On the Relevance of Auditory Feedback for Quality of Control in a Balancing Task

Matthias Rath, Robert Schleicher

Deutsche Telekom Laboratories
Technical University Berlin, Germany
matthias.rath@telekom.de, robert.schleicher@telekom.de

Abstract

A tangible audio–visual interface based on the metaphor of balancing a ball on a tiltable track allows the measurement of human control movements under different conditions of sensory feedback. This specific scenario of human–system interaction forms an example for the definition of various measures of performance and quality of interaction. The dependence of these measures on specific configurations of the interface with regard to the employed audio–visual feedback, and their relationship is discussed.

1. Introduction

The notion of introducing a more intuitive or natural quality to human–machine interaction has recently received strong interest. The concept and field of *Embodied Interaction* [1][2] is maybe one of the most prominent in this respect but the same scopes appear to be driving forces behind developments of multimodal interfaces in general. The terms of "intuitiveness" and "naturalness" however are far from being defined in a fixed manner, but are conceived rather ad hoc in very specific contexts of interfaces and tasks. They are generally connected to a tendency of, or at least hope for, spontaneous understanding of interaction principles on the side of a user. An interface or display technique is thus rather ascribed an intuitive or natural quality when it requires less explicit explanation and dedicated training to be used. Another way to approach perceptual quality of an interface is in terms of adequacy and efficiency of conveyance of information that is necessary for the task to be committed. A scenario of human–machine interaction will in this sense be attributed perceptual quality if the optimal amount of information needed by the user to interact is conveyed in the most efficient way, and "overhead" or "noise" of useless information is minimised. In many situation of sensory feedback however, in particular when dealing with non–speech-based interfaces as in the present contribution, it is not a priori clear which information is the most relevant, and how it is best transmitted. To gain clarity in this point may therefore also form part of interface design. Finally, the possible hedonic quality of an interaction scenario, such as "pleasantness" or "joy of use", is yet another aspect of perceptual quality and the question emerges how the different mentioned aspects are connected.

The study presented in the following forms an exploration in the direction of the above points at an example scenario of balancing a rolling ball on a track. It makes use of an experimental tangible audio–visual interface based on this metaphor, the "*Ballancer*" [3]. The work has its background in the development of sound generation algorithms that aim at an enhanced use of the auditory channel in human–machine interfaces [4][5][6][7] and focuses on the relevance of auditory feedback for continuous gestural control. With the Gibsonian, *ecological* viewpoint on perception [8], the last decades have seen an increasing interest in auditory perception as a source of information about the physical environments that surround us and that we interact with [9][10]. Most studies and applications of mechanisms of ecological auditory perception (*everyday listening* [11]) however focus on short, discrete acoustic events, mostly of notification character, or anyhow on information conveyed in classes that do generally not change or are not perceived dynamically, such as the material [12] or size [13] of environmental objects. Our interaction with the physical world is often continuous and accompanied by seamless flows of information through as well continuous auditory feedback. (One may only think of common experiences such as driving a car where the noises of motor, brakes and air streams continuously inform about states of the system and influence the drivers behaviour.) Interest in the relevance and application of *continuous* auditory feedback in human–system interaction is therefore one distinctive central point behind the development of the *Ballancer* and the experiments presented here. Conducted user experiments demonstrate several possibilities of measuring aspects of perceptual quality in the user–device interaction, which are based on objective measurements of control gestures and performance times on one hand and on "self-observation" in the form of questionnaires on the other. They depend in different ways on the sensory conditions of the interaction, specifically the presence or absence of sonic feedback, and correlate among each other in various ways.

1.1. The present example scenario and interface

The *Ballancer* is a tangible audio–visual interface developed to examine questions of the context described above. It is based on the metaphor of balancing a ball rolling along a tilt-able track. The device is handled by the user as if balancing a small marble on top of a 1-m long wooden track whereby feedback about the movement of the virtual ball can be given through different types of visual and/or auditory feedback; figure 1.1 shows a photo. Gestural control input to the computation engine of the virtual scenario is accomplished by means of a sensor connected to the wooden control track that allows the measurement of the track's inclination angle. The present setup of the *Ballancer* makes use of an *InertiaCube3* [14] orientation sensor. On the basis of incoming sensor data, the movement of the virtual ball in dependence on the changing inclination of the control track is realised through an implementation of a discrete–time version of simplified physical equations describing the underlying scenario. Theoretical and practical details behind the *Ballancer* interface have been described in a dedicated article [3]. The re-

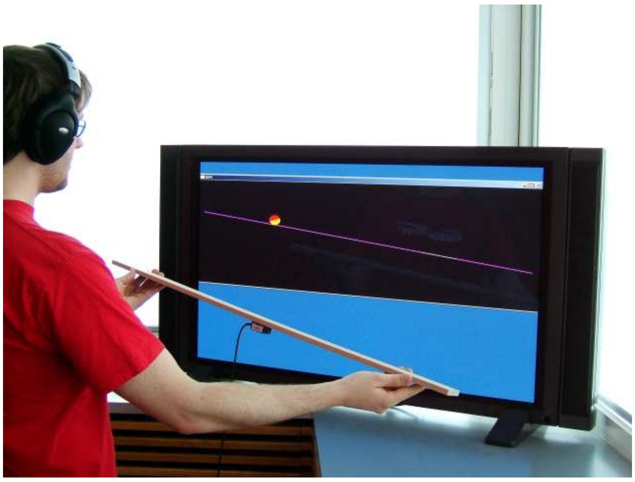

Figure 1: *The* Ballancer *in the configuration with a wide-screen display spanning the whole size of the 1*–m *physical control stick.*

active behaviour of the virtual ball (in connection with a *sound model* of rolling [7]), was, despite simplifications of the underlying physical description, perceived as convincing and the metaphor was understood spontaneously without any prior explanation by the subjects of a first user study [3].

The experimental setup used in the experiment described in this article includes a second, more abstract type of auditory feedback that will be described in the next section.

Visually the virtual scenario is presented in a simple graphical representation, in the present setup on a wide–screen display spanning the complete length of the (1-m–long) physical control stick (compare figure 1.1).

1.2. Previous work

The results of an initial user study at the *Ballancer* interface are basis and motivation for the work presented here. While detailed results can be found in dedicated publications [3][15], the main points are shortly summarised. 10 subjects participated in this experiment containing as its main part a task of moving the virtual ball into a fixed target area and stopping it herein. The target was characterised graphically on the (virtual) track by different colour and acoustically through a different, rougher, surface structure resulting in a different sound of the rolling ball. Subjects were asked to complete this task as fast as possible (in altogether about 90 trials) under the different feedback conditions of 4 graphical displays of different sizes (from the largest filling a 19" screen to the smallest spanning $1/12$ of this size) and with and without the auditory feedback from the *rolling model* [7]. It was found that

- when handling the interface blindfolded, receiving feedback only through sound from the *rolling model* test subjects generally understood the underlying metaphor without any explanation. While clearly having an artificial character the sound of the *rolling model* was nevertheless seen to be less ambiguous in representing a rolling object than the sound of a real glass marble rolling on the tilted track (in an identifications task with the 10 blindfolded subjects).
- the 10 subjects concluded the task of stopping the virtual ball inside the target area in average significantly faster when feedback from the *sound model* of rolling was present than with purely visual feedback. This effect of improved performance with sound was stronger for graphical displays of smaller size but present also for the full size 19" display.
- the noted faster task completion with additional sonic feedback could be connected to significant differences in indices of qualities of control movement such as the time after which the controlled virtual ball reaches its maximal velocity. In this sense the performance improvement with sonic feedback must be assumed to be (at least partly) due to increased control in acceleration and stopping with sound.

2. New *Ballancer* study: abstraction of auditory feedback and velocity information

Apart from its central interest in abstraction of auditory feedback that is discussed below, the work presented here differs from the first study (subsection 1.2) at the *Ballancer* in several smaller points. First, the focus of the present study does not directly include the specific potential of auditory feedback in the context of constrained conditions of visual display, such as limited screen size. Roughly speaking, interest is here more general in the question if sonic feedback in a typical situation of sensory–motor control conveys information that is in principle not or less efficiently perceived from visual feedback alone. As a consequence, in the setup of the *Ballancer* used in the following study subjects are supplied with visual feedback conditions that are optimal in the sense of the graphical display spanning the whole range of the control stick and subjects' arms during task performance (please compare figure 1.1) — as would be the case in the "real" physical scenario. For the same general reasons, the graphical display differs from the one of the previous study in that the ball is here not monochrome but with two differently coloured halves (figure 1.1) so that the turning movement itself may serve as an additional visual cue of ball movement. In the configuration of the *Ballancer* used in the previous study (subsection 1.2) the sonic feedback roughly reflected (and therefore potentially informed about) the position of the virtual ball through simple amplitude stereo panning and by the fact that the target area was marked by a different surface structure and as a consequence different acoustic behaviour. It was however shown [3][15] that the found performance improvement through sonic feedback was accompanied by optimisations of control movements already in the phase of the task before reaching the target area. Since positional information in amplitude panning is very coarse it must be assumed that the noted performance effects must be at least partly attributed to velocity information conveyed by the sonic feedback. The present work isolates this latter aspect and to this end employs conditions of sonic feedback that *exclude any possibility of positional information in the sound*. The momentary velocity of the virtual ball is used as the only input parameter. The sonic feedback does not depend on the controlled virtual ball's location in- or outside the target area and is entirely monophonic, i.e. does not contain any attributes of localisation.

2.1. Further abstraction of the sound of rolling

The *sound model* of rolling used in the initial configuration of the *Ballancer* is based on a simplified physical model of the interaction of the rolling object and the plane to roll on. The involved surface profiles that are "tracked" during the rolling movement form here the initial source of vibration that is processed as part of the *sound model* following considerations on the physical and geometrical principles of the scenario [7]. The development of the model followed the notion of *cartoonification* [11] (see also [3]), i.e. it aims at informativeness or "expressiveness" and clearness and simplification in its sonic appearance rather than realism. Success in these scopes was confirmed by the study reported in subsection 1.2 were the sound feedback from the model was, while being clearly recognisable as synthesised, immediately understood by subjects in the sense of the intended metaphor and lead to improved performance with the interface.

After these observations the notion of further "abstracting" the sonic feedback and possible implications on the conveyance of velocity information is central to the experiment reported in the following. A second, very simple and rather abstract sound model has been derived, by widely ignoring any idea of realism or even immediate similarity with the mechanical sound of a rolling object. This sonic feedback however still aims at expressing in a possibly intuitive way what is considered the main parameter of interest [1] for auditory display in our context, the one of velocity of a controlled movement. To this end the processing that accounts for the physical interaction in rolling is strongly reduced and the model is "stripped down" to tracing, at audio rate [2] the profile of a (hypothetic) surface, along the macroscopic track of contact that the rolling object follows. This strategy may be compared to an ideal needle as of a record player that follows what would be the "essential" trajectory of the movement (of the centre of the virtual ball). Finally (in contrast to the original rolling model that uses bandpass-filtered noise for modelling the surface), a highly artificial and rather unrealistic surface profile is now chosen, one of the shape of a lowpass-filtered sawtooth signal, to optimise the low-level psychoacoustic properties of the resulting signal:

- The Fourier spectrum of the sawtooth spreads over a very wide range of the frequency sensitivity of human hearing, up to its upper limit even for fundamental frequencies at the lower end of the hearing range.
- The sawtooth is periodic and thus stimulates a clear and strong sensation of pitch, e.g. in contrast to filtered noise.

The signal is slightly lowpass filtered in order to minimally smoothen the otherwise extremely harsh aesthetic appearance of the sawtooth. Summing up the consequences of the described derivation of the record–needle model, the fundamental frequency of the used sawtooth signal follows proportionally the ball velocity to be sonified. This condition of sonic feedback shall in the reminder for simplicity be denoted as "*abstract sound* (*as*)".

2.2. Objective performance measures for auditory feedback — experimental design

In order to examine the effects of the more abstract, very simple sonic feedback on users' handling and perception of the Ballancer device a second target reaching experiment, analogous to the one of the previous study (subsection 1.2) was conducted, containing both conditions of sonic feedback (again along with a condition of purely visual feedback). Thereby visual feedback is as described at the beginning of section 2, identical for the whole course of the experiment. 6 subjects participated in this experiment, all of them students at the Technical University of Berlin, aged between 21 and 27, four male and two female.

Even with this rather small number of participants (that is increased in current experiments) some statistically significant and interesting results were found. In the first, longer part of the study participants were again (as in the initial experiment of subsection 1.2) asked to move the virtual ball inside a graphically marked target area (compare figure 1.1) as fast as possible, this time under the three different conditions of *a)* sound feedback from the rolling model, "*rolling sound (rs)*", *b)* feedback from the more abstract record–needle model, "*abstract sound (as)*" and *c)* without sonic feedback, "*no sound (no)*". Thereby subjects were given no feedback whatsoever about the measured time they needed to conclude the task. The task was counted as fulfilled when the virtual ball completely stayed at rest inside the target area for at least $500ms$. The different conditions appeared in *sets* of 20 "*games*" each. The order of the sets/conditions was counterbalanced across subjects and the whole series of all conditions was repeated once for each subject so that the whole test consisted of 6 sets, e.g. for subject 1 of the form: "*rs, as, no, rs, as, no*", subject 2: "*rs, no, as, rs, no, as*"... Due to the repetition of the whole series, each condition appeared twice for each subject, as one set in a less "trained" state and again in "trained" circumstances in the second half of each test. Table 1 gives a quick summary of the distribution of feedback conditions for the 6 subjects in the design of the experiment. By counterbalancing the order of conditions we can hope that training effects during performance of the test "cancel out" in the comparison of different conditions of feedback averaged over all subjects.

sub-	Sound conditions of sets					
	"untrained"			"trained"		
ject	1	2	3	4	5	6
1	*rs*	*as*	*no*	*rs*	*as*	*no*
2	*rs*	*no*	*as*	*rs*	*no*	*as*
3	*as*	*rs*	*no*	*as*	*rs*	*no*
4	*as*	*no*	*rs*	*as*	*no*	*rs*
5	*no*	*as*	*rs*	*no*	*as*	*rs*
6	*no*	*rs*	*as*	*no*	*rs*	*as*

Table 1: *Counterbalanced order of feedback conditions (*"rolling"*, *"abstract"* and* no sound*) in the 6 sets of all 6 subjects.*

2.3. Subjective evaluation and quality of auditory feedback

While the measurement of performance parameters in first part of the experiment allows objective conclusions about information conveyance through the two types of auditory feedback, subjective assessments of quality were addressed by a questionnaire. These questions were presented to subjects after having concluded the performance part of the experiment. Altogether, each test (for one subject) took between 40 and 60 minutes and test subjects were paid 15 Euro for participating in the study. .

[1] The *sound model* of rolling in contrast to our specific focus here offers a potential to acoustically reflect a wider range of ecological attributes of the underlying scenario, such as the size, shape and weight of the rolling object, surface structure or material.

[2] i.e. with a temporal precision high enough to represent a continuous process to the human auditory system, exactly: at "cd audio rate" 44100 Hz...

3. Experimental results

3.1. Average *task times*

Again, as in the initial study, it is found that on average over all subjects the target reaching task is concluded faster with additional auditory feedback than without: table 2 shows a comparison of the mean performances under different conditions in the form of relative difference [3] and according statistical significances, i.e. p-values resulting from t-test comparison of the two respective sets of measurements. Hereby p-values of or close to a statistical significance of 5% are highlighted green. The shorter average "*task times*" with sound are equivalent to the negative $\%$-values only in the right columns of both ("untrained" and "trained") sub-charts of the table. It is thus a general result of the new study that this effect of faster performance with sound is present also for the very good visual feedback conditions of the wide–screen display. In further differentiation, the strength of the performance improvement through sound apparently interacts with the actual type of auditory feedback ("rolling" or "abstract") and the factor of evolving time in the course of the experiment. The effects of these two influences do not simply add up: It can be seen that in the untrained series task performance is significantly faster with *rolling sound* than without sonic feedback and still faster with *abstract sound* than without, but the latter improvement does not reach statistic significance. In the trained series *rolling sound* and *abstract sound* somewhat "switch roles"; still, with both types of sonic feedback average performance is better than without sound but now the *abstract sound* leads to the best average *task time* (that is significantly shorter than without sound). Figure 2 depicts this general behaviour of mean *task times* throughout training and conditions in a quickly perceived way, in a plot of the actual absolute respective mean values.

Relative differences of *task times*, statistic significance	"untrained"		"trained"	
	as	*no*	*as*	*no*
rs $\delta(\%)$ p	-9.1 0.1345	-17.9 0.0052	8.1 0.1414	-3 0.5977
as $\delta(\%)$ p		-9.7 0.1825		-10.2 0.0206

Table 2: *Differences in average* task times *(in %) under the different conditions in the untrained and trained set. Below each difference value the according statistical significance, p, is given.*

The development of the presented average *task times* can be summarised in simple words in the following way: The "abstract sound" leads to fastest average performance after some training, when the effect of the "rolling sound" becomes rather small. The latter however is more intuitive in the sense of fastest average task performance at the beginning of the test, when subjects are less familiar with the task. It is natural to ask to what extent these global tendencies are found for single subjects. The individual mean *task times* of single subjects are discussed further in subsection 3.3.

[3]This table is to be read as e.g.: mean *task time* under "rolling sound" "untrained" is 9.1% smaller than the respective value under "abstract sound" (upper left cell), mean *task time* under "abstract sound" "trained" is 10.2% smaller than the respective value without sound (lower right cell)...

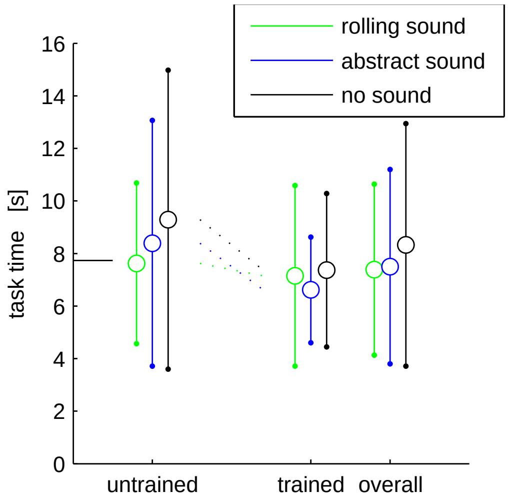

Figure 2: *Average* task times *(in* s*) across all subjects, under the different conditions of auditory feedback, over the "untrained" and "trained", and over all ("overall") games.*

3.2. Indices of quality of control movements

Because of the design of the present setup of the *Ballancer* (section 2) the phenomenon of faster mean *task times* with sound described in the previous subsection must here clearly be attributed to velocity information in the auditory feedback, since the position of the controlled ball is not reflected in the sonic feedback in any way. In order to examine in more detail the mechanism of the exploitation of auditory information in control movements, several characteristic indices in the recorded movement trajectories have been derived and analysed.

3.2.1. Ball oscillations

One characteristic of the virtual ball's movement in the target reaching task is the number of oscillations, more precisely changes of direction until being stopped inside the target. For a quick picture of these "*ball oscillations*" (and another movement index, the number *inclination swaps* introduced below) figure 3 shows the trajectories over time of virtual ball (blue) and track angle (red) at the example of one game. Table 3 shows the mean values of this index over all subjects and the "untrained" and "trained" series for the different feedback conditions, in absolute values (above) and in the format of relative differences (as in table 2). It is seen that all these average values are smaller with sonic feedback (negative values only in the right columns of both sub-charts of the lower table). Probably most striking is the value in the untrained series under the *rolling sound* condition as the number of oscillations is here even smaller than without sound *after training*. This fact reflects a spontaneous understanding of this sonic feedback by subjects.

3.2.2. Inclination swaps

Distinctive points in the control movement of the subject are the moments where the direction of acceleration of the virtual ball

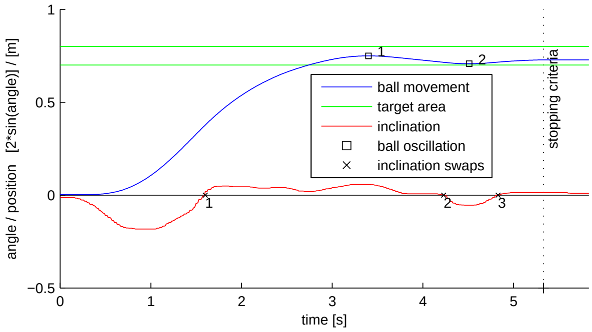

Figure 3: *Two characteristic indices in the temporal (x–axis in s) trajectories of the position of the virtual ball (blue, in* m*), and the underlying simultaneous development of the angle of the held control track (the red curve depicts the sine of the angle, stretched by a factor of* 2 *for better visibility) in one game. As an effect of a simple model of "stick" and "roll" friction the virtual ball stays at rest near the start and end of the game also in short phases where the control track is not being held absolutely horizontal (not even completely still, at the resolution of the direction sensor).*

Average *ball oscillations*	*rs*	*as*	*no*
"untrained"	2.76	3.86	4.18
"trained"	2.57	2.52	3.18
$\delta(\%)$	-6.9	-34.8	-23.8
p	0.622	0.032	0.066

Relative differences of *ball oscillations*, statistic significance	"untrained" *as*	"untrained" *no*	"trained" *as*	"trained" *no*
rs $\delta(\%)$	-28.5	-33.9	2	-19.4
rs p	0.0636	0.0019	0.909	0.2055
as $\delta(\%)$		-7.6		-20.9
as p		0.635		0.1691

Table 3: *Average number of changes of direction of the ball's movement.*

changes. These represent "decisions" [4] of the controlling subject to actively accelerate the ball further in the direction of its current movement or to actively stop, i.e. accelerate against the momentary direction. It can be seen from the values in table 4 that the mean number of such "*inclination swaps*" is at both conditions of acoustic feedback smaller than without sound (only negative %-values in the right columns of table 4, lower chart). As was the case with the *ball oscillations*, the situation for this index under *rolling sound* sticks out in that it is from the first series on smaller than without sound, even after training.

[4]The quotes are used to indicate that the question, if such an *inclination swap* is the result of conscious (as the word "decision" tendentially implies) or unconscious mechanisms, is left open.

Average *inclination swaps*	*rs*	*as*	*no*
"untrained"	3.88	4.66	5.35
"trained"	3.96	3.73	4.1
$\delta(\%)$	1.9	-19.9	-23.4
p	0.820	0.008	0.001

Relative differences of *inclination swaps*, statistic significance	"untrained" *as*	"untrained" *no*	"trained" *as*	"trained" *no*
rs $\delta(\%)$	-16.6	-27.4	6	-3.5
rs p	0.0404	0.0001	0.4441	0.6508
as $\delta(\%)$		-12.9		-8.9
as p		0.1069		0.1356

Table 4: *Average number of changes of the direction of inclination.*

3.3. Individual mean *task times*

Figure 4 gives an overview over the development of mean *task times* under the different conditions for all individual subjects. The behaviour of mean *task times* across all subjects as depicted in figure 2 is here generally not found in the values of single subjects. This is of course not surprising when considering that the mean *task times* are here influenced by the order in which the different feedback conditions were presented. A strong training effect, in particular in the beginning of the experiment, can be observed:

- The condition occurring first in the test always shows the longest mean *task time* over the "untrained" set, regardless of its quality. The fact of starting with one condition here clearly dominates over any effect of the sound used in this very condition. Only subject 2, who generally sticks out by his very small overall, and for two conditions even negative training behaviour, forms a slight exception.
- Again with the exception of subject 2, for each subject the condition presented first shows the strongest relative improvement of mean *task times* from the "untrained" to the "trained" set, regardless of its specific quality. (This is quickly seen from the steepness of the respective dotted lines in figure 4.)
- The difference of mean "untrained" and "trained" *task times* of one condition is generally larger than the differences between sound conditions within the "trained" series.

While the general point of the previous discussion may be read as the overall temporal training effect dominating the individual plots of figure 4 and these allowing small insight into the relevance of the different sound conditions on task performance, one aspect is noteworthy: the abstract sound condition is the only one where an improvement of mean performance from the "untrained" to the "trained" set is found *for all single subjects*. This is easily seen from the fact that all blue (and only these) dotted lines in figure 4 decrease to the right. While it was noted in the previous section 3.1 that in the "trained" series mean performance over all subjects is significantly best with abstract

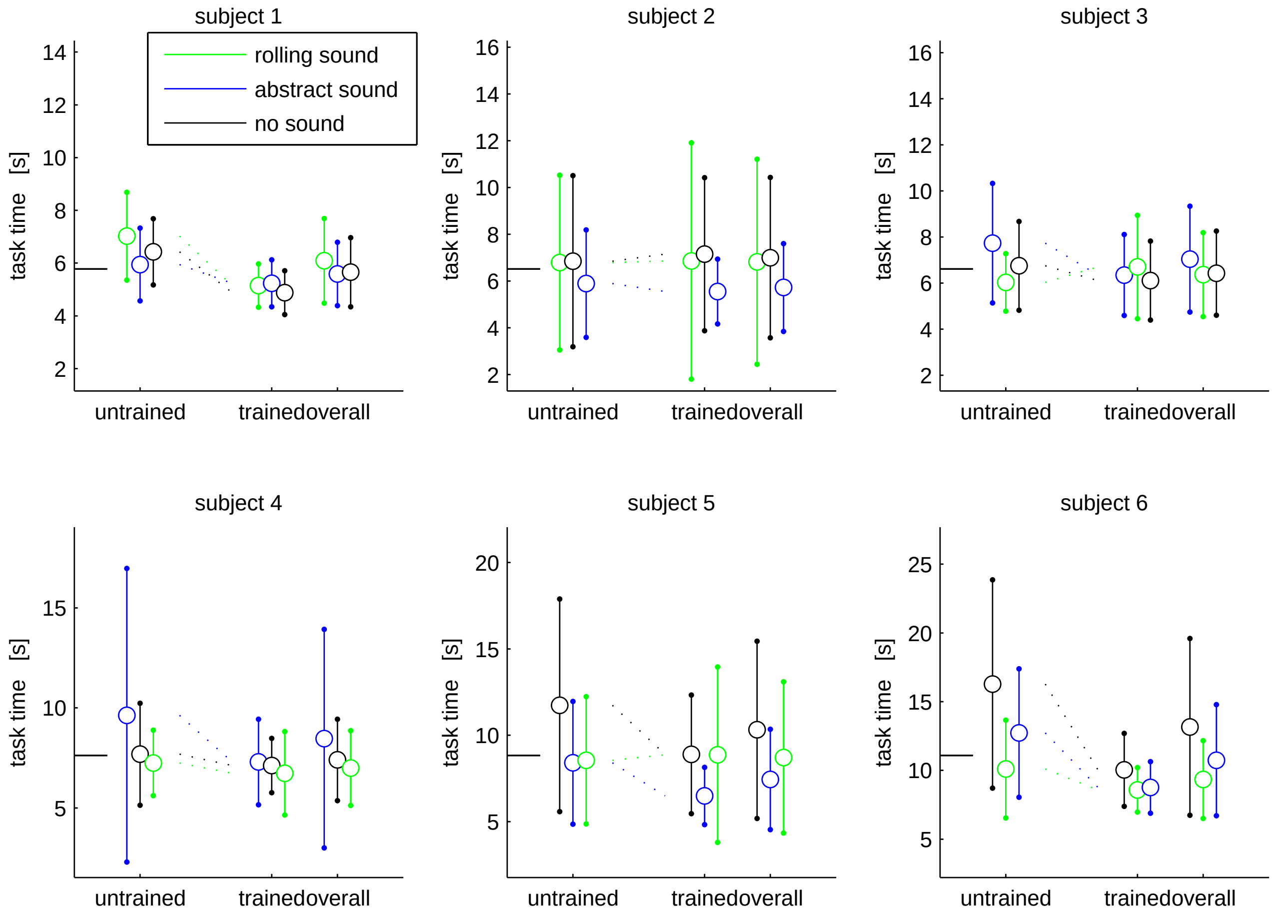

Figure 4: *Individual mean* task times *(in* s*) under the three different conditions of sonic feedback, over the "untrained" and "trained", and all ("overall") games, represented by coloured circles with standard deviations. The dotted connecting lines allow for an easy comparison of these "untrained" and "trained" values. A black line on the y–axis marks the overall mean performance (over all games, i.e. across all conditions) of each subject. For better comparison and visibility all vertical axes are scaled relative to this individual mean performance (which, as a consequence, appears at fixed height in all graphs).*

sonic feedback, the observation just remarked is a hint that this effect may be based on perceptual mechanisms that are shared by all participating subjects. The hypothesis that the positive effect of the abstracted acoustic feedback is limited to only some particular subjects who "understand" this feedback while others do not at all learn to make use of the contained information, appears improbable after this qualitative analysis of individual mean results.

3.4. Subjective assessment and self-evaluation

After having performed the target reaching experiment whose measured results where described in the previous subsection 3.1 subjects where asked to fill a questionnaire referring to their personal impressions during task performance. The first two of the questions aim at possible effects of the sound conditions on task difficulty:

1. Did you feel that under some conditions of acoustic feedback ('rolling sound', 'abstract sound', 'without sound' [5]) it was easier or more difficult to perform the presented task than under others?
   Please fill the gaps in the following comparison chart accordingly with the attributes 'easier than', 'more difficult than' and 'equally difficult as'.
   a) Performance with 'rolling sound' was . . . performance without sound.
   b) Performance with 'abstract sound' was . . . performance without sound.
   c) Performance with 'rolling sound' was . . . performance with 'abstract sound'.
2. How do you estimate the time you needed to perform the task under the different conditions of acoustic feedback? Please fill the gaps in the following comparison chart accordingly with the attributes 'slower', 'faster' and 'equally fast'.

[5] These terms where verbally explained to subjects to avoid any misunderstandings.

a) In average I performed the task ... with 'rolling sound' than/as without sound.
b) In average I performed the task ... with 'abstract sound' than/as without sound.
c) In average I performed the task ... with 'rolling sound' than/as with 'abstract sound'.

While question 1 aims at the subjective feeling of difficulty under the different acoustic conditions, question 2 asks for an estimation of the own performance which may possibly not exactly correlate, in particular because of an effect of training that may lead to faster task performance for conditions occurring later in the order of sets. Question 3 addressed exactly this last point:

3. In case you noted differences in your performance under 2. a), b) or c), please mark if you think that one of the following factors applies:
○ I was faster in one case because of the supportive effect of the respective sound feedback.
+ I was faster under the condition occurring later in the test because I was then better trained.
− I was faster under the condition occurring earlier in the test because later I was more tired.

Finally subjects were asked to

4. describe in free words the felt effects of the different conditions of acoustic feedback ('rolling sound', 'abstract sound', 'without sound') when performing the task respectively controlling the virtual ball?

Table 5 shows subjects' answers to questions 1 and 2 as represented by the attributes from comparison charts described above. In most cases, corresponding answers to questions 1 and 2 match in the expected way, "easier" ↔ "faster", "equal" ↔ "equal", "harder" ↔ "slower".. I.e. where the task was perceived as *easier* under one condition than under another the own performance under this condition was generally estimated faster than under the other; exceptions, 5 out of 18 comparisons, are highlighted by colour. 3 of these mismatching cases are the judgements of subject 1 who estimated his own performance as " 'rolling' slower than 'abstract' slower than without sound" while stating to feel no influence of acoustic condition on task difficulty. These comments may appear strange at first sight, but it is noted that the order of conditions for this subject was precisely "rolling", "abstract", "no sound" (compare table 1), so that this self-estimation would fit an expectable effect of training. In question 3 however subject 1 did not explicitly state any guess of an influence of training. As a matter of fact, subjects' remarks under question 3 in general are rather sparse and not revealing; this rough observation will be picked up again below. The remaining two cases of not matching responses to questions 1 and 2, subject 3 ("abstract" vs. "no sound") and subject 6 ("rolling" vs. "abstract sound"), do not coincide with any expected effects of training and are thus distinguished in table 5 in red rather than green as for subject 1.

	Answers to questions 1 and 2 for subject no.					
	1	2	3	4	5	6
rs vs. *no*	equal slower	easier faster	easier faster	easier faster	easier faster	easier faster
as vs. *no*	equal slower	equal equal	harder equal	equal equal	easier faster	easier faster
rs vs. *as*	equal slower	easier faster	easier faster	easier faster	harder slower	easier equal

Table 5: *Answers of the 6 subjects to the questions of the felt influence on difficulty (above in each cell) of the different conditions (*"rolling"*, *"abstract"* and* no sound*) and the estimated respective own average performance (below in each cell).*

For a comparison of individual objective measured *task times* and the subjective assessments of acoustic conditions and guesses of self-estimation (at least) three different average values may be looked at, namely the average *task times* in the *untrained* and *trained* sets and over the whole test. These individual values are represented in figure 4. Since several comparisons of individual subjective and objective results are possible these are not discussed in detail here but only in their overall tendency. Only one case of strongly contradicting subjectively estimated and objectively measured *task times* is found and shall be noted: Subject 2 rated the abstract auditory feedback as having no influence on task difficulty and estimated her own average performance under this condition as equally good as without sound but consistently reached clearly shorter average *task times* with "abstract sound" (see table 4). Apart from this one case, no extreme discrepancies of subjects' individual self-estimations and individual objective measurements are found. In their sum however, these subjective guesses would give a rather wrong picture of the actual measured average effect of the auditory feedback. As seen in table 5, 5 out of 6 subjects labelled the *rolling sound* as helpful for performance of the task and accordingly estimated their own performance as faster under this condition while this effect is statistically significant only in the first "untrained" series of sets (table 2). On the other hand only 2 of the 6 subjects ascribed the same positive effect to the abstract auditory feedback, one subject even thought that this sound would make task performance more difficult (table 5, line 2), while the "abstract sound" is seen to lead to the best average performance after some training (table 2). Similarly, only 1 subject thought that the "abstract sound" would be more useful in performing the task than the "rolling sound", while 4 out of 6 were convinced of the opposite. In connection with the individual comparisons above it appears that subjects have great difficulties in distinguishing the separate and possibly counteracting effects of training and the different feedback conditions on task performance. This assumption may be supported by the general tendency of subjects to avoid references to training effects in question 3 which explicitly aimed at this point. More precisely, only 2 such references to the temporal order (within the 18 comparisons) were made.

When finally asked to give an overall impression in their own words as the last point of the questionnaire, only one subject preferred the abstract sound, whereas all others rejected it as 'inappropriate', 'disturbing', 'annoying' or 'not fitting'. The question of possible reasons for this subjective rejection was not examined any further in the present study. Besides the abstract feedback clearly not reminding a rolling ball at all, it may be guessed that the very static, inorganic or artificial character of this sound plays a roll. Nevertheless, three of these five stated the task to be less attractive with no sound at all ("... without any sound the whole thing was rather awkward", "Without sound was boring, you really had to concentrate", "... without sound it was rather strenuous in the long run"). The most common complaint about the "abstract sound" was that it did not fit the task or represent the ball movement appropriately. On the other hand, the abstract sound led more reliably to an improvement from the untrained to the trained condition than any other feedback condition which suggests that subjects very well utilised the information conveyed by this feedback.

In the whole the supportive effect of the abstract auditory feedback used in the *Ballancer* setup here would remain basically undetected when relying on subjects' conscious answers. It may be a strength of the approach of assessing mechanisms of information conveyance through a measurement of gestural reactions that it can uncover perceptual phenomena which are not detected by conventional techniques of evaluation such as questionnaires. While this approach is rather rarely used in psychoacoustic research, the apparent dissociation between subjective assessment and objective behavioural measures is indeed well established in psychology [16] and has led to a re-appreciation of unconscious components in learning and decision processes. (For a recent review see [17].) Applied to auditory feedback in human-machine-interaction one clearly has to distinguish different aspects of quality: increases in subjective hedonic quality ('joy-of-use') and objective performance-oriented improvements do not necessarily co-occur. Whereas the first aims at enhancing a realistic experience by mimicking naturalistic sounds, the second benefits from increased information content in the respective signal.

4. Conclusions

The study presented here has dealt with possible aspects of perceptual quality of objective as well as subjective nature and their relation in the context of a task of continuous gestural control under different conditions of audio–visual feedback. It has been shown that the human–system interaction is improved through auditory feedback, in that a target reaching task is on average concluded faster by 6 test subjects when sound feedback was present. This effect (that had been found for smaller graphical displays and one type of auditory feedback in a previous study) was shown also under "optimal" conditions of visual feedback, and for a type of auditory feedback of rather abstract nature. In accordance with faster task performance with auditory feedback, two example indices, the numbers of *ball oscillations* and of *inclination swaps*, were seen to be smaller with sound in average — clear indications for an improvement of control movements. As a consequence of the distinctive design of the auditory feedback, it must be assumed that the noted effects must be attributed to an improvement in the conveyance of velocity information. Both types of auditory feedback thus improve the perceptual quality of the interface as far as concerns objective performance measurements. The two types of feedback however strongly differ in their scaling of the effect of performance improvement with training and in participants' subjective assessment. The closer ecologically oriented *sound model* is understood faster as it lead to the strongest and significant improvements in average performance in the first half of the test. In contrast, after some training (and only then) the more abstract auditory feedback leads to better performance. Subjective estimation of the influence of the types of auditory feedback on task difficulty and the own performance would predict a stronger effect for the "rolling sound" and not allow clear conclusions about the effects of the more abstract auditory feedback, an example for the value of assessment by measurement of gestural responses. From the aspect of hedonic quality the more ecological auditory feedback was clearly preferred by the test subjects, with the more abstract one generally not being appreciated.

As a general summary, differently defined measures of quality may strongly diverge, and in particular an integration of subjective and objective qualities can be a challenging task.

5. References

[1] Kenneth P. Fishkin, Thomas P. Moran, and Beverly L. Harrison, "Embodied user interfaces: Towards invisible user interfaces," in *Proceedings of the Seventh Working Conference on Engineering for Human-Computer Interaction, EHCI 98*, (Heraklion, Crete, September 13-18 1999, pp. 1–18, Kluwer, B.V.

[2] Kenneth P. Fishkin, Anuj Gujar, Beverly L. Harrison, Thomas P. Moran, and Roy Want, "Embodied user interfaces for really direct manipulation," *Communications of the ACM*, vol. 43, no. 9, pp. 75–80, 2000.

[3] M. Rath and D. Rocchesso, "Informative sonic feedback for continuous human–machine interaction — controlling a sound model of a rolling ball," *IEEE Multimedia Special on Interactive Sonification*, pp. 60–69, April 2005.

[4] Davide Rocchesso and Federico Fontana, Eds., *The Sounding Object*, Mondo Estremo, Firenze, Italy, 2003, < http://www.soundobject.org >.

[5] Matthias Rath and Federico Fontana, "High-level models: bouncing, breaking, rolling, crumpling, pouring," in *The Sounding Object*, Davide Rocchesso and Federico Fontana, Eds., pp. 173–204. Mondo Estremo, Firenze, Italy, 2003.

[6] M. Rath, *Interactive Realtime Sound Models for Human–Computer Interaction — A Sound Design Concept and Applications*, Ph.D. thesis, Department of Computer Science, University of Verona, 2004, available online under <http://www.di.univr.it/dol/main?ent=pubbdip>, "research reports".

[7] M. Rath, "An expressive real-time sound model of rolling.," in *Proceedings of the 6th "International Conference on Digital Audio Effects" (DAFx-03)*, London, United Kingdom, September 2003.

[8] J. J. Gibson, *The senses considered as perceptual systems*, Houghton Mifflin, Boston, USA, 1966.

[9] W. W. Gaver, "What in the world do we hear? an ecological approach to auditory event perception," *Ecological Psychology*, vol. 5, no. 1, pp. 1–29, 1993.

[10] W. W. Gaver, "How Do We Hear in the World? Explorations in Ecological Acoustics," *Ecological Psychology*, vol. 5, no. 4, pp. 285–313, Apr. 1993.

[11] W. W. Gaver, *Everyday listening and auditory icons*, Ph.D. thesis, University of California, San Diego, 1988.

[12] R. Wildes and W. Richards, "Recovering material properties from sound.," in *Natural Computation*, W. Richards, Ed., pp. 356–363. MIT Press, Cambridge, MA, 1988.

[13] M. M. J. Houben, A. Kohlrausch, and D. J. Hermes, "Auditory cues determining the perception of size and speed of rolling balls," in *ICAD01*, Espoo, Finland, 2001, pp. 105–110.

[14] "*InterSense*," < http://www.isense.com/ >.

[15] M.Rath, "Gestural exploitation of ecological information in continuous sonic feedback — the case of balancing

a rolling ball," in *Proceedings of the 8th International Conference on Digital Audio Effects (DAFx'05)*, Madrid, Spain, 2005, pp. 98 – 103.

[16] A. Bechara, H. Damasio, D. Tranel, and A. R. Damasio, "Deciding advantageously before knowing the advantegeous strategy.," *Science*, vol. 275, pp. 1293–1294, 1997.

[17] A. Bechara and A. R. Damasio, "The somatic marker hypothesis: a neural theory of economic decision.," *Games and Economic Behaviour*, vol. 52, no. 2, pp. 336–372, 2005.