

Different effects of auditory feedback in man-machine interfaces

Matthias Rauterberg

Institute of Hygiene and Applied Physiology (IHA)
Swiss Federal Institute of Technology (ETH)
Clausiusstrasse 25, CH-8092 ZURICH
+41-1-632-7082, rauterberg@iha.bepr.ethz.ch

Abstract

Two experiments were carried out to estimate the effect of sound feedback (e.g., auditory alarms) on the operator's work of a man-machine system. The first experiment tested the hypothesis of additional, but redundant auditory feedback. Twelve subjects were instructed to define queries on a simple database. In this experiment, we could not find a general superiority of auditory feedback. In a second experiment we designed a process simulator so that each of eight machines made tones to indicate its hidden status over time. Eight students of computer science operated this process simulation program of an assembly line with robots. Relevant information of disturbances and machine breakdowns was given in auditory forms, too. The results indicate, that the additional feedback of auditory alarms improves significantly the operator performance and increases positively some mood aspects.

Keywords: Auditory feedback, audible alarms, computer simulation, man-machine interaction.

1 Introduction

The hearing of sounds (e.g., alarms) is based on the perception of events and not on the perception of sounds as such. For this reason, sounds are often described by the events they are based on. Sound is a familiar and natural medium for conveying information that we use in our daily lives, especially in the working environment (Momtahan et al. 1993). The following examples help to illustrate the important kinds of information that sound can communicate (Mountford & Gaver 1990):

- *Information about normal , physical events* – most of all real world events (e.g., putting a pencil on the desk) is associated with a (soft-)sound pattern. This type of auditory feedback is parallel to the visual feedback of the actual status of all interacting objects (e.g., pencil and desk).
- *Information about [hidden] events in space* – all audible signals out of our visual field (e.g., footsteps warn us of the approach of another person).
- *Information about abnormal structures* – a malfunctioning engine sounds different from a healthy one and/or alarms in supervisory control environments.
- *Information about invisible structures* – all hidden structures that can be transformed into audible signals (e.g., tapping on a wall is useful in finding where to hang a heavy picture).
- *Information about dynamic change* – all specific semantics of real world dynamics (e.g., as we fill a glass we can hear when the liquid has reached the top).

The textual representation of information is of most use when the operator is familiar with the domain area and can demonstrate much experience and knowledge in that domain area (Marmolin 1992). In comparison, more concrete (visual and audible) repre-

sentations of information that the operator can query are of most use when the domain area is new and unknown. Most of all user interfaces stresses the visual perception. Auditory feedback can probably help to reduce eye strain. New possibility for the interactive representation of complex sound generating events and processes are possible, especially in multimedia interfaces.

The parallel use of different media and the resulting parallel distribution of information, for example by simultaneously showing a predecessor through a concrete representation and its explanation through audio distribution, leads to a denser sharing of information. In this case, the operator can dedicate his attention solely to the visual information, which has parallel audio support. This reduces the need to change the textual or other visual delivery and prevents the overflow of visual information (Edwards 1988).

Sounds can be utilised to improve the operators' understanding of visual predecessors or can stand alone as independent sources of information. Gaver, Smith and O'Shea (1991) used sounds as diagnostic support applied with the direction of a process simulation. But, they did not prove the hypothesis that an interface with audible feedback is superior to an interface without sound feedback. The authors describe only some global impressions of different operator reactions to sound feedback.

In the context of supervisory control an alarm is a signal, that informs the operator of a dangerous or problematic process state. Wanner (1987) classified alarms in the following two categories: programmed and not-programmed alarms. The first alarm class is divided by Riera et al. (1995) into two groups: (1) *breakdown alarms*, that correspond to internal failures of components, and (2) *process alarms*, that show an abnormal performance of a process. The not-programmed alarms are not defined at the time of system design, but these audible cues are used by the operator (e.g., abnormal noise, smoke, steam, explosion etc.).

Stanton, Booth and Stammers (1992) classify alarms by their input modality (visual vs. auditory) and their information processing code (verbal vs. spatial). Spatial alarm processing requires a manual response to maximize performance while a verbal alarm requires a vocal response. Typical problems with alarms are: "the avalanche of alarms during a major transient or shift in operating mode, standing alarms, alarm inflation, nuisance alarms, and alarms serving as status messages" (Stanton et al. 1992, p. 87).

First, we investigated the effects of auditory feedback for a situation where the sound is given additionally to the visual feedback. In the second experiment we investigated the effects of auditory feedback of hidden events which were produced by a continuous process in the background.

2 Experiment-I: Individualized sound feedback

We carried out a first experiment to test the hypothesis that sound feedback is particularly helpful, if the user can choose his or her individually preferred sound pattern as a redundant information additionally to the visual output on the screen. To test this hypothesis we implemented a simple database system: a stack with cocktail recipes. In the sound feedback condition, the result of each database query was associated with an additional and customized sound pattern.

2.1 Method

Subjects: Twelve subjects (4 female, 8 male; mean age 22 ± 2 years) were instructed to define queries on a database to find a particular recipe. The experience of all test subjects with computer systems and cocktail shaking was equally distributed over both test conditions (determined by a test design with repeated measurements).

Material: The database has a direct manipulative interface and contains 352 different cocktail recipes. The database was implemented under HyperCard on an Apple Mac-

intosh IIfx. The user could choose his or her preferred sound feedback with the customisation interface (see Figure 1). Each discrete result feedback was one of 49 different sound patterns for the following six output conditions: 'fit exactly', 'fit except one part', 'fit except 2 parts', 'fit except 3 parts', 'fit except more than 3 parts', 'does not fit'. The user could choose among the three classes: noise, speech, and music.

<input type="radio"/> No Sound	<input checked="" type="radio"/> Noise	<input type="radio"/> Speech	<input type="radio"/> Music
<input type="radio"/> fit exactly:.....	Spaceballs		
<input type="radio"/> fit except 1 part	Triangle		
<input type="radio"/> fit except 2 parts	Rings		
<input type="radio"/> fit except 3 parts	Dishcrash		
<input type="radio"/> fit except ≥ 4 parts	Cow		
<input type="radio"/> does not fit	Wuääh		
	Train Whistle		
	beduwee		
	Ping		
	crinkle		
	Sam SysMeow		
	Nuke		
	Cricket		

Figure 1. The customisation interface to select the individual sound feedback.

Tasks: Each test subject was introduced with the following scenario: "You are in front of your housebar. The bar contains the following components... (e.g., gin, rum, apricot brandy, lemon, sugar). Which four different cocktails can be produced with different components each?". The task was to search for appropriate cocktails, when the components are given (e.g., type of liquor, type of juice, etc.). Each user had to look for four different recipes. The database system output all possible recipes that fulfill at least one selection criteria. In general: the more selection criteria were fulfilled, the better the solution was. To support the user finding the best recipe, result feedback with the customized sound pattern was provided.

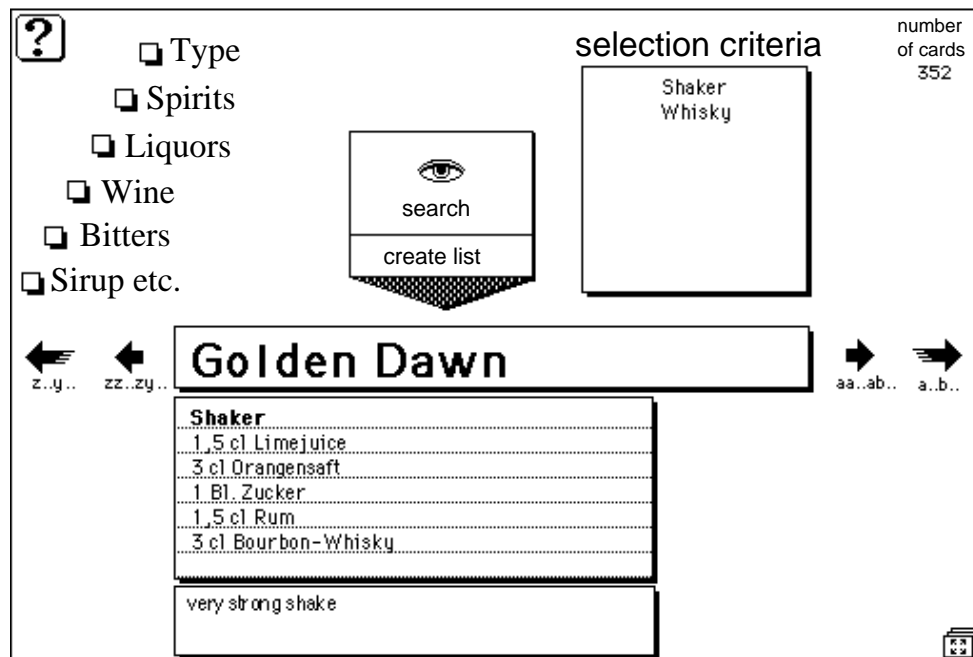


Figure 2. The user interface of the cocktail database.

Procedure: First, each user chose individually the most convenient sound for each output condition from one of the three sound classes (see Figure 2). We run the experiment with a two-factorial test design. Factor A was 'with' or 'without' audible feedback. Test condition 1 was only visual feedback given in form of an inverse output of all 'selection criteria' that were part of the selected recipe. Test condition 2 was visual and auditory feedback about all involved 'selection criteria'.

Factor B was a repeated measurement design. Six subjects started the experiment without auditory feedback (test condition 1) and repeated a very similar task with audible feedback (test condition 2). The other six subjects started with audible feedback (test condition 2), and repeated the task without auditory feedback (test condition 1). Both task-solving trials lasted maximally 15 minutes each. Each individual session took about 60 minutes.

Measures: The *first dependent variable* is the total search time (in seconds). The *second dependent variable* is the search time per recipe (in seconds). The *third dependent variable* is the number of dialog operators ('# of dialog operators'; e.g., mouse clicks). Before and after each task trial the user has to answer a mood questionnaire (eight scales with overall 36 items as monopolar rating scales). After the task trial with sound feedback we measured each personal opinion (subjective 'sound preference' questionnaire with five monopolar rating scales).

2.2 Results

The result of the personal selection, the individual customization shows the following distribution: speech in 42%, music in 25%, and noise in 33%. Most of the users' sound selections did not reflect the ordinal rank structure of the search results (seven sound customizations without ordinal structure, four sound customizations with ordinal structure).

The results of the three dependent measures (total search time, search time per recipe, and number of dialog operators) showed, that the overall performance of direct manipulative database queries could not be significantly improved, when additional, redundant feedback was given in an audible form (see Table 1).

Table 1: Results of the three dependent variables that measure users' performance of database queries for both test-conditions 'with sound feedback' and 'without sound feedback' (N=12).

Variable	With sound	Without sound	P sign.
search time total	685 ± 193	649 * 288	.709
search time per recipe	201 ± 99	182 ± 113	.627
# of dialog operators	98 ± 38	112 ± 60	.577

In a second analysis of subjective sound preferences we distinguished between those users that prefer sound (N=5) and those users that do not prefer sound (N=7). For this post-hoc analysis we constructed the dimension 'preference' with two levels: 'sound preferred' versus 'sound refused'. Overall both test-conditions ('with' and 'without' sound feedback), when sound was preferred, the mean search time is 633 ± 248 sec; mean search time, when sound was not preferred (refused, resp), is 692 ± 241 seconds. This difference is not significant ($F(1,10) = 0.28, p \leq .609$).

But, if we analyse only the test-condition 'with sound feedback' then we get the following result for the variable 'search time total' between the two groups 'preferred' versus 'refused' (see Table 2): mean search time for 'sound preferred' is 544 ± 215; mean search time for 'sound refused' is 786 ± 95. This difference is significant ($F(1,10) = 7.18, p \leq .023$). Similar results are found for the two other dependent variables, too.

Table 2: Results of the three dependent variables that measure users' performance of database queries only for the test-condition 'with sound feedback' (N=6).

Variable	Sound preferred	Sound refused	P sign.
search time total	544 ± 215	786 * 95	.023
search time per recipe	136 ± 54	246 ± 99	.046
# of dialog operators	71 ± 26	117 ± 32	.029

The significant difference of the variable 'total search time' can be explained by the significant interaction term between the dimension 'test-condition' and the dimension 'preference' (mean search time for 'with sound' and 'sound preferred' is 544 ± 215; mean search time for 'without sound' and 'sound preferred' is 721 ± 270; mean search time for 'with sound' and 'sound refused' is 786 ± 95; mean search time for 'without sound' and 'sound refused' is 570 ± 310; $F(1,10) = 5.04, p \leq .049$).

2.3 Discussion of Experiment-I

The non significant outcome of this comparison study could be explained by the uncontrolled factor 'individualisation': most of the users did not consider the implicit rank order in the search results. On one side we can assume a general, positive effect on the performance determined by customization, but on the other side this customization should reflect important structures of the task itself.

In the post-hoc analysis we divided our test sample into two groups: users that liked their sound choices, and users that disliked their own choices. The results of this experiment-I show, that auditory feedback can be helpful, but only if the user chooses a sound pattern that he or she really likes.

3 Experiment-II: Sound feedback of hidden events

Our main interest in this second experiment was to test the hypothesis of (Buxton 1989) and (Gaver et al. 1991), that human operators in a 'real' process control situation monitor multiple background activities simultaneously through auditory sound feedback (*tones* as auditory and spatial alarms, see Stanton et al. 1992). So, we designed a system that produces audible cues and tones to help operators to monitor the status of ongoing processes.

Diagnosing and treating problems with the plant were aided by alert sounds (e.g., breakdown alarms; see also Gaver et al. 1991). We carried out this experiment, to test our hypothesis in laboratory environment with a high alarm rate during a supervisory control task.

3.1 Method

Subjects: Eight male students of computer science took part in the experiment as untrained operators (mean age of 24 ± 1 years). The experience of all test subjects with computer systems and simulation software was equally distributed over both test conditions (determined by a test design with repeated measurements).

Simulator: The simulation is based on a flexible manufacturing system, that produces cases made of aluminium (see 'work pieces' in Figure 3). The whole system consists of eight computer numeric controlled (CNC) manufacturing centres and eight loading robots for these centres. In the input directing station all work pieces are automatically directed on the assembly line. The assembly line transports each work piece through different stations to the CNC manufacturing centres and back to the output directing station. The whole plant was deliberately designed to be too large to fit on the computer screen, so operators could only see about half the robots and CNC machines at any time (see 'actual screen clipping' in Figure 3).

A work piece could have one of the following status: (1) loading on the assembly line at the input directing station, (2) transportation on the assembly line, (3) fixation on the carrier at the reset station, (4) final fixation and twist on the carrier, (5) fixation on a pallet with three other work pieces at the robot, (6) processing one of two sides in the CNC station, (7) change from one side to the other at the reset station, (8) to be provided with a serial number at the labeling station, (9) loading off the assembly line at the output directing station. Steps (3) to (7) are carried out twice, once for each side of the work piece.

We designed our simulator so that each of the machines made tones to indicate its status over time. Each tone was designed to reflect the semantic of the actual event. For instance, a splashing tone indicated that cooling liquid was being spilled. Because of the complexity of our system, as many as 38 tones made be placed at once.

We attempted to design the sounds so that none would be masked (rendered inaudible) by other sounds. Gaver, Smith and O'Shea (1991) describe two strategies to be useful in avoiding masking. First, sounds were spread fairly evenly in frequency, so that some were high-pitched and others lower. Second, we avoided playing sounds continuously and instead played repetitive streams of sounds, thus maximising the chance for other sounds to be heard in the gaps between repetitions. CNC 0 and CNC 4 are characterised by a high-pitched sound. CNC 3 and CNC 7 are low-pitched (cf. Figure 3).

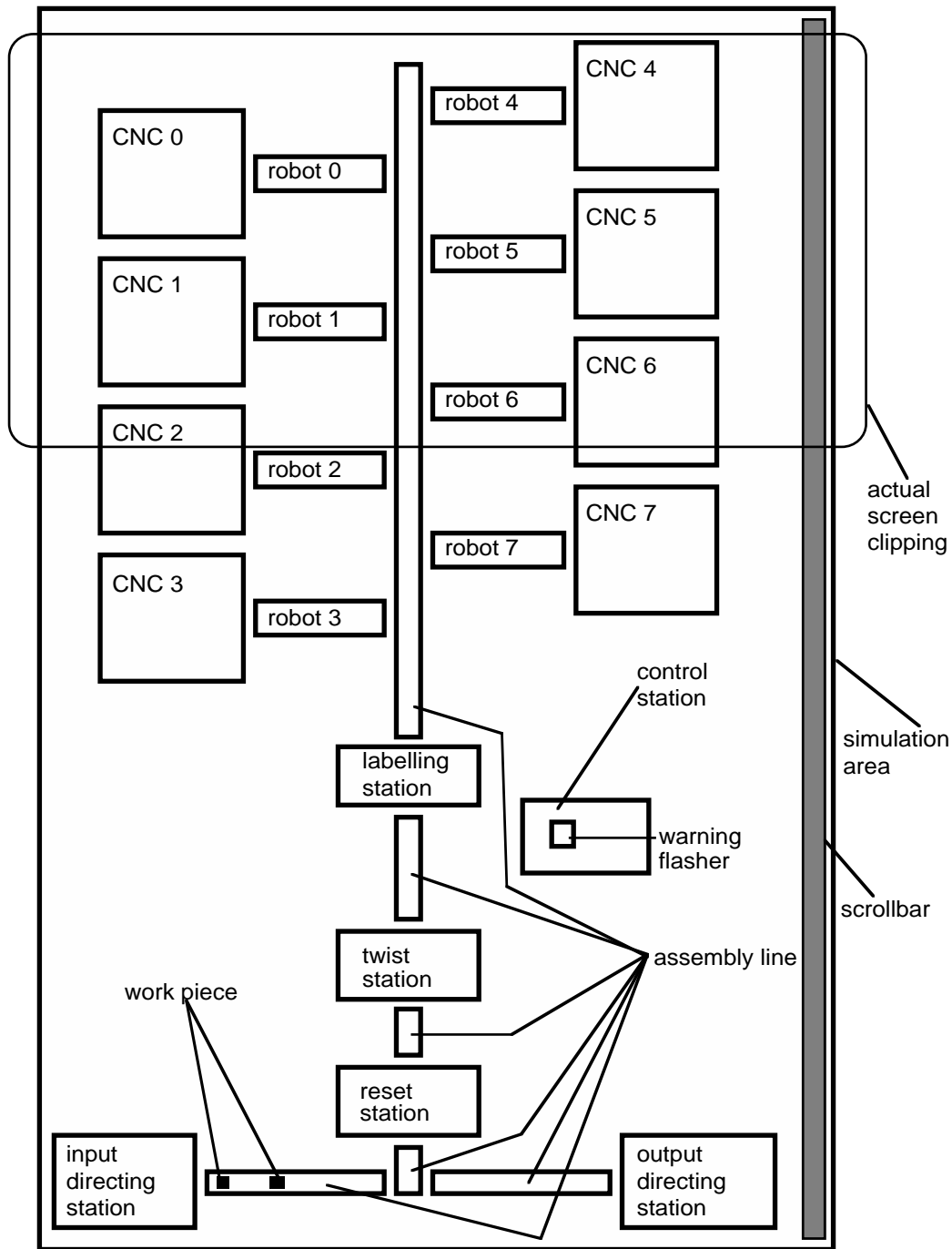


Figure 3 The groundfloor of the plant simulator. The rectangle shows the actual screen output each operator sees at a given time.

Normal running of a machine was coupled with a characteristic sound pattern. Each machine breakdown generated instead of the normal sound a specific alert tone: the auditory alarm (see Table 3). If a robot or a CNC centre breakdown occurs, then this centre can not process the pallet of four work pieces further on.

The breakdown of a machine, that will not be repaired immediately, leads to a jam on the assembly line. The most important—but not dangerous—consequence of an overlooked alarm is the decrease of the productivity and performance of the whole plant.

Table 3: Sound types, alarm, duration, and size (KB = kilo byte).

machine	sound	alarm	duration	size
CNC 0-7	normal	no	1.20 s	51 KB
robot 0-7	normal	no	0.39 s	16 KB
input station	normal	no	0.41 s	17 KB
output station	normal	no	0.78 s	33 KB
reset station	normal	no	1.40 s	60 KB
twist station	normal	no	0.40 s	17 KB
labelling station	normal	no	0.49 s	21 KB
CNC 0-7	no cooling	yes	1.08 s	46 KB
CNC 0-7	jammed pipe	yes	1.38 s	59 KB
robot 0-7	lost piece	yes	1.04 s	44 KB
robot 0-7	tear off pipe	yes	1.04 s	44 KB
control station	warning	yes	0.24 s	10 KB

Material: We ran the experiment on an IBM compatible PC (Olivetti® i386, 25 MHz, 6 MByte main storage, 17" VGA color screen) with an extra sound card (Logitech® 16 Bit, 44 kHz, stereo). A special simulation program was developed in Turbo Pascal® 1.0 to present the signals on the screen. Operators heard the auditory alarms out of two small active speakers (maximal 3 watt). All machines at the left side (see Figure 3) could be heard out of the left speaker. The right speaker gave out the sound of all machines at the right side.

Task: Subjects were instructed to operate a plant simulator and to take care for a high productivity rate. The task was to trouble-shoot the whole manufacturing system. First, each subject had to detect that a breakdown happened. Then he has to find the interrupted machine (robot or CNC machine). The actual breakdown event shows the operator how to repair the machine. The operator can get this information visually in a modal dialogue box with the status report at the control station or in an audible form through auditory alarm feedback. The choice of one of both possibilities was a tradeoff for the user: the costs for the visual feedback of the status report was the scrolling effort to the control station; the costs for the auditory feedback was the memory load to discriminate the different sound pattern.

A CNC machine could have two breakdown events ('jammed outlet pipe of cooling agent', 'empty cooling agent'). A robot could breakdown with two different events ('lost work piece', 'tear off a pressure pipe').

Each interrupted machine could be repaired by entering an appropriate repair code (a four-digit number, see Table 4) in a repair dialogue box located at the machine. The operator sees only a part of the whole plant (see 'actual screen clipping' in Figure 1). He moves the actual screen up and down by clicking with the mouse in the scrollbar area to 'go to' the interrupted machine. A mouse click on the machine symbol pops up the repair dialog box. Entering the correct repair code transfers the interrupted machine in the normal state. If an incorrect repair code is entered, then no internal state change happens and the operator could hear only a short beep.

Table 4: All breakdown types that lead to an alarm, and their repair codes.

Machine	breakdown	code
CNC 0-7	no cooling	3713
CNC 0-7	jammed pipe	8319
robot 0-7	lost piece	1731
robot 0-7	tear off pipe	1733
control station	status request	8700

Operators' view of the plant behavior was that robots and CNC centres breakdown accidentally. The plant simulator was programmed so, that all breakdowns appeared in the same sequence. This approach guarantees that the trials between operators are maximally comparable.

Procedure: We run the experiment with a two-factorial test design. Factor A was 'with' or 'without' audible feedback. Test condition 1 was only visual alarm feedback with a warning flasher and a modal dialogue box with status information of each manufacturing system located at the control station. Test condition 2 was visual and/or auditory alarm feedback of each machine breakdown.

Factor B was a repeated measurement design. Four subjects started the experiment without auditory alarm feedback (test condition 1) and repeated the same task with audible [and visual] feedback (test condition 2). The other four subjects started with audible [and visual] feedback (test condition 2), and repeated the task without auditory alarm feedback (test condition 1).

Each subject filled out a questionnaire to estimate the individual experiences with computers (about 10 minutes). The subjects were introduced in operating the simulation tool through 'learning by using' (about 15 minutes). The simulation ran for the trouble-shooting task exactly 20 minutes. Before and after each trouble-shooting task the operator has to answer a mood questionnaire (eight scales with overall 36 items as monopolar rating scales). This mood questionnaire measures the mental workload at a rough estimate. After each trouble-shooting task we measured the subjective satisfaction with a semantic differential (11 bipolar items). Each individual session took about 90 minutes.

Measures: Our first dependent variable is a point scale that measures the productivity of the plant. Each work piece, that entered the assembly line at the input direction station, counts one point. One point is counted for each side, that was processed at a CNC machine (see chapter 2.2). Each work piece, that left the assembly line at the output direction station, counts an extra point. Each work piece on the assembly line counts one to four points. The productivity score after 20 minute's simulation time is the sum over all work pieces that entered the assembly line.

The second dependent variable is the number of requested status reports at the control station. The third and fourth dependent variables are number of correct and number of incorrect repairs. The eight scales of the mood questionnaire and the 11 items of the semantic differential are dependent variables to measure operators' satisfaction.

2.2 Results

First, we present the results of the four dependent variables that measure operators' trouble-shooting activities. We find a significant difference between the two test-conditions for two of four dependent measures ('productivity score' and '# of status reports'; see Table 3).

Table 3: Results of the four dependent variables that measure operators' trouble-shooting activities for the two test conditions: with or without auditory alarm.

Variable	With alarm	Without alarm	P sign.
productivity score	70 ± 5.6	65 ± 5.3	.052
# of status reports	17 ± 5.8	23 ± 4.0	.032
# of correct repairs	36 ± 2.5	36 ± 2.3	.999
# of incorrect repairs	16 ± 11.0	9 ± 7.1	.184

Without auditory alarm feedback operators moved to the control station and requested the status report significantly more than in the test condition with sound feedback (see Table 3). We could observe, that most of the operators in test condition 2 with auditory alarm go first to the control station to look for all breakdowns (visual feedback), and go after that through the whole plant to repair machine by machine. During this walk through they could remember all machines which have to be repaired by listening to the different sound pattern of each alarm type.

On one side, we can observe a significant improvement through auditory alarm feedback, on the other side we can find, that operators perceive the simulation with auditory alarms more intransparent and feel slightly more confused than without auditory alarms (see Table 4).

Table 4: Results of the eleven items of the semantic differential for the two test conditions: with or without auditory alarm. (bipolar rating scale: -2, -1, 0, +1, +2)

Variable (-)	(+)	With alarm	Without alarm	P sign.
time-consuming.....	time-saving	-1.1 ± 0.7	-1.0 ± 0.9	.791
rigid.....	flexible	-0.9 ± 1.3	-0.8 ± 0.8	.735
circumstantial....	simple	+0.5 ± 2.3	+0.4 ± 3.1	.889
intransparent....	transparent	+0.4 ± 1.1	+1.4 ± 0.6	.064
confuse.....	unequivocal	+0.1 ± 2.7	+1.1 ± 1.0	.179
unclear.....	clear	0.0 ± 2.6	-0.4 ± 1.4	.596
complicated.....	uncomplicated	0.0 ± 1.1	-0.3 ± 1.9	.712
prescribed.....	free	-0.5 ± 0.9	-0.4 ± 1.1	.816
unforeseeable....	foreseeable	0.0 ± 2.3	+0.1 ± 1.8	.871
unsusceptible....	susceptible	-0.8 ± 1.1	-0.9 ± 1.0	.781
angry.....	pleasing	-0.4 ± 1.7	-0.1 ± 1.3	.709

Operators felt significantly more self-assure and more social accepted after working with auditory alarm feedback than without auditory feedback (see Table 5). Their readiness for endeavour, restfulness, and mood increased in the test condition with sound.

Table 5: Results of the differences (after - before) of the eight scales of the mood questionnaire for the two independent test conditions: with or without auditory alarms. (monopolar rating scale).

Variable	With alarm	Without alarm	P sign.
readiness of endeavour	+2.4 ± 4.1	-0.5 ± 4.1	.199
restfulness	+1.3 ± 2.7	+0.4 ± 3.3	.589
readiness for contacts	+0.9 ± 2.5	-0.8 ± 2.2	.219
drowsiness	-1.1 ± 2.4	-1.5 ± 3.2	.801
self-assurance	+1.8 ± 2.0	-0.6 ± 1.7	.022
social acceptance	+0.1 ± 1.0	-1.1 ± 1.0	.031
to feel excited	0.0 ± 6.1	-1.0 ± 5.9	.738
mood-laden	+1.3 ± 2.2	-0.3 ± 1.0	.128

If we assume that the mood questionnaire measures the mental workload at a rough estimate then we can suppose that in this investigation the auditory alarm feedback does not increase the mental strain. It is quite unclear to which number of different alarms this assumption will be correct.

3.3 Discussion of Experiment-II

The results of this experiment give us a strong support for the assumption that continuous auditory feedback of hidden events can improve the productivity of man-machine interactions. Not only performance can be improved, but also some mood aspects (e.g., 'self-assurance', and 'self-acceptance'). One possible explanation of this outcome can be given in the way that the continuous auditory feedback help the operator during his 'walk through the plant' to remember which machine has to be repaired next, even if the total sound pattern is complex (e.g., several alarms overlapped with all other normal sounds). The result, that operators feel the test condition with auditory feedback more intransparent than the test condition without sounds, can be interpreted in the way that auditory feedback supports the operator's task solving process on a pre- or unconscious level, and not on a conscious one.

4 GENERAL DISCUSSION

The sense of hearing is an all-round sense. This aspect is an important difference to visual perception, that is a directional sense. An auditory interface can be much larger than the visual interface (screen). Visually hidden aspects of parallel processes in the background can be made perceptible with auditory feedback (Cohen 1993). The results of our experiment support this design approach. Auditory feedback of concurrent processes, that are important for task solving, improves the usability of interfaces.

Audition is a spatial sense; we can be aware simultaneously of many sounds coming from different locations. But spatial patterns in audition are much more limited than those of vision. It is primarily a time sense, for its main patterns are those of succession, change, and rhythm. Auditory feedback typically arrives sequentially in time, whereas visual pattern may be presented either sequentially or simultaneously. Of course many perceptual experiences depend on the operation of several senses at once; then the prominence of sense over another becomes a matter for study (Hartman 1961).

Auditory feedback has poor 'referability', meaning that they usually cannot be kept continuously before the operator, although they can be repeated periodically. Visual patterns offer good referability, because the information usually can be 'stored' in the dis-

Publish in: N.A. Stanton and J. Edworthy (1999, eds.), Human Factors in Auditory Warnings (pp. 225-242). Aldershot, UK: Ashgate Publishing

play. One the possible advantage of auditory feedback is its 'attention-demanding'; it 'breaks in' on the attention of the operator. Visually stimuli, however, do not necessarily have this captive audience. The operator has to be looking toward the display in order to perceive the stimulus. Hearing is somewhat more resistant to fatigue than vision (McCormick 1957, p. 427).

How many different concurrent tones can be discriminated? Operators reacted up to 38 different tones in our simulation study. Momtahan et al. (1993) could show that staff in operating rooms was able to identify only a mean of between 10 and 15 of the 26 alarms. Nurses were able to identify only a mean between 9 and 14 of the 23 alarms found in their intensive care unit. Momtahan et al. explain their results with the poor design of auditory warning signals. Standardisation of auditory feedback can minimise this perceptual problem.

Cohen (1993) found that it is a difficult task to design tones "which tell the right story and are also pleasant and emotionally neutral." Good auditory feedback needs sound patterns that are interpretable without visual redundancy (e.g., door creaks open, door slams). We have to look for sound patterns that 'stand for themselves'. Given these sounds we have to map them in a metaphorical sense to new events introduced by technology (e.g., door creaks open => login, door slams => logout; see Cohen 1993). For simulation tools, that deal with real world events, we can easily use the corresponding real world sounds.

The results of our study support the 'real sound' approach. To avoid boredom and fatigue—caused by outputting always the same sound pattern—the design of tones for auditory feedback should be highly context sensitive. E.g., listening to everyday sounds is based upon the perception of events and not upon the perception of sounds in and of themselves. This fact becomes clear in the following example (Rauterberg et al. 1994):

"A pen dropped upon a piece of paper from a height of about 15 cm created a different sound than when it is dropped upon the hard surface of a desk. An altogether different sound is created when a rubber eraser is dropped upon the paper or, respectively, on the desk."

The sound created in each case of the previous example is neither a characteristic of any of the participating objects (pen, rubber eraser, sheet of paper, desk surface) nor a characteristic of the event 'dropped' itself. The four different sounds in the examples are, with an observation that holds true to the reality of the situation solely determined by their respective interaction and environmental conditions. Most of the natural sounds are a result of one or more interactions between two or more objects in a definite place and in definite surroundings and can be defined as the following:

Auditory feedback = f (process objects, interaction, process environment)

Every interaction possesses attributes that have an influence on the produced sound (cf. Darvishi et al. 1995). A framework concept for the description of auditory feedback is needed, in which auditory alarms can be represented as auditory signal patterns along several descriptive dimensions of various objects interacting together in a certain environment (cf. Munteanu et al. 1995). This approach is appropriate especially for the design of auditory feedback signals of the process alarms. To make auditory alarms context sensitive leads directly to a design strategy that reduces the number of context-free alarms (cf. the discussion of 'reduction techniques' in Stanton et al. 1992).

5 CONCLUSION

The results of these experiments showed, that the performance of operating a plant simulator could be significantly improved, when feedback of machine break downs and

Publish in: N.A. Stanton and J. Edworthy (1999, eds.), Human Factors in Auditory Warnings (pp. 225-242). Aldershot, UK: Ashgate Publishing

other disturbances was given in an audible form, too. We can also observe a significant increase of different aspects of operator's mood. Overall, we can say that operators feel better and less stressed with sound feedback, than without sound.

We found that auditory alarm feedback was effective in the following way. Auditory alarm feedback helped operators keep track of the ongoing processes. Auditory alarms allowed operators to track the activity, rate, and functioning of normally running machines. Without auditory feedback, operators overlooked machines that were broken down. With auditory feedback these problems were indicated either by the machines' sound ceasing or by the various alert sounds. Continuous auditory feedback allowed operators to hear the plant as an integrated complex process. The sounds merged to produce an auditory pattern, much as the many sounds of everyday machines.

Using non speech sounds to provide system information is appealing for several reasons. First, by adding sound to the interface the bandwidth of communication can be significantly increased. Second, the information conveyed by sounds is complementary to that available visually, and thus sound can provide a mean for displaying information that is difficult to visualise, especially with limited screen real estate. Auditory alarm feedback can help to improve the usability of interfaces in the following ways: most of all interfaces stresses the visual perception, so that auditory feedback can help to reduce eye strain and fatigue.

Acknowledgements

We have to thank the following persons for their generous support: E. Styger, A. Baumgartner, A. Jenny, and M. de Lange for developing the software and all students participating as test subjects.

6 REFERENCES

- Buxton, W. (1989) Introduction to this special issue on non speech audio. *Human-Computer Interaction* 4(1): 1-9.
- Cohen, J. (1993) "Kirk Here:" Using Genre Sounds To Monitor Background Activity. In S. Ashlund, K. Mullet, A. Henderson, E. Hollnagel and T. White (eds.) INTER-CHI'93 Adjunct Proceedings. (pp. 63-64), New York: ACM.
- Darvishi, A., Munteanu, E., Guggiana, V., Schauer, H., Motavalli, M. & Rauterberg, M. (1995) Designing environmental sounds based on the results of interaction between objects in the real world. In K. Nordby, P. Helmersen, D. Gilmore & S. Arnesen (eds.) *Human-Computer Interaction – Interact'95*. London: Chapman & Hall, pp. 38-42.
- Edwards, A. (1988) The design of auditory interfaces for visually disabled users. In E. Soloway, D. Frye and S. Sheppard (eds.) CHI'88 Conference Proceedings "Human Factors in Computing Systems". (pp. 83-88), New York: ACM.
- Gaver, W. (1989) The Sonic Finder: an interface that uses auditory icons. *Human Computer Interaction* 4(1):67-94.
- Gaver, W., Smith, R. & O'Shea, T. (1991) Effective sounds in complex systems: the ARKola simulation. In S. Robertson, G. Olson and J. Olson (eds.) CHI'91 Conference Proceedings "Reaching through Technology". (pp. 85-90), Reading MA: Addison-Wesley.
- Hartman, F. (1961) Single and multiple channel communication: a review of research and a proposed model. *Audio-Visual Communication Review* 9(6):235-262.

Publish in: N.A. Stanton and J. Edworthy (1999, eds.), Human Factors in Auditory Warnings (pp. 225-242). Aldershot, UK: Ashgate Publishing

- Marmolin, H. (1992) Multimedia from the perspective of psychology. In L. Kjelldahl (ed.) *Multimedia: systems, interaction, and applications*. (pp. 39-52), Berlin, Heidelberg: Springer.
- McCormick, E. (1957) *Human Engineering*. New York: McGraw-Hill.
- Momtahan, K., Hetu, R. & Tansley, B. (1993) Audibility and identification of auditory alarms in the operating room and intensive care unit. *Ergonomics* 36(10):1159-1176.
- Mountford, S. & Gaver, W. (1990) Talking and Listening to Computers. In B. Laurel and S. Mountford (eds.) *The Art of Human-Computer Interface Design*. (pp. 319-334), Reading, MA: Addison-Wesley.
- Munteanu, E., Guggiana, V., Darvishi, A., Schauer, H., Rauterberg, M. & Motavalli, M. (1995) Physical modelling of environmental sounds. In F. Pedrielli (ed.) *Proceedings of '2nd International Conference on Acoustic and Musical Research' – CIARM '95*. Ferrara: Universita di Ferrara, pp. 107-112.
- Rauterberg, M., Motavalli, M., Darvishi, A. & Schauer, H. (1994) Automatic sound generation for spherical objects hitting straight beams based on physical models. In T. Ottmann & I. Tomek (eds.) *Educational Multimedia and Hypermedia*. (Proceedings ED-MEDIA'94, pp. 468-473), Charlottesville: Association for the Advancement of Computing in Education.
- Riera, B., Vilain, B., Demeulenaere, L. & Millot, P. (1995) A proposal to define and to treat alarms in a supervision room. In *Preprints of the 6th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems at Massachusetts Institute of Technology, Cambridge, MA USA, June 27-29, 1995*, pp. 463-468.
- Stanton, N.A., Booth, R.T. & Stammers, R.B. (1992) Alarms in human supervisory control: a human factors perspective. *International Journal of Computer Integrated Manufacturing* 5(2):81-93.
- Wanner, J.C. (1987) Facteurs humains et sécurité. Séminaire: Erreurs humaines et automatisation. Ecole Nationale de l'Aviation Civile, May 19-20, 1987.