Applied Ergonomics 116 (2024) 104204

Contents lists available at ScienceDirect

Applied Ergonomics

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

Evaluation of different feedback designs for target guidance in human controlled robotic cranes: A comparison between high and low performance groups

Felix A. Dreger *, Gerhard Rinkenauer

Leibniz Research Centre for Working Environment and Human Factors, IfADo, Dortmund, Germany

ARTICLE INFO

Keywords:
Machine operator
Concurrent visual feedback
Sonification

ABSTRACT

Labour shortages and costly operator training are driving the need for digital on-board robotic crane operator support in forestry and construction. This simulator study investigated the effects of sonification (auditory, pitch/loudness) and continuous visual (brightness/size) feedback on aiming movements with a robotic crane for low and high performers. The feedback was designed non-linear and linear. Thirty-six participants controlled a robotic crane bimanually using joysticks across 320 movements. Performance and skill indicators (movement time, accuracy, trajectory, smoothness) as well as satisfaction, and usefulness were assessed. Low-performing participants showed higher movement accuracy, particularly with non-linear pitch feedback compared to visual feedback. High performers exhibited no significant performance improvement in movement time, accuracy, or smoothness. There was no effect of linear or non-linear mapping of the feedback. Additionally, perceived satisfaction was lower with auditory than visual feedback. These results suggest that real-time auditory feedback can enhance operator accuracy whereas acceptance remains challenging.

1. Introduction

Costly machine operator training and labour shortages are driving the forestry and construction industries to improve machine operator performance through digital operator support. The goal to raise productivity requires the development of effective support systems that can be used on-board of machines during operation. Crucial to operator performance is the ability to bimanually control a robotic crane, which determines the productivity and safety of operations in the forestry and construction industries (Purfürst, 2010). Even experienced operators show large productivity differences (Ovaskainen et al., 2004, 2011), much of which can be attributed to poor robotic crane control (Hartsch et al., 2022). Reduced training needs could already be achieved by introducing technical support via direct crane tip control (also referred to as end-effector control) (Manner et al., 2017). End-effector control allows the operator to directly control the movement of the tip of the robotic crane, rather than controlling each joint separately. However, joint-based control systems are still on the market and will continue to be for some time to come. Regardless of the type of control system, simulator studies on tele-operation show that the achievable movement accuracy is a challenge with both types of control systems (Dreger et al., 2023). Therefore, the aim of this study is to investigate the extent to which the accuracy of robotic crane control can be improved through enhanced sensory operator feedback.

1.1. Forms of feedback

Since bimanual robotic crane control is largely a motor control task, feedback must be suitable for improving human movement capabilities. There are two types of information that humans use as feedback to refine motor movement. First, information from inherent feedback that is accessible to the performer by perceiving the environment and proprioception. Second, information from augmented feedback that is not generally available to the performer or is difficult to access (Schmidt et al., 2019). The augmented feedback can be in the form of knowledge of results (KR), which is the information about the movement outcome after the movement, or in the form of knowledge of performance (KP), which is information about the movement characteristics. Both types of information can be of help to the performer to improve motor movement (Keogh and Hume, 2012; Magill and Anderson, 2017; Oppici et al., 2021; Schmidt et al., 2019; Sharma et al., 2016; Zhu et al., 2020).

* Corresponding author.
E-mail address: dreger@ifado.de (F.A. Dreger).

https://doi.org/10.1016/j.apergo.2023.104204
Received 22 August 2023; Received in revised form 22 November 2023; Accepted 7 December 2023
Available online 20 December 2023

1.1.1. Concurrent movement feedback

Concurrent feedback combines KP (feedback on movement execution) and KR (feedback about movement result) and is referred to as feedback that is provided continuously throughout the supported movement. Feedback can be given by experienced operators and trainers or via digital support. Concurrent feedback can be provided as information on the whole movement or a specific part of the movement to be improved. Concurrent movement feedback was found to be especially helpful, to support learning of complex movements (Wulf et al., 1999). Particularly, in early learning stages, movement execution with concurrent feedback was improved (Wulf et al., 1998). Even in virtual environments, concurrent feedback could outperform real-task training and coaching for example in table tennis (Todorov et al., 1997). The feedback of KP and KR is more effective when it leads to an external rather than an internal focus of attention (Shea and Wulf, 1999). In robotic crane control, the position of the end-effector attracts the external focus of the operator's attention in movement control (Häggström et al., 2015), therefore, providing information on the end-effector appears useful. In addition, the usefulness of concurrent feedback is closely related to the complexity of the task and the skill level. Complex tasks and low skill levels can benefit more from concurrent feedback than high skill levels and less complex tasks (Wulf et al., 1998). Auditory feedback has been shown to be particularly suitable for complex 3D movements such as rowing (see Effenberg et al., 2011) and was effective in early stages of skill acquisition. Conversely, proficient performers can be distracted from concurrent feedback (Wulf et al., 1998). Robotic crane control requires learning complex bimanual hand movements, especially the transformation from joystick input to robotic crane movement needs to be mastered. Therefore, concurrent feedback may be of help to improve the performance of robotic crane control.

1.1.2. Visual movement feedback

Concurrent feedback was often provided visually via displays and improved rowing movements using symbolic oar and blade representations (Sigrist et al., 2011, 2013). In robotic crane control, visual concurrent feedback has yet not been investigated. It is worth noting that in robotic crane control, visual feedback must be carefully designed. This is because the control task is already predominantly visual. Providing feedback through the same modality can therefore overload the visual system since the same cognitive resources are claimed (Wickens, 2002). One way to reduce the visual load of presentation is to use augmented reality (AR) displays, which overlay the real world and thus allow feedback to be presented in the visual field in such a way that it is only slightly distracting. This is also referred to as contact analogue presentation, which showed benefits over conventional information presentation in, e.g., controlling robotic cranes or driving cars (Ding et al., 2022; Eriksson et al., 2019).

1.1.3. Auditory movement feedback

Typically, auditory feedback for motor behaviour was used in the form of alarms that alerted to indicate deviations from the correct movements such as in dance and gymnastics (Baudry et al., 2006; Clarkson et al., 1986). Auditory feedback on the hand position in swimming has been shown to improve crawling technique (Chollet et al., 1988). More complex concurrent auditory feedback effectively improved dancing, rowing, cycling performance and also piloting performance (Effenberg et al., 2016; O'Brien et al., 2020; Sigrist et al., 2016; Sigrist et al., 2011; Valéry et al., 2017). The feedback was used to provide information on a single movement element but also on the whole movement (Effenberg et al., 2016). A special type of concurrent auditory feedback is sonification (see Walker and Nees, 2011). Sonification as concurrent feedback sets sound for movement variables such as spatial position or dynamics of certain limbs during movement, e.g. paddle forces in cycling, or of variables external to the human body such as oar position in rowing or roll and pitch in flying and distance to hazards in driving (Effenberg, 2005; Song et al., 2022; Valéry et al., 2017; Vidal et al., 2020). To use the information provided by sonification, a certain amount of movement representation/training must already be present (Sigrist et al., 2013). Sonification uses sound properties such as timbre, stereo balance, volume, and pitch to provide concurrent feedback. These auditory properties can be mapped to different movement variables at a time to create a combined sound pattern. In particular, stereo balance and pitch feedback were useful to support optimal rowing performance (Sigrist et al., 2011).

In addition to movement dynamics, movement errors can also be signalled in sonification feedback. Multidimensional error feedback has been applied in cycling and rowing training (Effenberg, 2005; Sigrist et al., 2016). In both, sonification led to higher and more accurate performance. Within the above studies, feedback on the movement error showed fast adaptations of the rower.

In conclusion concurrent auditory feedback in form of sonification and concurrent visual feedback can benefit the execution and learning of motor tasks (Sigrist et al., 2013). Both visual and auditory feedback are helpful, specifically for low performers and when more complex skills are required, which according to Wulf & Shea (2002) involve the control of multiple degrees of freedom and have an ecological application. As bimanual control of robotic cranes is considered complex, concurrent feedback is assumed to be useful to support operators. Nonetheless, studies on sonification as feedback for movement dynamics or task errors are rare in sports science and even rarer in human factors.

1.2. Feedback in robotic crane control

Auditory feedback has already been applied in robotic crane control. A study with sounds of different discretised frequencies showed shorter movement times in the feedback condition (Mavridis et al., 2015). The movements were conducted with a 2D stick figure manipulator visualised on a notebook controlled with two joysticks.

Visual concurrent feedback in augmented reality (AR) was used to improve the performance of a hydraulic robotic crane. AR aided the operator in overcoming control shortcomings introduced by the asymmetric workspace mapping of master (joystick deflection) and slave (crane movement) of a hydraulic manipulator (Ding et al., 2022). Improvements were found in terms of task completion time, where kettle bells placed on barrels were lifted and moved from one barrel to another.

To make use of auditory and visual feedback in training and support systems, the design of feedback during the movement, the timing of feedback, and auditory and visual properties such as pitch and shape of feedback remain to be investigated. Both auditory feedback and visual AR feedback appear suitable to support robotic crane operators. Therefore, it seems worthwhile further investigating the effectiveness of designing different auditory and visual feedback when operating robotic cranes.

1.3. Aim of the study

The aim of this study was to analyse the effectiveness of sonification (auditory) or continuous visual feedback to enhance robotic crane movement precision and to evaluate the cognitive load, acceptance, and usefulness of the feedback. Additionally, the effects of feedback on movement time and skill of robotic crane movements should be assessed. Feedback for both modalities was designed to follow a non-linear or linear dependence on target distance. The non-linear feedback of pitch (higher), volume (louder), brightness (less bright), and size (larger), increased the resolution of the reported movement changes of the target distance the closer the robotic crane tip got to the target. Previous research suggests that auditory feedback is superior to visual feedback (Mavridis et al., 2015; Wickens, 2002), and non-linear feedback is superior to linear feedback. Both are hypothesised to be more effective for low than for high performers in robotic crane control (Wulf et al., 1998).

2. Method

2.1. Participants

Thirty-six participants (male = 20, female = 16) novice to the task of controlling a robotic arm, right-handed (self-reported), with normal or corrected to normal vision (stereoscopic vision test, Walraven, 1972) and self-reported normal hearing ability provided written consent to participate in the study. Participants were recruited students and employees from the Technical University Dortmund and were between 18 and 35 years old ($M = 24.19$ years; $SD = 4.31$ years). The Participants were split in two groups of low ($n = 18$, female = 9, male = 9, $M = 22.78$ years; $SD = 4.10$ years) and high ($n = 18$, female = 7, male = 11, $M = 25.61$ years; $SD = 4.16$ years) performers for the comparison of feedback effectiveness. The study was approved by the ethics committee of the Leibniz Research Centre for Working Environment and Human Factors.

2.2. Simulator

The simulator was a fixed-base robotic crane simulator consisting of a Chicago truck seat and two joysticks. Two Xiao Mii 55-inch TV screens combined with a semi-permeable mirror and two speakers provided the visual and auditory environment. The simulator setup is shown in Fig. 1. The screens were placed at 90° angle to each other, with one screen in front of the participant and the other on the left side. In between the front screen and the participant was the semi-permeable mirror placed at 45° to the line of sight of the participant. The participant could see information on the front screen and information from the left screen mirrored in the semi-permeable mirror (see Fig. 1a). The robotic crane had 4 degrees of freedom (DoF) and was velocity controlled via the joints that mapped to the joysticks. The exact mapping is shown in Fig. 2. The simulator was running on Ubuntu version 20.04. The visualisation of the robotic crane ("Robotis Open Manipulator"; ROBOTIS Inc., Korea) was rendered using GAZEBO. The experimental control was implemented in C++ and Python 3. The visual feedback was rendered using RViz and concurrent auditory feedback was created with PureData. All software communicated via ROS noetic.

The simulation showed the robotic crane from a bird's-eye view on a gridded floor and a white background (cf. Fig. 3). Movement targets were pairs of flat circles laid out on the ground with a diameter of $d = 0.1$ m. The circles determined movement start and end. Four targets indicated the start of movement on the left side by blue colour. The same targets, mirrored, indicated the start of movement on the right side in purple colour (cf. Fig. 3). Eight different pairs of circles were presented and thus eight different movements were performed using the robotic crane. For each target pair a series of five alternating taps was executed.

Four different distances and orientations of the circular target pairs were presented. Movements were executed diagonally to ensure the use of all joints with the goal to tap the centre of the target circle with the tip of the closed gripper (cf. Dreger et al., 2023).

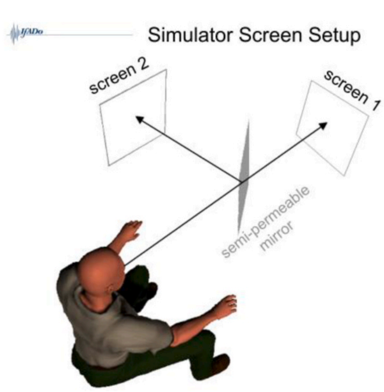
2.3. Concurrent movement feedback design

Continuous visual or sonification feedback on the distance to the target was given concurrently to the movement for the last 66% of the total distance between the two target circles. This part of an aiming movement is deemed the landing or homing in phase after the peak velocity of the movement (Meyer et al., 1988). Thus, the feedback onset was during the movement. All resulting mappings of visual and auditory feedback are shown in Fig. 4. The starting point of all mappings was a non-linear function to which the corresponding linear mapping was derived. The continuous distance feedback (visual or auditory) during the movement provided KP while moving and KR in form of the achieved distance to the target centre at the end of the movement. The non-linear mapping increased the resolution of reported movement changes and ought to allow the participants to align the crane tip precisely with the target centre.

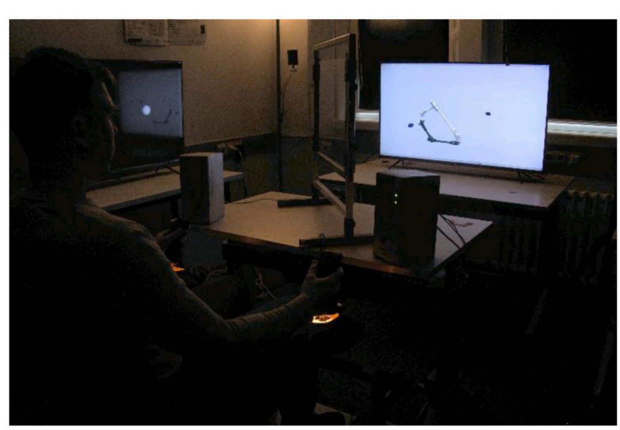
2.3.1. Auditory feedback

Continuous auditory feedback was provided in the form of the sonified movement of the crane tip distance to the target, by either modulating the loudness or the pitch. For this, loudness and pitch were mapped onto the distance of 66% of the total distance between each target pair.

Pitch was modulated by frequency. The predominant frequency was chosen such that the sound was comfortable for the participants by using three overtones and at the same time had a mostly flat isophone loudness contour for the given frequency range. The frequency range was 180 Hz–246.67 Hz, which was mapped onto the distance to the target. This range is comparable to a lower female voice. The mapping resulted in higher tones the closer the crane tip got to the target centre. Thus, the increase of pitch and current pitch (feedback) was indicative for performance. Loudness was manipulated by sound pressure level expressed in dB. For this, the range from 46.67 to 60 dB, was mapped on the distance to the target. This range is comparable to the loudness of whispering to a normal conversation. The loudness increased with closing in the target centre with loudness increase and current loudness (feedback) being performance indicators. Both loudness and pitch were mapped



(a)



(b)

Fig. 1. Schematic drawing (a) of the simulator setup with screen 2 displaying the feedback and the semi permeable mirror bringing the feedback in the field of view of the participants directed to the main front screen 1. The used simulator setup (b).

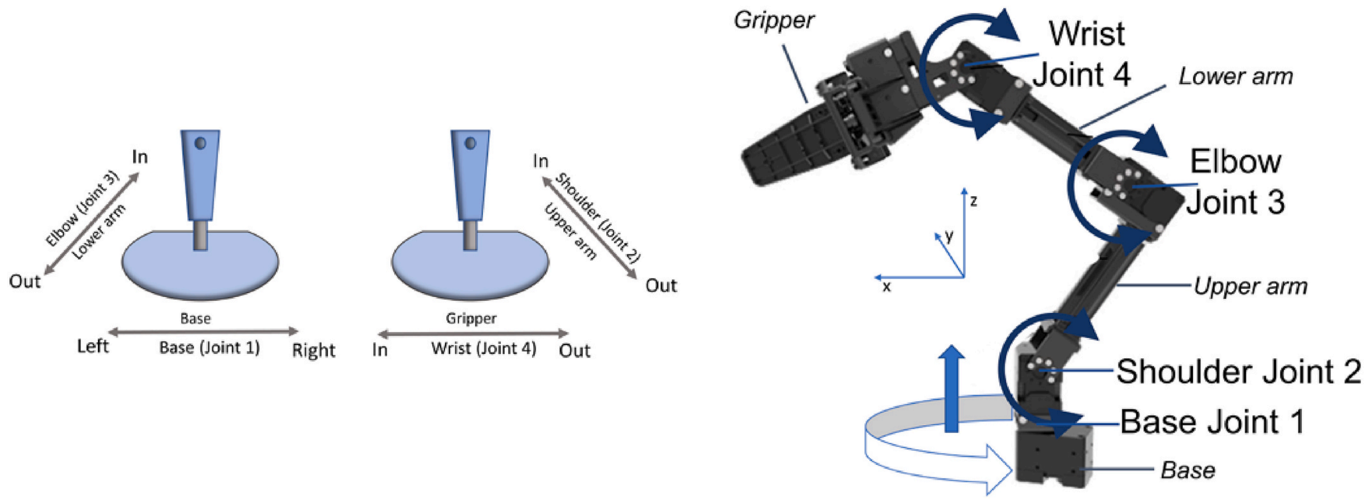
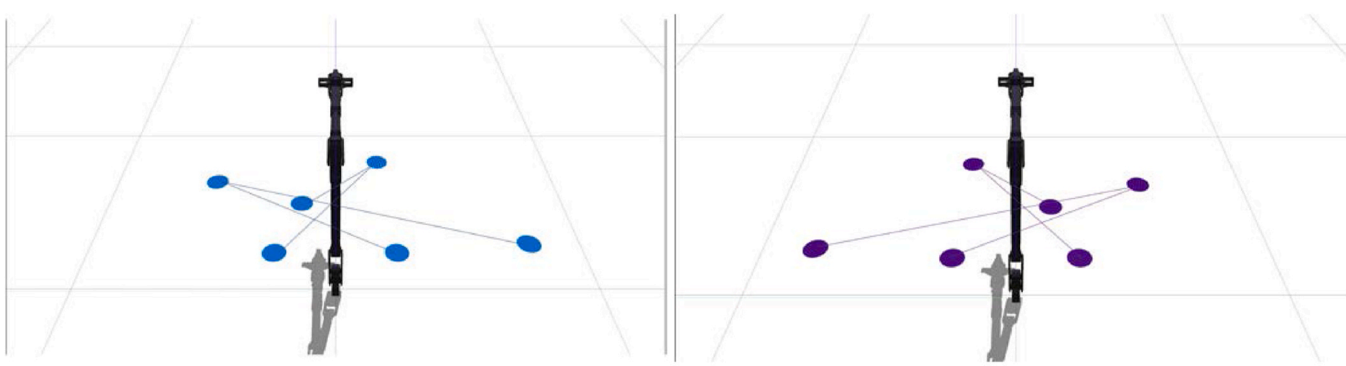


Fig. 2. Joysticks with control mapping (left). The arrows indicate the joystick movement direction and the effect on the controlled joint. Robotic crane with movement directions and labels (right) (Figure from Dreger et al., 2023a,b).



(a) Start left (b) Start right

Fig. 3. Shown are four target pairs that indicated the movement start left (a, blue) and indicating the movement start on the right (b, purple). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

linearly and non-linearly. The non-linearity was aimed to be in line with human perception (cf. sound intensity).

2.3.1.1. Loudness. For the loudness feedback, the distance information of the end-effector to the target was simulated as the approach to a spherical sound source. The sound intensity was set to be 60 dB at the cranes end-effector distance to the target $r = 0$. The change in sound intensity as a function of distance was simulated according to the inverse square law $(r) \propto \frac{1}{r^2}$. The sound pressure level was calculated as the logarithmized ratio of the sound intensity with a reference sound source I_0. The reference sound intensity is usually assumed to be $I_0 = 10^{-12}\ W/m^2$ (W = Watt) (Weinzierl, 2008). This results in the sound intensity level as a function of distance (see also Weinzierl, 2008) with:

$$\text{Sound Intensity Level:} \quad L_I(r) = 10\log_{10}\frac{I(r)}{I_0} \tag{1}$$

The design of the feedback mapping was then implemented in the experiment in such a way that the loudness within the feedback range decreased from 60 dB to 46.7 dB as the distance to the target r increases. The nonlinear behaviour was determined by equation (2) and the linear behaviour by equation (3). Both functions are shown comparatively in Fig. 4.

$$\textit{Non-linear mapping:} \quad \textit{Sound Intensity Level}\ (r) = 40 + \left(20 \cdot \frac{L_I(r)}{60}\right) \tag{2}$$

$$\text{Linear mapping:} \quad \textit{Sound Intensity Level}\ (r) = 60 - \mathrm{r} \cdot 13.33 \tag{3}$$

2.3.1.2. Pitch. For pitch feedback, an analogous scenario to loudness was used for distance information. Therefore, the same nonlinear function (see Eq. (1)) was used to map frequency (in Hz) to distance ($F(r)$) in the range from 246.7 Hz to 180 Hz. That is, the pitch decreases as the distance r increases. The mapping is determined by equation (4) for the non-linear and equation (5) for the linear behaviour of the pitch feedback (see also Fig. 4)

$$\text{Non-linear } \textit{mapping:} \quad \textit{Frequency}\ (r) = 146.67 + \left(100 \cdot \frac{F(r)}{60}\right) \tag{4}$$

$$\text{Linear mapping:} \quad \textit{Frequency}\ (r) = 180 + (1 - r) * 66.67 \tag{5}$$

2.3.2. Visual feedback

Visual feedback was provided in the visually attended region, which means that the visual feedback was always close to the gripper (end-effector) of the robotic crane and followed the movement. Either the brightness of a grey circle or the size of a grey circle was mapped. Both brightness and size were mapped to 66% of the total distance for each target. Examples of the visual feedback are shown in Fig. 5. The increase in size of a grey circle and the current size in reference to an outer circle were indicators of performance.

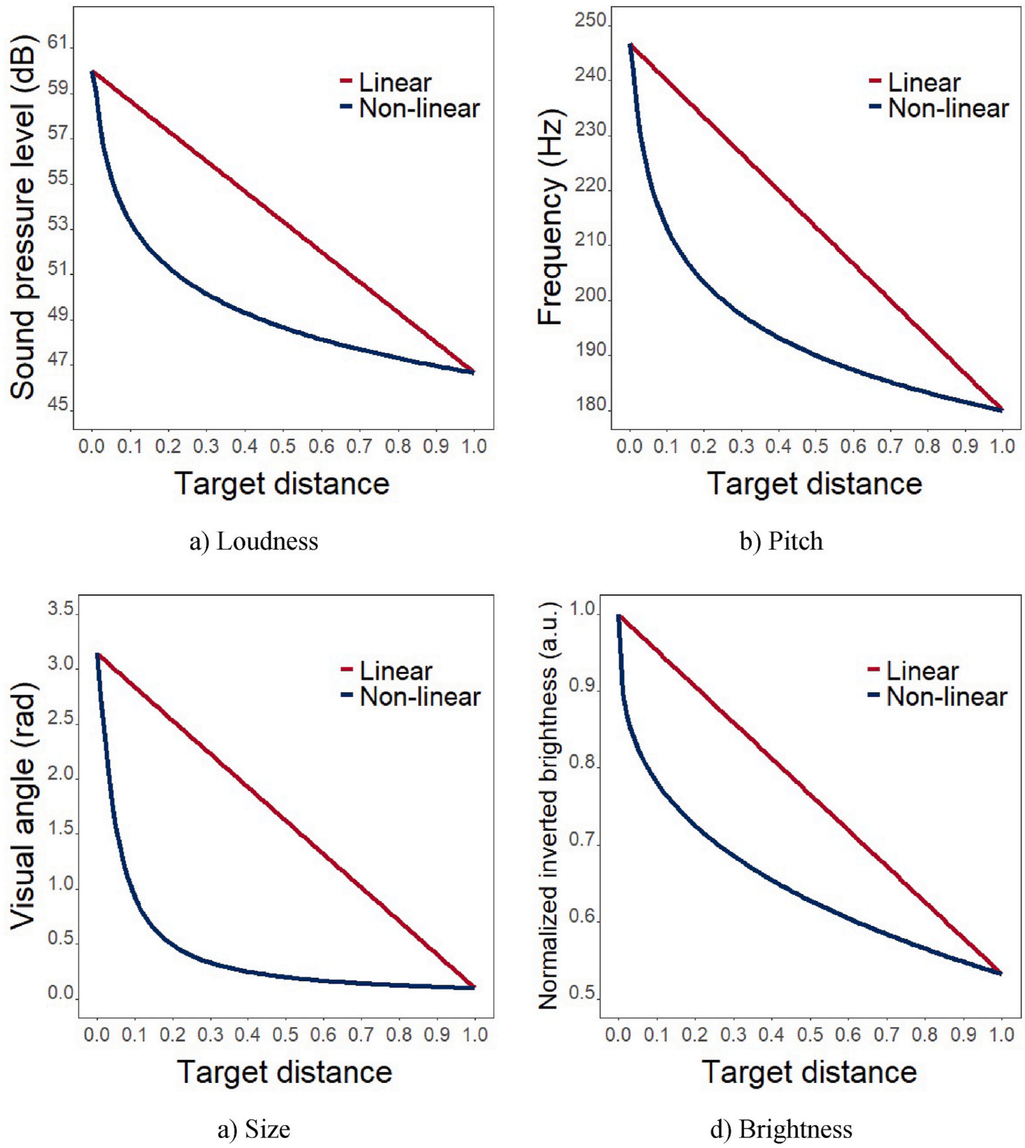


Fig. 4. Auditory and visual linear and non-linear feedback design of a) loudness, b) pitch, c) size, and d) brightness. The target distance is the distance at which feedback is present, which is 66% of the total distance between the targets. The curves show the change of feedback while approaching the target centre. *Note. Brightness was inverted for better comparability.

2.3.2.1. Size. Size refers to an expansion of a grey circle that was nested within another circle of fixed size ($d = \pi$). With approaching the target, the inner grey circle increased in size until both circles matched. A complete match occurred when the end effector was in the target centre. The expansion of the inner circle was mapped either non-linearly (eq. (6)) or linearly (eq. (7)). The non-linear mapping was modelled on the change in visual angle (in rad), analogous to the change in retinal size as a person approaches an object.

$$\text{Non}-\text{Linear mapping:}\quad Size_{Visual\ Angle}(r)= 2*atan\left(\frac{0.1}{2*r}\right) \tag{6}$$

$$\text{Linear mapping:}\quad Size\ (r)= 0.09+(\pi-0.09)*(1-r) \tag{7}$$

2.3.2.2. Brightness. To manipulate brightness with the RViz software, the alpha value (range 0–1) of the grey circle was scaled. Humans are more sensitive to decreases in brightness than increases, therefore, the brightness was mapped to decrease as the target was approached until it disappeared when the end-effector was in the centre of the target. Stevens' power law (Marks and Stevens, 1966) was used to map brightness non-linearly to reflect human perception of brightness (eq. (8), see Fig. 4).

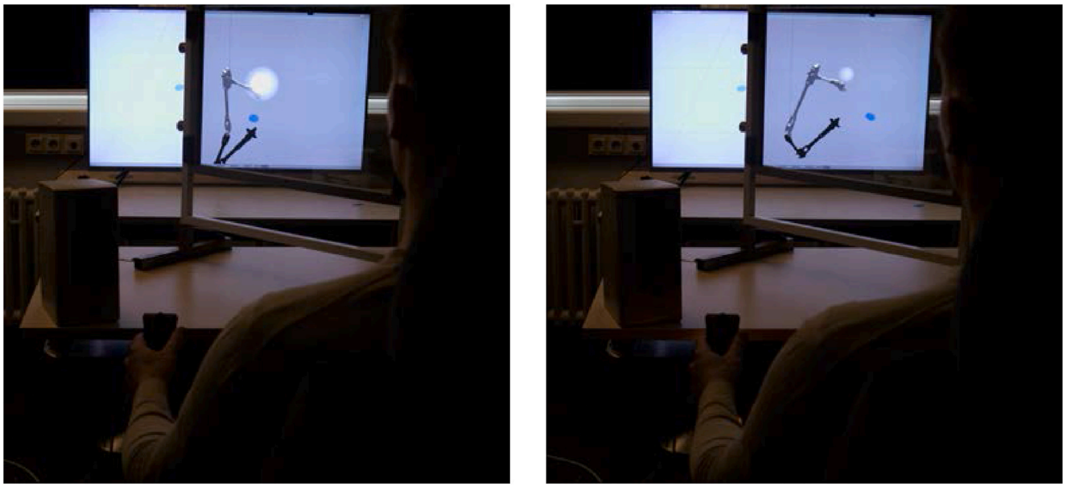
$$Non-linear\ mapping:\quad Brightness(r) = (r*0.1)^{0.33} \tag{8}$$

The linear mapping of brightness used a threshold constant to ensure that the circle disappeared when the end-effector was at the centre of the target, not earlier. This threshold is simulator dependent and was 0.37 in the current study. The linear mapping was implemented as described in equation nine.

$$\text{Linear mapping:}\quad Brightness(r)=\text{Threshold}*(1-r) \tag{9}$$

2.4. Procedure and instructional design

All participants were informed of the purpose of the study, the current corona regulations during the experiment and were given written task instructions afterwards. Then, participants went to the simulator where the seat and joystick positions were adapted to their preferences. The experiment started with a short demonstration of two movements with two oversized targets (circles) to help memorise the mapping of the



a) Size b) Brightness

Fig. 5. Example of visual feedback designs as implemented in the robotic crane simulator showing a) size and b) brightness feedback.

joysticks and familiarize with the procedure. Then, the actual robotic crane control task started. The participants' task was to alternatingly tap the two presented target pairs with the gripper tip in the centre. The first block was always a block without movement feedback for performer group assessment. This was followed by eight feedback and one no feedback block resulting in nine (counterbalanced) blocks after the performer group assessment. The sequence of the nine blocks was predetermined by a Latin square design and thus balanced across all participants to avoid learning effects. Each block consisted of eight targets with four movements each. Thus, each feedback category had 32 movements per participant (320 in total). A short survey was administered after each feedback block. This survey assessed acceptance with the Van der Laan scale and the mental load using the NASA TLX-R. The robotic crane control task was followed by a short demographic survey, after which the laboratory session ended. Overall, the experiment took 3.5 h to complete. The experiment was conducted as a within-subjects design, so that every participant was exposed to each feedback condition.

2.5. Dependant variables

2.5.1. Objective measures

Performance: Control performance was based on objective measures of the robotic crane movement. Movement time, accuracy in terms of constant error (distance to target centre at movement end) and variable error (standard deviation of constant error) were calculated to infer overall performance.

Skill: Skilled movements are characterised by a reduction in variability that should be reflected in the movement smoothness of the robotic crane. As with human movements, smoothness is a general criterion of movement skill. Smoothness was assessed by the spectral arc length of the robotic crane movement (for details see Balasubramanian et al., 2012, 2015). As additional indicator the lateral and vertical deviation of the executed trajectory from a straight line was evaluated using the root mean squared error (RMSE), (cf. Dreger et al., 2023).

2.5.2. Self-report measures

Self-report measures were used to evaluate the acceptance, usability, and mental load of the feedback. Acceptance was surveyed using a semantic-differential scale (Van Der Laan et al., 1997) on nine bipolar one-dimensional (e.g. useful-useless) ratings from -2 to 2 with five steps. Mean responses on the usefulness and satisfaction scales were calculated considering reverse phrasing of respective items. The NASA TLX-R was used to infer subjective cognitive load on six dimensions: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. The NASA TLX was used in the 21-item version, providing task load measures based on 5% increments for each dimension (Hart and Staveland, 1988).

2.6. Statistical analyses

The data pre-processing was carried out with MATLAB 2021a and R version 4.1.1. Performance and skill measures were compared using repeated measures and mixed-effects ANOVA. All statistical tests were conducted using an alpha level of 0.05.

3. Results

Data recording problems of two participants in which not all feedback could be provided required to recruit two additional participants.

The initial training block without feedback was used to classify low and high performers using a median split based on movement time, as low performers and early learning stages benefit from concurrent feedback (Wulf et al., 1998). Movement time in learning robotic crane control has been shown to be emphasised over accuracy and is more relevant for discriminating performance of robotic crane movements (Dreger et al., 2023). The no-feedback condition was excluded due to the training effects of performing the condition twice compared to other feedbacks. To rule out the possibility that the median split affected the balancing of learning effects, chi-squared tests for each feedback were conducted to compare the feedback distribution across the experimental session for low and high performance (see Table 1). All feedbacks were unaffected by learning ($p > .265$).

Table 1
Results of comparing the distribution of low and high performers across the experimental session for each feedback.

$N = 36$ Low vs. high performer	χ^2	df	p
Pitch-L	2	8	.981
Loud-L	8	8	.434
Pitch-NL	4	8	.434
Loud-NL	6	8	.647
Scale-L	6	8	.647
Scale-NL	8	8	.434
Bright-NL	4	8	.857
Bright-L	10	8	.265

3.1. Effects of high and low performer group on performance

A mixed-effects ANOVA was performed to compare the effect of Performer Group and Feedback on movement time. High performers showed significantly shorter movement times than low performers ($F(1, 32) = 18.98$, $p < .001$, $\eta_p^2 = .37$). The different types of feedback showed a tendency to be significantly different in terms of movement time ($F(3.65, 116.86) = 2.18$, $p = .081$, $\eta_p^2 = .04$). No significant interaction effect of Feedback and Performer Group on movement time was found ($p > .5$). A further mixed-effects ANOVA was performed to analyse the effect of Performer Group and Feedback on the constant error (accuracy). No main effects of either Performer Group or Feedback were found ($p > .05$). However, the analysis revealed a significant interaction between Performer Group and Feedback on accuracy, which was examined further below ($F(6.01, 192.35) = 2.15$, $p = .05$, $\eta_p^2 = .06$).

In addition, a mixed-effects ANOVA showed that trajectory smoothness was higher for high ($M = -3.23$, $SD = 0.94$) compared to low ($M = -3.65$, $SD = 1.15$) performers ($F(1, 32) = 8.40$, $p = .007$, $\eta_p^2 = .21$). Feedback Modality and Linearity did not show significant effects on trajectory smoothness.

3.2. Analysis of performer group feedback interaction

A repeated measures ANOVA revealed no effect of feedback on constant error for high performers (p > .05). In contrast, this analysis revealed a significant difference in constant error between the feedback conditions for low performers $F(4.37, 69.90) = 2.65$, $p = .036$, $\eta_p^2 = .23$ (descriptive statistics in Table 2). Therefore, the following performance and skill analyses will focus on low performers.

3.2.1. Analysis of low performers

Tukey adjusted post-hoc pairwise comparisons showed that nonlinear pitch feedback outperformed visual linear size feedback ($p = .022$) (see Fig. 6). Furthermore, a tendency was found that non-linear pitch feedback outperformed non-linear size feedback ($p = .056$) and non-linear loudness feedback showed a tendency to outperform visual linear size ($p = .095$) feedback. Variable error had no effect and therefore performance was constant across all feedbacks. ($p > .05$).

3.2.2. Modality and linear/non-linear feedback of low performers

A two-factor repeated measures ANOVA testing the effects of Modality and Linearity on constant error showed that accuracy was higher with auditory compared to visual feedback ($F(1, 17) = 6.02$, $p = .027$, $\eta_p^2 = .26$). Thus, low performers could make use of the auditory information to improve their movement accuracy (cf. Fig. 7). Linearity was not found to have a significant effect on constant error or to interact with Modality. There was no difference between loudness and pitch feedback within the auditory conditions. ($p > .05$). Descriptively non-linear pitch feedback showed the lowest constant error (cf. Table 2). Despite the effect on constant error, neither auditory nor visual feedback reduced movement time. ($p > .05$). Variable error, i.e. the movement stability and smoothness, was similar for auditory and visual feedback ($p > .05$).

Furthermore, no significant effect of Modality was found on the lateral and vertical expansion of the trajectory measured by the root mean square error of the distance (in cm) of the trajectory from a straight line between the start and end of the movement.

3.3. Self-report measures

Two two-factor repeated measures ANOVAs comparing the scores of the NASA TLX Scales and Modality or Linearity revealed no significant differences between the different modalities and the linear and non-linear mapping of the feedback ($p > .05$). There were also no significant differences between low and high performers observed ($p > .05$). However, the main effect of Scale showed significant differences between the six NASA TLX scales ($F(2.59, 90.74) = 9.79$, $p < .001$, $\eta_p^2 = .22$). Post-hoc analysis showed that effort was perceived higher than temporal demand, frustration, and physical demand ($p < .05$). Mental demand was perceived higher than physical demand and temporal demand ($p < .05$). In addition, performance was perceived lower than physical and temporal demand ($p < .05$, see Fig. 8).

Fig. 9 shows the usefulness and satisfaction with the feedback. Nine (out of 324) missing data points from the usefulness scale were replaced by the mean of the respective feedback rating to ensure that all data could be used in the following analysis. Two mixed-effects ANOVAs were performed with Scale and either Modality or Linearity as the within factor and Performer Group as the between factor. Significant main effects were found for Performer Group ($F(1, 34) = 4.29$, $p = .046$, $\eta_p^2 = .11$)) and (usefulness and satisfaction) Scale ($F(1, 34) = 53.22$, $p < .001$, $\eta_p^2 = 0.61$)). Additionally, a tendency to significance was found for Modality ($F(1, 34) = 3.98$, $p = .054$, $\eta_p^2 = .10$)). A significant interaction effect was found between Performer Group and Scale ($F(1, 34) = 7.52$, $p = .010$, $\eta_p^2 = .18$)) as well as Scale and Modality ($F(1, 34) = 29.65$, $p < .001$, $\eta_p^2 = .47$)).

The post-hoc analyses showed that high performers rated usefulness and satisfaction with feedback higher than low performers. Both performer groups rated usefulness of feedback higher than satisfaction with feedback. The Tukey-adjusted post-hoc pairwise comparison of the interactions (Performer Group x Scale; Scale x Modality) showed that low performers rated satisfaction lower than usefulness ($p < .05$). Both performer groups showed an overall tendency to rate the visual feedback higher than auditory feedback.

Generally, auditory feedback had low satisfaction, but high usefulness scores ($p < .05$) and auditory feedback was less satisfying than visual feedback ($p < .05$). In contrast, satisfaction and usefulness were rated similarly for visual feedback.

4. Discussion

This study aimed to investigate the performance, skill, and self-reported task load, usefulness, and satisfaction of visual or auditory concurrent feedback to improve movement accuracy in manual robotic crane control. The effect of the mapping function (non-linear or linear) of the feedback was also investigated.

The results show that the effectiveness of the feedback given depended on the participants' level of performance after the training. Feedback benefited low performers but not high performers. This is in line with (Wulf et al., 1998, 1999) who found concurrent feedback more effective in the early stages of learning a complex skill. High performers may deem feedback as irrelevant or even distracting information, or conversely, accuracy feedback was more useful for low performers. It could also be a deliberate choice to use or ignore feedback.

In spite of efforts to present visual feedback close to the end-effector and without visual interference, neither performance nor skill could be improved on the basis of concurrent visual movement information. This finding is consistent with the claims of multiple resource theory that would expect lower performance if two information draw on the same cognitive resource (Wickens, 2002). However, these findings are in contrast to Sigrist et al. (2011), who found that visual feedback was helpful to improve oar position and blade orientation in rowing. This

Table 2
Mean constant error with standard deviations in paratheses ($n = 18$).

Feedback		Pitch NL	Pitch L	Loud L	Loud NL	Size L	Size NL	Bright NL	Bright L
Constant Error	*M (SD)*	3.34 (1.38)	3.75 (1.46)	3.58 (1.46)	3.45 (1.38)	4.21 (1.89)	4.09 (1.60)	3.59 (1.53)	3.60 (1.53)

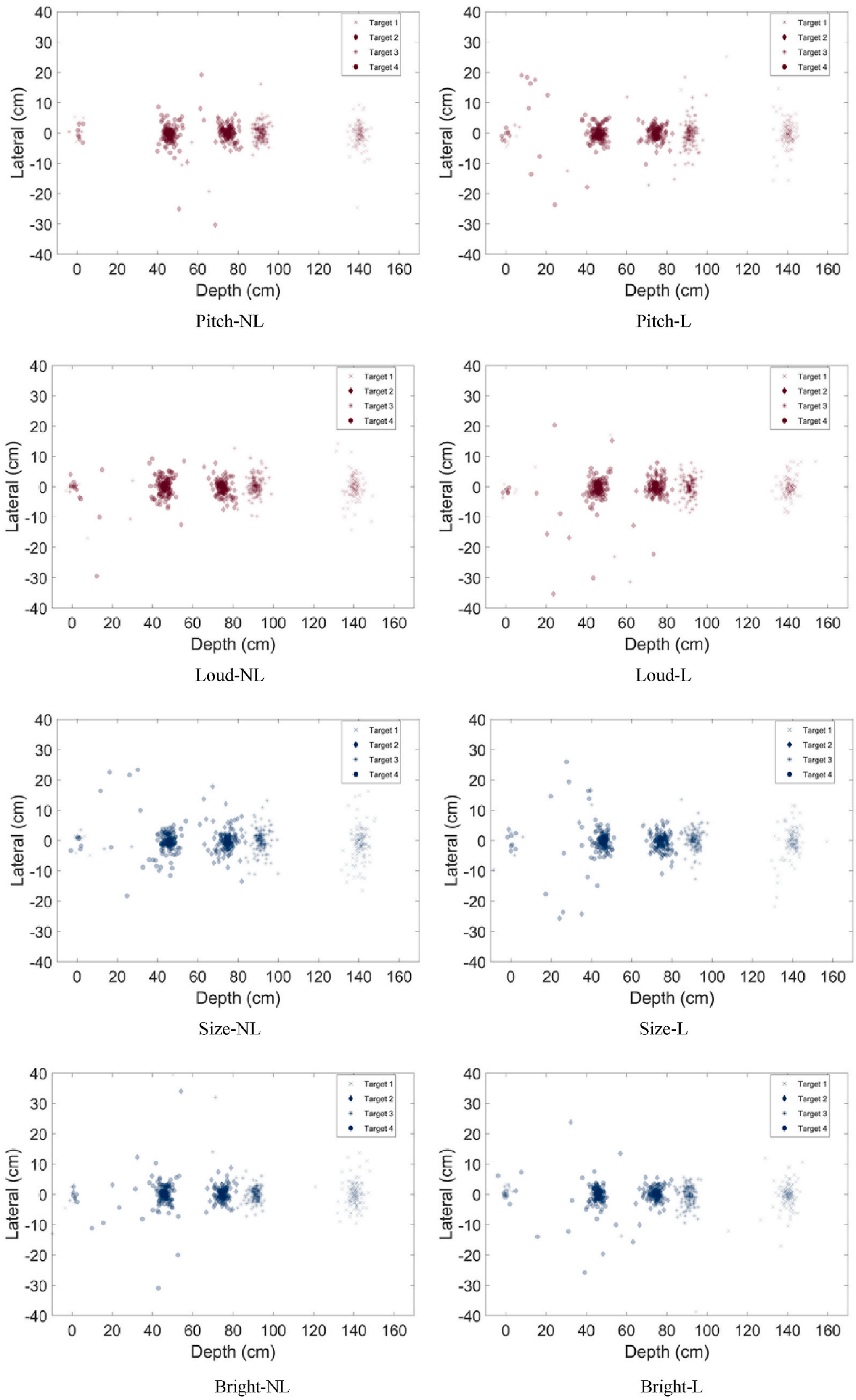


Fig. 6. The constant error of each target is displayed for each feedback condition. The colours indicate the modality, with dark red for auditory, blue for visual feedback. The target location is rotated on the depth axis and thus represents movement length. The distribution of target endpoints is shown for four different movement difficulties that are averaged at block level in the analysis. Data points around zero represent start errors. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

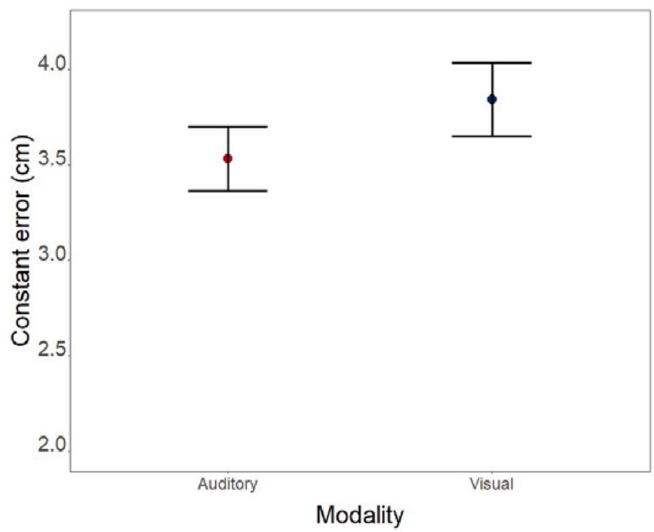


Fig. 7. Distance to the target centre for the aiming movement measured as mean constant error for the auditory and visual modality and without feedback. Error bars represent standard errors.

may lie in the nature of the rowing task, where each hand of the rower is separately handling the oar and blade. In addition, the focus in rowing is on speed and course of the boat, where information from the blades is not as much interfering as visual feedback on end-effector position in robotic crane control. Furthermore, feedback characteristics such as a simple visual representation of the paddle by a curved line may help visual appeal. Consistent with the predictions of Wickens (2002) and the findings of Chen et al. (2016); Effenberg et al. (2016); Sigrist et al. (2013); and Vidal et al. (2020), auditory feedback was useful in improving the accuracy of low performers. Within auditory feedback, no difference in accuracy was found between pitch and loudness feedback. However, non-linear pitch feedback was descriptively superior to linear pitch and loudness feedback. The usefulness of pitch feedback in robotic crane control was also demonstrated by Mavridis et al. (2015) in reducing movement times, although accuracy was not assessed in this study. In contrast, the movement time was not affected by the different types of feedback, which can be because accuracy was targeted in the feedback design of the current study.

The non-linear or linear feedback mapping showed no significant effects. About the absence of the effect can only be speculated. Either because the effects of linear and non-linear conditions compensated for one another across different feedback designs, or because feedback design needs to be further tuned to the respective feedback concept and human perception.

The accuracy measures were found to be supported by the self-report measures. No differences were found between the feedback types in terms of task load. The ratings indicate a low cognitive task load despite perceived mental effort. The effort may represent the mental load of the participants associated with controlling the robotic crane and learning the motor transformations, which may be conflated with the effects of the designed feedback in the evaluation.

Acceptance in terms of usefulness showed no difference between modalities which contrasts with the objective measures on constant error. However, there was a difference between perceived usefulness and satisfaction for auditory feedback. Notably, auditory feedback is perceived as more useful than satisfactory. Satisfaction is associated with attributes such as pleasant, nice, and desirable. This is in line with other experiments with sonification, where auditory feedback is perceived as annoying (Bazilinskyy et al., 2019), especially when sinusoidal tones are used instead of sounds (e.g. chords) (Effenberg et al., 2005). No differences between usefulness and satisfaction were found for visual feedback. Generally, satisfaction with feedback in both performer groups was rated lower than usefulness. This may be partially due to the increased visual load of visual feedback and the non-linear mapping, which failed to improve performance in terms of movement times (especially in the visual modality).

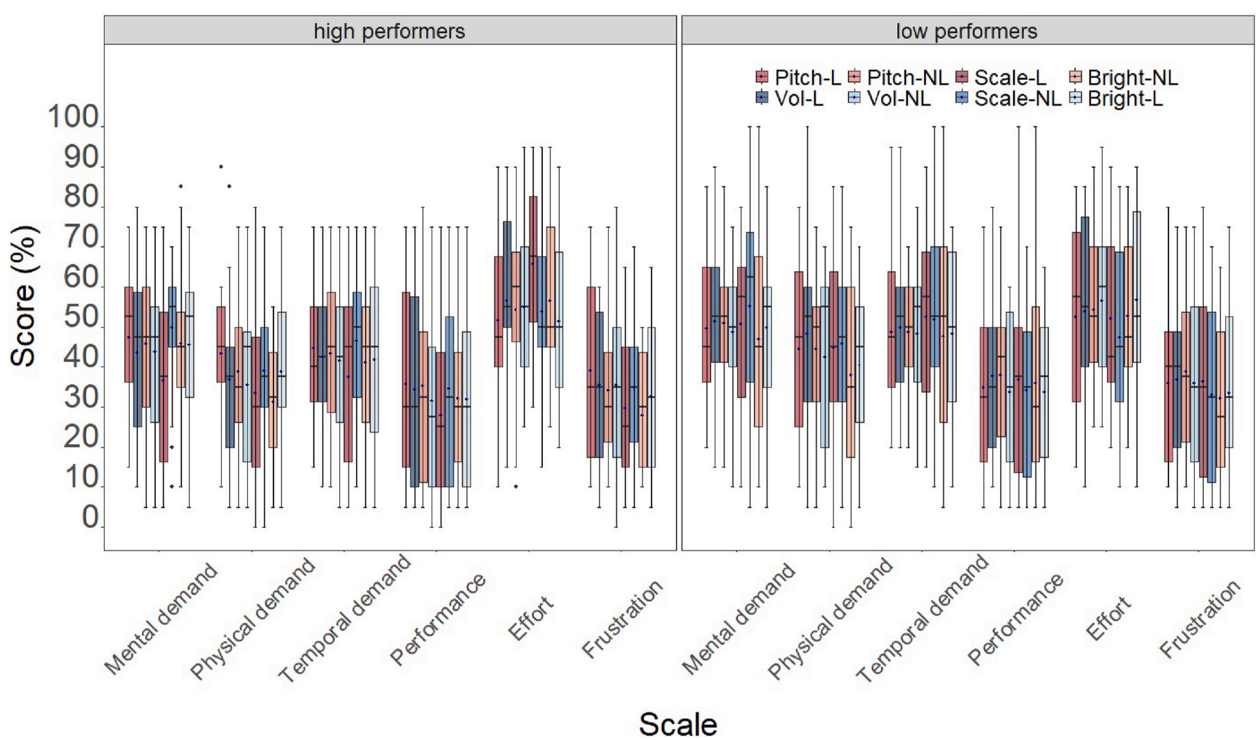


Fig. 8. NASA TLX scores of high and low performers in percent (%) for the feedback conditions shown for all NASA TLX scales for low and high performer. The blue diamond indicates the mean score. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

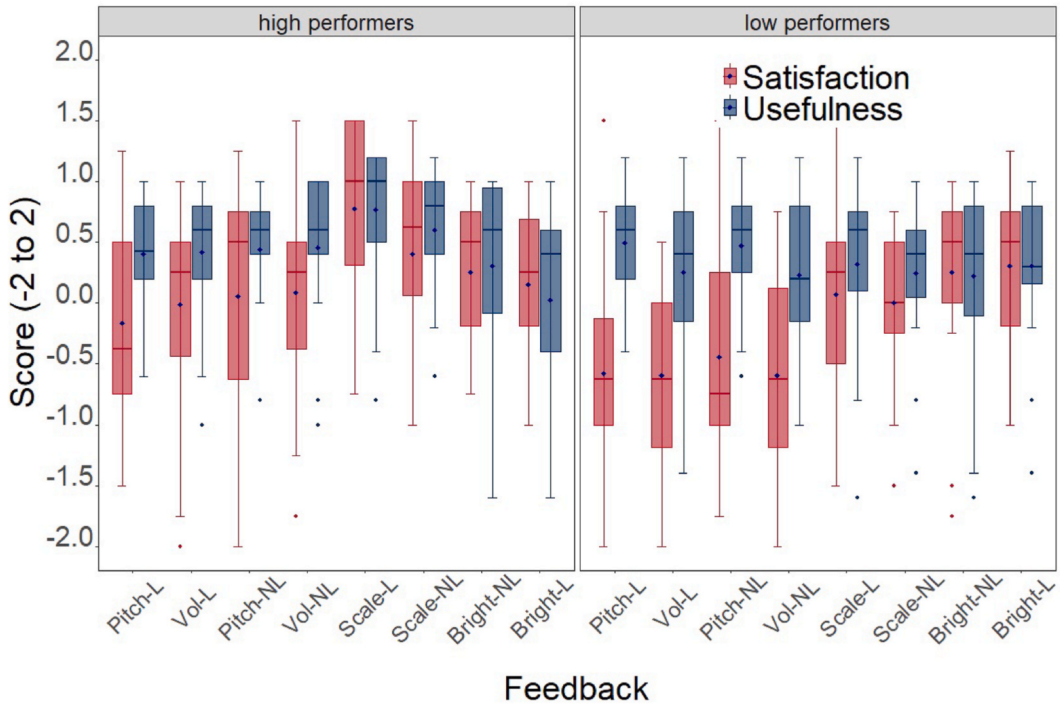


Fig. 9. Usefulness and satisfaction scores of high and low performers assessed with the Van der Laan acceptance scale on a range from −2 (negative) to 2 (positive) for each movement feedback. The whiskers show the interquartile range, the lines in the boxes show the median and the diamonds show the mean.

4.1. Limitations

When interpreting the results, several limitations need to be considered. Bimanual robotic crane movements are highly complex, and motor control can change significantly with different movement targets. This means that aiming for real objects in 3D space may require additional feedback information to maintain high accuracy. Furthermore, results are only applicable to low performers in the early stages of building experience in robotic crane control. Whether high performers use feedback quickly and efficiently, or whether high performers do not require feedback cannot be answered within the current study design. Nonetheless the authors believe, that with appropriate feedback design, more experienced operators can benefit from auditory feedback. In addition, the simulator used provided excellent controllability of the experiment but remains a simplification in terms of fidelity and ecological validity of the task. Although transfer to real-world application can be successful (Ovaskainen, 2005; Ranta, 2009), the feedback must be tested in real-world operations. The experiment was designed such that learning across participants was reduced to avoid corresponding confounds. This means that concurrent feedback, when present, improves performance, but the effect must still show improved long-term performance. Furthermore, in future studies, performance with feedback should be compared with performance without feedback to better assess the gain and the relief or additional mental effort associated with feedback. Finally, the tested sample comprised students and staff from an academic background that may limit the generalisability of the results, however, this ensured the complete novelty of the task and the same training level while using the feedback.

4.2. Conclusions

To conclude, this study was the first attempt to systematically provide feedback based on human perceptual characteristics to enhance the accuracy of bimanually controlled robotic crane movements. Auditory feedback yielded higher accuracy and was perceived as more satisfying than visual feedback for low performers. Further auditory mappings and characteristics need to be explored to extend use to more experienced operators and to improve the usefulness of the feedback regarding the improvement of movement time.

Funding

This work was supported by the ERA-NET Cofund Action “Forest-Value—Innovating the forest-based bioeconomy” and by Fachagentur Nachwachsende Rohstoffe (FNR, Germany), Forskningsradet (The Research Council of Norway), and VINNOVA, The Swedish Innovation Agency. ForestValue received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement N° 773324.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Balasubramanian, S., Melendez-Calderon, A., Burdet, E., 2012. A robust and sensitive metric for quantifying movement smoothness. IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 59 (8), 2126–2136. https://doi.org/10.1109/TBME.2011.2179545.

Balasubramanian, S., Melendez-Calderon, A., Roby-Brami, A., Burdet, E., 2015. On the analysis of movement smoothness. J. NeuroEng. Rehabil. 12 (1), 112. https://doi.org/10.1186/s12984-015-0090-9.

Baudry, L., Leroy, D., Thouvarecq, R., Chollet, D., 2006. Auditory concurrent feedback benefits on the circle performed in gymnastics. J. Sports Sci. 24 (2), 149–156. https://doi.org/10.1080/02640410500130979.

Bazilinskyy, P., Larsson, P., Johansson, E., De Winter, J.C.F., 2019. Continuous auditory feedback on the status of adaptive cruise control, lane deviation, and time headway: an acceptable support for truck drivers? Acoust Sci. Technol. 40 (6), 382–390. https://doi.org/10.1250/ast.40.382.

Chen, J.L., Fujii, S., Schlaug, G., 2016. The use of augmented auditory feedback to improve arm reaching in stroke: a case series. Disabil. Rehabil. 38 (11), 1115–1124. https://doi.org/10.3109/09638288.2015.1076530.

Chollet, D., Micallef, Jean Paul, Rabischong, Pierre, 1988. Biomechanical signals for external biofeedback to improve swimming techniques. In: Swimming Science V. Human Kinetics.

Clarkson, P.M., James, R., Watkins, A., Foley, P., 1986. The effect of augmented feedback on foot pronation during barre exercise in dance. Res. Q. Exerc. Sport 57 (1), 33–40. https://doi.org/10.1080/02701367.1986.10605386.

Ding, R., Cheng, M., Han, Z., Wang, F., Xu, B., 2022. Human-machine interface for a master-slave hydraulic manipulator with vision enhancement and auditory feedback. Autom. ConStruct. 136, 104145 https://doi.org/10.1016/j.autcon.2022.104145.

Dreger, F.A., Chuang, L.L., Rinkenauer, G., 2023a. Analysis of learning the bimanual control of (tele)operating joint space controlled robotic arms with 4 degrees of freedom using the two-timescales power law of learning. Ergonomics 0 (0), 1–13. https://doi.org/10.1080/00140139.2023.2270784.

Dreger, F.A., Chuang, L.L., Rinkenauer, G., 2023b. Comparing Operator Learning, Performance, Cognitive Load, and Trajectories of End Effector and Joint Controlled Robotic Arms for Support System Design. Submitted for Publication.

Effenberg, A., 2005. Movement sonification: effects on perception and action. IEEE Multimedia 12 (2), 53–59. https://doi.org/10.1109/MMUL.2005.31.

Effenberg, A., Bliss, J.P., Weber, A., Zinke, A., 2005. MotionLab sonify: a framework for the sonification of human motion data. Ninth Int. Conf. Inf. Vis. (IV'05), 17–23. https://doi.org/10.1109/IV.2005.84.

Effenberg, A., Fehse, U., Schmitz, G., Krueger, B., Mechling, H., 2016. Movement sonification: effects on motor learning beyond rhythmic adjustments. Front. Neurosci. 10 https://doi.org/10.3389/fnins.2016.00219.

Effenberg, A., Fehse, U., Weber, A., 2011. Movement Sonification: Audiovisual benefits on motor learning. BIO Web of Conferences 1, 00022. https://doi.org/10.1051/bioconf/20110100022.

Eriksson, A., Petermeijer, S.M., Zimmermann, M., de Winter, J.C.F., Bengler, K.J., Stanton, N.A., 2019. Rolling out the red (and green) carpet: supporting driver decision making in automation-to-manual transitions. IEEE Transactions on Human-Machine Systems 49 (1), 20–31. https://doi.org/10.1109/THMS.2018.2883862.

Häggström, C., Englund, M., Lindroos, O., 2015. Examining the gaze behaviors of harvester operators: an eye-tracking study. Int. J. For. Eng. 26 (2), 96–113. https://doi.org/10.1080/14942119.2015.1075793.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (Eds.), Advances in Psychology, vol. 52. North-Holland, pp. 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9.

Hartsch, F., Dreger, F.A., Englund, M., Hoffart, E., Rinkenauer, G., Wagner, T., Jaeger, D., 2022. Positive and Negative Work Practices of Forest Machine Operators: Interviews and Literature Analysis.

Keogh, J.W.L., Hume, P.A., 2012. Evidence for biomechanics and motor learning research improving golf performance. Sports BioMech. 11 (2), 288–309. https://doi.org/10.1080/14763141.2012.671354.

Magill, R.A., Anderson, D., 2017. Motor Learning and Control: Concepts and Applications, eleventh ed. McGraw-Hill Education.

Manner, J., Gelin, O., Mörk, A., Englund, M., 2017. Forwarder crane's boom tip control system and beginner-level operators. Silva Fenn. 51 (2) https://doi.org/10.14214/sf.1717.

Marks, L.E., Stevens, J.C., 1966. Individual brightness functions. Percept. Psychophys. 1 (1), 17–24. https://doi.org/10.3758/BF03207815.

Mavridis, N., Pierris, G., Gallina, P., Papamitsiou, Z., Saad, U., 2015. On the subjective difficulty of Joystick-based robot arm teleoperation with auditory feedback. In: 2015 IEEE 8th GCC Conference & Exhibition, pp. 1–6. https://doi.org/10.1109/IEEEGCC.2015.7060097.

Meyer, D.E., Kornblum, S., Abrams, R.A., Wright, C.E., 1988. Optimality in human motor performance: ideal control of rapid aimed movements. Psychol. Rev. 95 (3), 340–370.

O'Brien, B., Hardouin, R., Rao, G., Bertin, D., Bourdin, C., 2020. Online sonification improves cycling performance through kinematic and muscular reorganisations. Sci. Rep. 10 (1), 20929 https://doi.org/10.1038/s41598-020-76498-0.

Oppici, L., Dix, A., Narciss, S., 2021. When is knowledge of performance (KP) superior to knowledge of results (KR) in promoting motor skill learning? A systematic review. Int. Rev. Sport Exerc. Psychol. 1–25. https://doi.org/10.1080/1750984X.2021.1986849.

Ovaskainen, H., 2005. Comparison of harvester work in forest and simulator environments. Silva Fenn. 39 (1) https://doi.org/10.14214/sf.398.

Ovaskainen, H., Palander, T., Tikkanen, L., Hirvonen, H., Ronkainen, P., 2011. Productivity of different working techniques in thinning and clear cutting in a harvester simulator. Balt. For. 17 (2), 11.

Ovaskainen, H., Uusitalo, J., Väätäinen, K., 2004. Characteristics and significance of a harvester operators' working technique in thinnings. Int. J. For. Eng. 15 (2), 67–77. https://doi.org/10.1080/14942119.2004.10702498.

Purfürst, F.T., 2010. Learning curves of harvester operators. Croat. J. For. Eng. 9.

Ranta, P., 2009. Added values of forestry machine simulator based training. Proceedings of the International Conference on Multimedia and ICT Education 6.

Schmidt, R.A., Lee, T.D., Winstein, C.J., Wulf, G., Zelaznik, H.N., 2019. Motor Control and Learning: A Behavioral Emphasis, sixth ed. Human Kinetics.

Sharma, D.A., Chevidikunnan, M.F., Khan, F.R., Gaowgzeh, R.A., 2016. Effectiveness of knowledge of result and knowledge of performance in the learning of a skilled motor activity by healthy young adults. J. Phys. Ther. Sci. 28 (5), 1482–1486. https://doi.org/10.1589/jpts.28.1482.

Shea, C.H., Wulf, G., 1999. Enhancing motor learning through external-focus instructions and feedback. Hum. Mov. Sci. 18 (4), 553–571. https://doi.org/10.1016/S0167-9457(99)00031-7.

Sigrist, R., Fox, S., Riener, R., Wolf, P., 2016. Benefits of crank moment sonification in cycling. Procedia Eng. 147, 513–518. https://doi.org/10.1016/j.proeng.2016.06.230.

Sigrist, R., Rauter, G., Riener, R., Wolf, P., 2013. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: a review. Psychon. Bull. Rev. 20 (1), 21–53. https://doi.org/10.3758/s13423-012-0333-8.

Sigrist, R., Schellenberg, J., Rauter, G., Broggi, S., Riener, R., Wolf, P., 2011. Visual and auditory augmented concurrent feedback in a complex motor task. Presence Teleoperators Virtual Environ. 20 (1), 15–32. https://doi.org/10.1162/pres_a_00032.

Song, J., Wang, Y., An, X., Ma, S., Wang, D., Gan, T., Shi, H., Yang, Z., Liu, H., 2022. Novel sonification designs: compressed, iconic, and pitch-dynamic auditory icons boost driving behavior. Appl. Ergon. 103, 103797 https://doi.org/10.1016/j.apergo.2022.103797.

Todorov, E., Shadmehr, R., Bizzi, E., 1997. Augmented feedback presented in a virtual environment accelerates learning of a difficult motor task. J. Mot. Behav. 29 (2), 147–158. https://doi.org/10.1080/00222899709600829.

Valéry, B., Scannella, S., Peysakhovich, V., Barone, P., Causse, M., 2017. Can an aircraft be piloted via sonification with an acceptable attentional cost? A comparison of blind and sighted pilots. Appl. Ergon. 62, 227–236. https://doi.org/10.1016/j.apergo.2017.03.001.

Van Der Laan, J.D., Heino, A., De Waard, D., 1997. A simple procedure for the assessment of acceptance of advanced transport telematics. Transport. Res. C Emerg. Technol. 5 (1), 1–10. https://doi.org/10.1016/S0968-090X(96)00025-3.

Vidal, A., Bertin, D., Drouot, F., Kronland-Martinet, R., Bourdin, C., 2020. Improving the pedal force effectiveness using real-time sonification. IEEE Access 8, 62912–62923. https://doi.org/10.1109/ACCESS.2020.2984211.

Walker, B.N., Nees, M.A., 2011. Theory of sonification. In: Hermann, T., Hunt, A., Neuhoff, J.G. (Eds.), The Sonification Handbook, first ed. Logos Publishing House, pp. 9–39.

Walraven, J., 1972. TNO Test for Stereoscopic Vision. Laméris Ootech B.V.

Weinzierl, S., 2008. Grundlagen. In: Weinzierl, S. (Ed.), Handbuch der Audiotechnik. Springer, pp. 1–39. https://doi.org/10.1007/978-3-540-34301-1_1.

Wickens, C.D., 2002. Multiple resources and performance prediction. Theor. Issues Ergon. Sci. 3 (2), 159–177. https://doi.org/10.1080/14639220210123806.

Wulf, G., Hörger, M., Shea, C.H., 1999. Benefits of blocked over serial feedback on complex motor skill learning. J. Mot. Behav. 31 (1), 95–103. https://doi.org/10.1080/00222899909601895.

Wulf, G., Shea, C.H., 2002. Principles derived from the study of simple skills do not generalize to complex skill learning. Psychon. Bull. Rev. 9 (2), 185–211. https://doi.org/10.3758/BF03196276.

Wulf, G., Shea, C.H., Matschiner, S., 1998. Frequent feedback enhances complex motor skill learning. J. Mot. Behav. 30 (2), 180–192. https://doi.org/10.1080/00222899809601335.

Zhu, B., Kaber, D., Zahabi, M., Ma, W., 2020. Effects of feedback type and modality on motor skill learning and retention. Behav. Inf. Technol. 39 (4), 431–442. https://doi.org/10.1080/0144929X.2019.1599068.