483 during the C_{vis} conditions (median number of interactions per participant in C_{vis}: 23.0 ± 7.5), 448 during C_{sota} (23.5 ± 2.75) and 436 in C_{son} (23.5 ± 2.75), see Fig. 5.

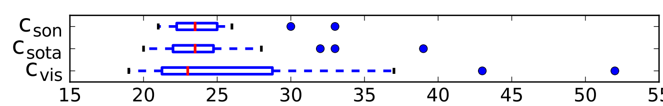

Fig. 5. Median number of mouse clicks per experiment part measured for the monitoring task.

6.2.1. Analysis of buffer values

Although there were no statistically significant differences concerning the mean buffer values of the respective machines when the participants performed the supply-action, the number of supplies that were performed when the respective input buffer was already completely depleted was substantially higher under C_{vis} (29 out of 216 supplies = 13.4%) and C_{sota} (34/199 = 17.1%) than under C_{son} (2/197 = 1%). Significant differences between the conditions were observed for one of the two machines that required maintenance ($p < 0.038$). The aggregated median machine condition when maintaining it was 18.003 ± 34.82%. The median condition at maintenance was 14.204 ± 9.315 under C_{vis}, 17.224 ± 12.966 under C_{sota} and 23.754 ± 12.260 under C_{son}. However, there have been no significant differences between C_{vis} and C_{sota} ($p > 0.127$), C_{vis} and C_{son} ($p > 0.0789$) or C_{sota} and C_{son} ($p > 0.249$) observed. In C_{vis}, 7 out of 83 maintenances were initiated when the machine had already stopped producing, compared to 6 out of 96 in C_{sota}, and 0 in C_{son}.

6.2.2. Anticipation optimal rationality

The aggregated median anticipation optimal rationality score over all conditions was 0.670 ± 0.209. There are highly significant differences ($p < 0.001$) between C_{vis} (0.632 ± 0.175), C_{sota} (0.578 ± 0.111) and C_{son} (0.787 ± 0.074). The results under C_{son} deviate less than under the other conditions (see Fig. 6), something that has been observed for many other metrics in this experiment as well.

The difference between C_{vis} and C_{sota} is not significant ($p > 0.25$), but between C_{vis} and C_{son} ($p < 0.001$) as well as between C_{sota} and C_{son} ($p < 0.001$) significant differences have been observed. As can be seen from Fig. 7, the difference between C_{son} and

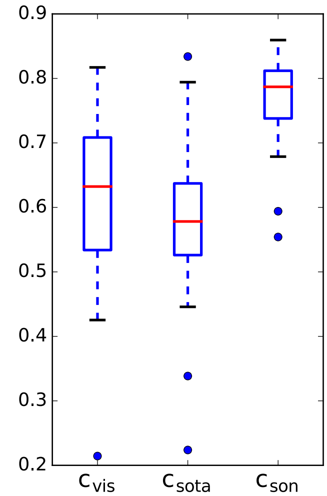

Fig. 6. Median anticipation optimal rationalities of the experiment parts.

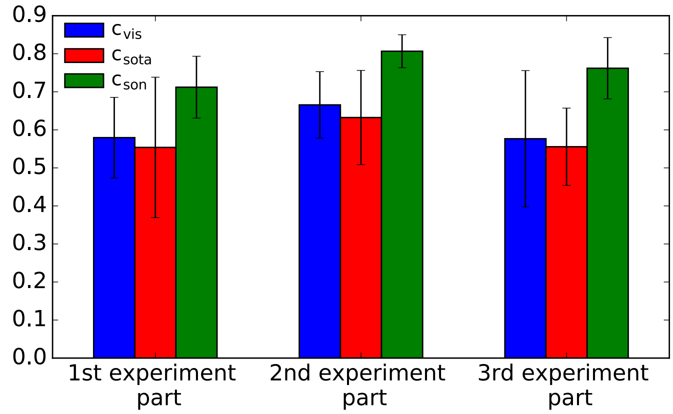

Fig. 7. Aggregated mean anticipation optimal rationality R_A, depending on if the respective condition was the participants first (left), second (middle) or third (right) condition of the experiment. The three sub figures compare the results of C_{vis} (left), C_{sota} (middle) and C_{son} (right).

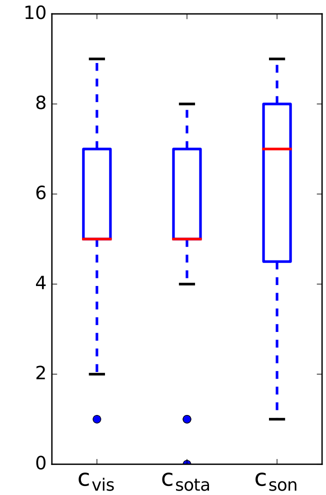

Fig. 8. Monitoring task performance improvement.

C_{vis} and C_{sota} has been higher, when these conditions have not been the participants' respective first condition.

6.3. Questionnaire results

There seems to be a clear preference for C_{son} concerning the responses to those items that compare the three conditions and for which significant differences between the conditions have been observed:

- Item "I have improved my performance in the process simulation over time." ($p < 0.009$, Fig. 8):

  $C_{son}(7.0 \pm 3.5) > C_{vis}(5.0 \pm 2.0) \quad p < 0.048$

  $C_{son}(7.0 \pm 3.5) > C_{sota}(5.0 \pm 2.0) \quad p < 0.048$
- Item "I have improved my performance in the main task over time." ($p < 0.040$)

  $C_{son}(5.5 \pm 4.0) > C_{sota}(4.0 \pm 2.0) \quad p < 0.040$
- Item "I have been informed in time about potential problems during process simulation." ($p < 0.001$, Fig. 9):

  $C_{son}(9.0 \pm 4.5) > C_{vis}(2.5 \pm 4.75) \quad p < 0.002$

  $C_{son}(9.0 \pm 4.5) > C_{sota}(3.0 \pm 4.5) \quad p < 0.003$
- Item "How helpful were the different modes of process monitoring that have been presented in the respective parts of the experiment?" ($p < 0.002$, Fig. 10):

  $C_{son}(8.5 \pm 1.0) > C_{vis}(6.0 \pm 3.5) \quad p < 0.018$

  $C_{sota}(7.5 \pm 1.0) > C_{vis}(6.0 \pm 3.5) \quad p < 0.028$

It is interesting that the mean of the items that are related to the intrusiveness of C_{son} (see Fig. 11) is slightly higher (5.04 ± 2.1) than of those associated with the sound design being pleasing (4.18 ± 1.9). However, the feedback related to information aspects was in average quite positive (6.73 ± 1.9).

7. Discussion

7.1. Main task results

As there have been no statistically differences between the three conditions observed, neither in terms of the number of

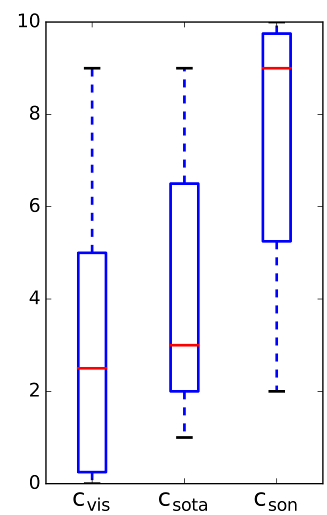

Fig. 9. Timely information about problems in process simulation.

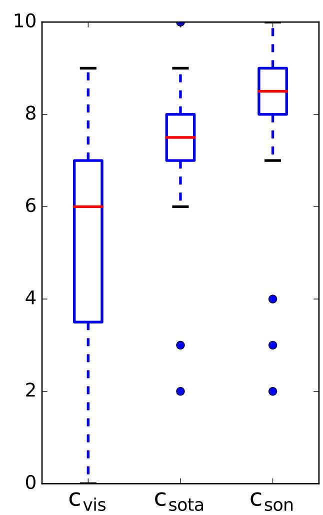

Fig. 10. Perceived helpfulness.

solved calculations, their correctness, nor concerning the overall main task score, the null hypothesis of H1.1 can be accepted.

7.2. Process monitoring results

The participants' performance in the process monitoring task was significantly higher under C_{son} than under C_{vis} and C_{sota}, while the performance under C_{sota} was in average not significantly different than that under C_{vis}. Thus, the null hypothesis of H2.1 can be accepted, while it can be rejected for H2.2. In tendency, substantially fewer interactions where performed under C_{son} compared to C_{vis} and C_{sota}, when a machine had already reached a critical state.

In general, the results of C_{son} seem to posses a lower variability than those of C_{sota} and C_{vis} concerning almost all aspects, as the results typically deviate less and contain less outliers. An explanation for this might be, that participants in C_{son} interacted less often *too late*, thus avoiding critical states more often, but also less often *too early*.

Fig. 12 reinforces these observations by showing the average R_A values for the three conditions within 10 time slots. What can be seen is that the participants achieved higher R_A values in C_{vis} and

C_{son} compared to C_{sota} in the beginning. After a short while, the R_A values for C_{sota} increase significantly and are on average higher than those of C_{vis} for the rest of the experiment. The highest scores throughout the experiment, except for a short period during the middle of the experiment, can be observed for C_{son}. In general, all three conditions show the same ups and downs throughout the experiment. However, towards the end of the experiment, a sudden drop in the R_A values of C_{vis} has been observed. Independent of these ups and downs, in tendency the R_A values of C_{sota} and C_{son} seem to be steadily improving throughout the experiment until the end, while the R_A values for C_{vis} remain more or less constant.

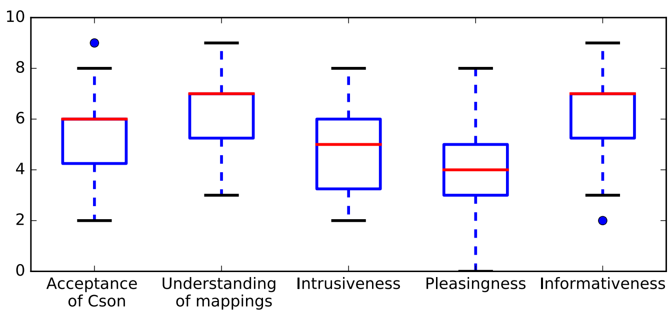

Fig. 11. Results of Likert scales related to the sound design of C_{son}.

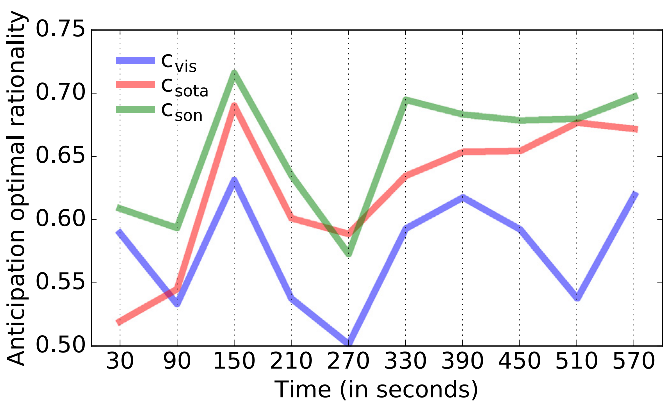

Fig. 12. Mean anticipation optimal rationalities over the course of the experiments' time in 10 time slots. Blue = C_{vis}, red = C_{sota}, green = C_{son}. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Fig. 13 shows all user interactions that have been performed by all participants for the conditions C_{vis}, C_{sota} and C_{son}. The different event types are highlighted by different symbols and at the same time different colors as explained in the caption. The plots for C_{vis} and C_{sota} look pretty similar, with events clustered more or less equally along the *y*-axes, however with larger clusterings at the lower end of the chart, indicating interactions that have been evaluated with an R_A of zero. In C_{son} on the other hand, there are almost no symbols at the bottom, but more towards the higher part, indicating higher R_A values. Furthermore, the C_{son} chart looks more organized and *tidy*. This is likely because of two factors: (a), in C_{son} there have been fewer interactions made in general, and (b) the events of the different types seem more grouped together, indicating that an interaction of the same type has more often been executed at around the same time in the experiment, and with a similar R_A value. This is partly due to the fact that as the different interactions are required at around the same intervals within the experiment (apart from a small random factor), interactions have more often been performed at the 'optimal time'.

7.3. Questionnaire results

C_{sota} was considered significantly more helpful than C_{vis}, as was C_{son}, thus in this regard the null hypothesis of H3.1 can be rejected, although C_{son} was not considered to be significantly more helpful than C_{sota}. There were no significant differences in terms of attention switching. Although there was also no significant difference in terms of estimated process monitoring performance increase between C_{vis} and C_{sota}, C_{son} scored significantly higher than both. Concerning H3.2, the null hypothesis cannot be rejected, as participants did not feel significantly more in control during C_{son} (8.0 ± 5.5), than during C_{vis} (5.5 ± 5.0) or C_{sota} (4.0 ± 4.75). As there could be no statistically significant differences in terms of perceived exhaustion observed, the null hypothesis of H3.3 cannot be rejected as well. However, several participants stated that it would be better not to use sounds that become continuously stronger,

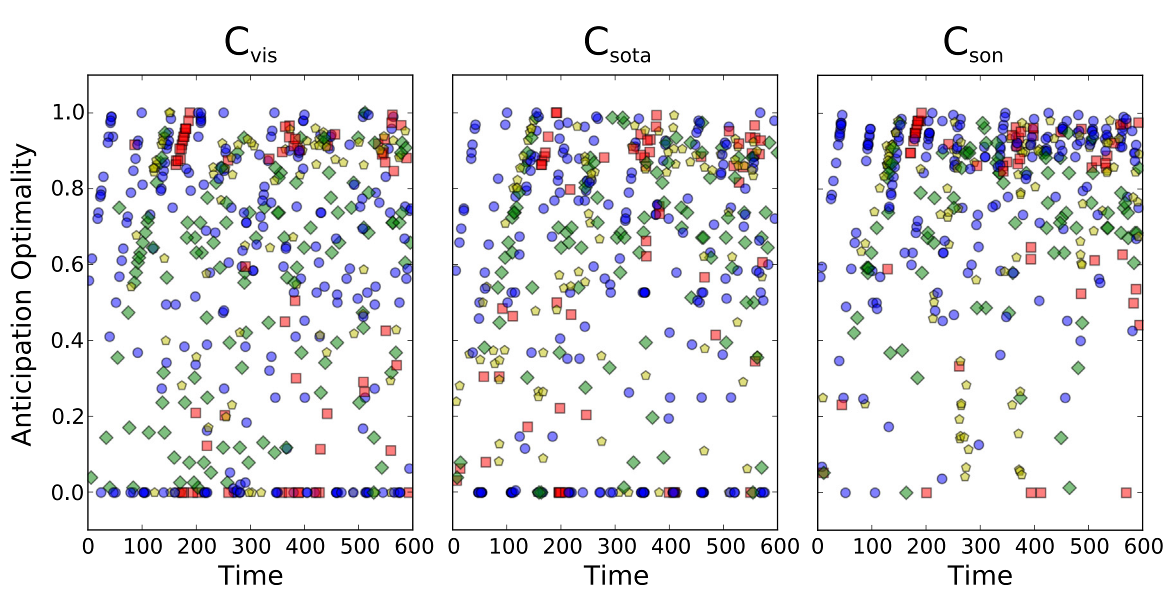

Fig. 13. All interactions of all participants during C_{vis}(left), C_{sota}(middle) and C_{son}(right) throughout the experiment time (*x*-axis), evaluated according to anticipation optimal rationality (*y*-axis). Supplies are symbolized with a blue circle, emptyings with a green diamond, maintenances of machine C with a red square, and maintenances of machine E with a yellow pentagon.

Table 3
Hypotheses of Section 3 and their answers.

Hyp	H_0/H_1	Result
1.1	H_0	No differences concerning main task performance
2.1	H_0	No differences between C_{vis} and C_{sota} in monitoring
2.2	H_1	Higher monitoring performance under C_{son}
3.1	H_1	C_{son} considered sig. more helpful than C_{vis}
3.2	H_0	Participants did not feel more in control with C_{son}
3.3	H_0	C_{sota}, C_{son} not considered more exhausting than C_{vis}

but to combine the problem-sounds with additional warning sounds (e.g. when the input stack is at only 10% etc.). All participants that commented negatively on C_{son} further indicated in their questionnaire responses that they found C_{son} to be more intrusive, less pleasing and less euphonious than the median participant (one subject even stated that he would *go crazy* listening to our sound design for a longer period of time.).

On the other hand, one participant noted after C_{sota}, that due to the fear of the appalling sounds, one automatically tries to observe the graphical representation more, thus it feels like the sounds distract more from the actual task. Another participant mentioned, that this mode (C_{sota}) was very shocking. In the mode with all sounds (C_{son}) it was a bit problematic to differentiate all sounds, but in principle this mode is the most pleasing one, as there is a smooth transition from what you are actually doing, to the monitoring. Verbal comments of two participants further suggest that their performance in C_{son} would most likely increase over time and the intrusiveness of the sounds would decrease. Almost all participants stated that they believe that sounds can in general be helpful for process monitoring (9.0 ± 2.0). The experiment data further suggests, that there is room for improvement concerning the sound selection and mapping strategies (see supplementary material). Such improvements might further increase both monitoring performance, as well as the participants' acceptance. Furthermore, there is a very strong correlation (Pearson: 0.900, $p < 0.001$) between the understanding of the mappings from data to sound applied in C_{son}, and the believe that the sonification of C_{son} is helpful. Not surprisingly, there is also a very strong correlation between the understanding of the C_{son} mappings, and the perceived informativeness of the sounds of C_{son} ($r = 0.884$, $p < 0.001$).

Table 3 provides an overview over the hypotheses described in Section 3, and their answers.

8. Conclusion

We wanted to find out how well *continuous sonification* can direct attention in comparison to *alert-based* sonification for process monitoring as a secondary task, something that to our best knowledge has not been investigated experimentally before. Furthermore, there already exist approaches that are based on continuous sonification for peripheral monitoring, but most employed sonification techniques left room to believe that they would not be considered pleasing (as they e.g. are based on synthesized sounds that are not very complex), and could lead to fatigue if listened to over a long period of time, e.g. a complete workday. Our approach therefore featured an event-based forest soundscape. We have developed a system that allows its users to compare the effectiveness of different sonifications for process monitoring in a fine-grained manner that extends beyond the typically used reaction times and binary correctness measures. The main task is simulated by means of simple arithmetic problems that have to be solved, whereas for process monitoring a simplified production process has been simulated which requires several user interactions. An experiment with 18 subjects has been conducted that compared three conditions in a within-subjects design: C_{vis} (visual only), C_{sota} (visuals + auditory alerts after reaching a critical state) and C_{son} (combining the two former with a continuous, event-based sonification that applies a forest soundscape). Each of the three experiment parts was conducted for 10 min.

8.1. Results

The main results are:

- Participants were significantly more effective in process monitoring with C_{son} compared to C_{vis} and C_{sota}.
- There were no significant differences in terms of main task performance observed between the conditions.
- Participants found C_{son} significantly more helpful for monitoring, with C_{sota} being less helpful and C_{vis} the least helpful.
- There seems to be a strong polarization concerning whether C_{son} or C_{sota} can be considered more intrusive and distracting. It is probably safe to assume, that a more pleasing sound design (e.g. more carefully selected sounds and improved mappings) could increase the acceptance of C_{son}.
- The experiment failed to prove that participants feel 'more in control of the process' under C_{son} compared to the two other conditions, but participants stated that the two modes that include audio were also not significantly more exhausting than C_{vis}.

In general, as Fig. 14 shows, it seems that a trade off between main task and process monitoring task has to be made. If the main task is of the highest importance while process monitoring can be neglected, C_{sota} seems to be the mode of choice, as it shows the highest main task scores (but the lowest monitoring performance).

On the other hand, if is not absolutely crucial, that the employees are not to be disturbed during their main task, while process monitoring is also important, C_{son} seems to be the best suited mode: C_{son} leads to slightly lower main task scores than C_{sota}, but by far the highest process monitoring results. C_{vis} on the other hand seems to be unsuitable for most cases, as it has a lower main task performance and at the same time a significantly lower process monitoring effectiveness than C_{son}.

8.2. Limitations and future work

In general, the process simulation of this experiment was designed to simulate a real-life production process. Such processes are typically 'bigger' and more complex, e.g. concerning the number of machines and the number of values that can and should be observed (e.g. temperature measures). Therefore, in real production

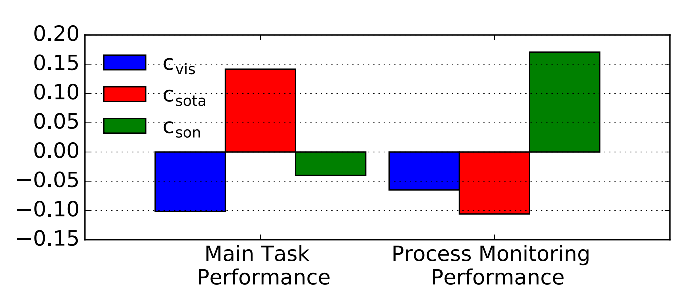

Fig. 14. Main task performance (left) versus process monitoring performance (right) for C_{vis}, C_{sota} and C_{son}. All numbers have been scaled to 0, compared to the respective mean experiment score in main and secondary task.

scenarios there can exist a significantly higher number of potentially critical states and situations, making it often difficult or nearly impossible to define them all beforehand. Such a more complex scenario would potentially be less feasible with C_{sota}, as the situations and states that would issue an auditory cue would have to be defined beforehand. A continuous sonification, that does not rely on pre-defined values and states, but instead conveys all interesting events and values might be able to better handle such a scenario. However on the one hand, the sound design of C_{son} would have to be adjusted to account for the fact that, when not all critical states and situations can be known beforehand, one can also not map the approaching of such situations to volume. Implications of this might e.g. be that a continuous sonification would have to be designed so aesthetically pleasing, that it is also acceptable when played at "normal" volumes over a complete work day. On the other hand, in a scenario with dozens of different machines, it would be difficult to distinguish and assign a unique sound to each machine. In such a scenario, the sonifications would either have to base on aggregated, process-level data (such as e.g. so called KPIs – Key Performance Indicators), or the individual machines/data points to be sonified would have to be interactively selectable by the user. With such techniques however, both a mapping of approaching critical states to volume in a continuous sonification, and an auditory-cue based sonification might be suitable again. Thus, for both approaches different mappings would have to be tested and compared for more complex scenarios, that better represent real-life working conditions as well. This could for example be achieved by modeling a more complex process simulation that has more machines, more data attributes requiring attendance, and also more different interaction possibilities.

However, an even better way would probably be to install a sonification system in a real-world monitoring context (e.g. in the control room of a factory), and let users actually use the system for a longer period of time (e.g. for several work days or weeks). With a mix of questionnaires and semi-structured interviews, specifically aspects of long-term usability and intrusiveness could be answered in more detail than it was possible in this study. Furthermore, as requested by two users, further experiments will be conducted that include a condition with more fine-grained warnings that are more pleasing in the beginning, but gradually become more intrusive. At the same time, different – potentially more pleasing – C_{son} designs will be tested to find out (a) if they are considered more pleasing and participants could imagine using them for a longer period of time and (b) if they would still enable the same level of effectiveness as C_{son} of this study. Such sonification designs could e.g. base on continuous soundscapes that are not based on short, repetitive events (such as in this experiment), but on longer, looped samples, or even *musical* concepts (see Hermann et al., 2015).

Furthermore, the literature suggests that the results of domain experts could differ from those of process monitoring novices, which would have to be evaluated as well. During this study, head tracking data has been collected. The data will be analyzed and presented elsewhere. One hypothesis is that the participants in C_{son} had to shift their focus of attention significantly less frequent than under the other conditions. Although C_{sota} would have alerted the users in case of critical situations, it can be expected that not in all cases the participants *trusted* the system to convey an alarm in time, therefore users possibly checked the visual display to be safe.

To conclude, continuous sonifications, like our forest-soundscape sonification, enhance the adequacy of interactions in peripheral process monitoring better than displays based on auditory cues and systems that rely solely on visual means, while they do not significantly affect the main task performance.

Acknowledgment

This research was supported by the Cluster of Excellence Cognitive Interaction Technology 'CITEC' (EXC 277) at Bielefeld University, which is funded by the German Research Foundation (DFG). This research was partly funded by COMET K1, FFG – Austrian Research Promotion Agency.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ijhcs.2016.06.002.

References

Alty, J.L., 1995. Can we use music in computer-human communication? In: Proceedings of the HCI'95 Conference on People and Computers X. HCI '95. Cambridge University Press, New York, pp. 409–423.
Anderson, J., Sanderson, P., 2004. Designing sonification for effective attentional control in complex work domains. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 48 (September (16)), 1818–1822, URL: ⟨http://pro.sagepub.com/content/48/16/1818⟩.
Ballora, M., Panulla, B., Gourley, M., Hall, D.L., 2010. Preliminary steps in sonifying web log datas. In: International Community for Auditory Display, Washington.
Ballora, M., Giacobe, N.A., Hall, D. L., 2011. Songs of cyberspace: an update on sonifications of network traffic to support situational awareness, vol. 8064. pp. 80640P–80640P–6. URL: http://dx.doi.org/10.1117/12.883443.
Brock, D., Ballas, J., Stroup, J., 2002. Using an auditory display to manage attention in a dual task, multiscreen environment.
Brock, D., McClimens, B., McCurry, M., 2010. Virtual auditory cueing revisited. In: Proceedings of the 16th International Conference on Auditory Display. Washington, USA. URL: ⟨http://www.nrl.navy.mil/itd/aic/sites/www.nrl.navy.mil.itd.aic/files/pdfs/brock-etal-icad2010.pdf⟩.
Crawford, J., Watson, M., Burmeister, O., Sanderson, P. M., 2002. Multimodal displays for anaesthesia sonification: Timesharing, workload, and expertise. In: Proceedings of the ESA/CHISIG Conference on Human Factors (HF2002), pp. 27–29.
Gaver, W.W., Smith, R.B., O'Shea, T., 1991. Effective sounds in complex systems: the ARKOLA simulation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching Through Technology (CHI'91). ACM, New Orleans, pp. 85–90.
Gilfix, M., Couch, A.L., 2000. Peep (the network auralizer): monitoring your network with sound. In: Proceedings of the 14th USENIX Conference on System Administration. LISA '00. USENIX Association, Berkeley, pp. 109–118.
Gopinath, M., 2004. Auralization of Intrusion Detection System using JListen (Master's thesis). Birla Institute of Technology and Science, Pilani (Rajasthan), India. URL: ⟨https://www.cs.purdue.edu/homes/apm/listen/JListen-1.0-BITS/Gopinath/MS-thesis/thesis-final.doc.
Hermann, T., 2008. Taxonomy and definitions for sonification and auditory display. In: Proceedings of the 14th International Conference on Auditory Display. Paris, France.
Hermann, T., Hildebrandt, T., Langeslag, P., Rinderle-Ma, S., 2015. Optimizing aesthetics and precision in sonification for peripheral process-monitoring. In: Proceedings of the 21st International Conference on Auditory Display (ICAD) 2015. Graz, Austria, pp. 318–319. URL: ⟨http://eprints.cs.univie.ac.at/4395/⟩.
Hildebrandt, T., Hermann, T., Rinderle-Ma, S., Nov. 2014a. A sonification system for process monitoring as secondary task. In: 2014 5th IEEE Conference on Cognitive Infocommunications (CogInfoCom), pp. 191–196.
Hildebrandt, T., Mangler, J., Rinderle-Ma, S., July 2014b. Something doesn't sound right: sonification for monitoring business processes in manufacturing. In: LABEM 2014: 1st International Workshop: Lowering the Adoption Barrier of Enterprise Modelling. IEEE, Geneva, Switzerland. URL: ⟨http://eprints.cs.univie.ac.at/4066/⟩.
Malone, T.W., Crowston, K., Herman, G.A. (Eds.), August 2003. Organizing Business Knowledge: The MIT Process Handbook. The MIT Press, Cambridge, Mass.
McClimens, B., Stroup, J.L., Ballas, J.A., Brock, D., July 2004. The design of mixed-use virtual auditory displays: recent findings with a dual-task paradigm. URL: ⟨https://smartech.gatech.edu/handle/1853/50920⟩.
McClimens, B., Brock, D., Brock, D., Brock, D., 2011. Modeling Attention Allocation in a Complex Dual Task With and Without Auditory Cues.
Poguntke, M., Ellis, K., May 2008. Auditory attention control for human-computer interaction. In: 2008 Conference on Human System Interactions, pp. 231–236.
Rauterberg, M., Styger, E., 1994. Positive effects of sound feedback during the operation of a plant simulator. In: Selected Papers from the 4th International Conference on Human - Computer Interaction. EWHCI '94. Springer, London, pp. 35–44.
Renwick, J.T., Babson, P.E., 1985. Vibration analysis—a proven technique as a predictive maintenance tool. IEEE Trans. Ind. Appl. IA 21 (2), 324–332.

Sanderson, P., Crawford, J., Savill, A., Watson, M., Russell, W.J., 2004. Visual and auditory attention in patient monitoring: a formative analysis. Cognit. Technol. Work 6 (April 3), 172–185, URL: ⟨http://link.springer.com/article/10.1007/s10111-004-0159-x⟩.
SAP SE, 2015. Monitoring PP/DS Alerts – SCM Basis – SAP Library. URL: ⟨https://help.sap.com/saphelp_scm50/helpdata/en/a5/cf893744ce9c71e10000009b38f842/content.htm⟩.
Sauer, O., 2004. Modern Production Monitoring in Automotive Plants. In: World Automotive Congress. Book of Abstracts: 23–27 May 2004, Barcelona.
Seagull, F.J., Wickens, C.D., Loeb, R.G., 2001. When is less more? Attention and workload in auditory, visual, and redundant patient-monitoring conditions. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 45 (18), 1395–1399.
Siemens AG, 2007. Siemens SIMATIC PCS 7. Technical Report. URL: ⟨http://w3.siemens.com/mcms/process-control-systems/SiteCollectionDocuments/efiles/pcs7/pdf/76/kg_stpcs7_v70_en.pdf⟩.
Vickers, P., 2011. Sonification for process monitoring. In: Hermann, T., Hunt, A., Neuhoff, J. G. (Eds.), The Sonification Handbook. Logos, Berlin, pp. 455–492.
Vickers, P., Laing, C., Fairfax, T., July 2014. Sonification of a network's self-organized criticality. [cs] arXiv:1407.4705ArXiv: 1407.4705.
Watson, M., Sanderson, P., 2004. Sonification supports eyes – free respiratory monitoring and task time – sharing. Hum. Factors: J. Hum. Factors Ergon. Soc. 46 (January (3)), 497–517, URL: ⟨http://hfs.sagepub.com/content/46/3/497⟩.
Watson, M.O., Sanderson, P.M., 2007. Designing for attention with sound: challenges and extensions to ecological interface design. Hum. Factors: J. Hum. Factors Ergon. Soc. 49 (January (2)), 331–346, URL: ⟨http://hfs.sagepub.com/content/49/2/331⟩.
Watson, M., Sanderson, P.M., Woodall, J., Russell, W.J., 2003. Operating theatre patient monitoring: the effects of self paced distracter tasks and experimental control on sonification evaluations. Ergon. Soc. Aust., 128–137, URL: ⟨http://espace.library.uq.edu.au/view/UQ:99479⟩.