Learning effect in joystick tactile guidance

ZIKMUND, P.; HORPATZKÁ, M.; MACÍK, M.

IEEE Transactions on Haptics

Early access

ISSN: 1939-1412

DOI: https://doi.org/10.1109/TOH.2024.3368663

Learning effect in joystick tactile guidance

Pavel Zikmund, Michaela Horpatzká and Miroslav Macík

***Abstract*—Haptic feedback is a method to provide tactile guidance in scenarios requiring multiple senses and divided attention like aviation. Earlier tests on a flight simulator and an in-flight test using the proposed tactile guidance method have shown the need to study its learning process. In this study, twelve participants completed two tactile guidance tasks without visual feedback across twelve sessions to analyze the learning effect. The paper shows an improvement between sessions in guidance accuracy, response time, and self-assessed workload. On the other hand, reaction delay is not affected by the training. The percentage improvement between the initial and trained skills reached 30 % in guidance accuracy performance.**

***Index Terms*—Human-Computer Interaction, Human Performance, Tactile Devices, Learning Effect.**

I. Introduction

Tactile guidance methods promise intuitive and easy operation. Despite the simplicity of use, many researchers show the presence of the learning effect, but only a few measure it. The learning effect is usually eliminated by prescribing different methods order for particular participants. That is possible in experiments comparing two or more guidance methods. Such experiments show the best guidance method among all tested methods. On the other hand, such experiments have a first impression character. The guidance methods performance of trained participants is not provided. The lack of participants' training increases results variance in the comparative experiments. One example is a recent study by de Rooij et al. [1], which presented a visual display to supplement haptic feedback on the side stick to maintain safe flight conditions. The experiment involved 15 professional pilots. The learning effect was observed, though it was not initially expected, as the pilots received training before the experiment.

The authors presented new hardware for tactile guidance [2]. The hardware consists of a joystick with a sliding element. The sliding element moves into or out of the handle surface of a joystick under the operator's fingers, see Fig. 1. The device has two main functions: warning and guidance. The vibration mode of the sliding element is dedicated to the warning. The front or back movement of the sliding element means guidance instruction. The primary motivation for the development of the device was the prevention of accidents caused by unwanted stalls and spins. Loss of control is the most common cause of general aviation accidents in recent decades. Consequently, authorities such as the NTSB [3] have repeatedly highlighted the loss of control in their most wanted list of transportation safety improvements. One of the possible methods for reducing the number of these accidents is by improving pilot-aircraft interaction. A loss of control occurs when the aircraft reaches a critical angle of attack and starts to stall. Stall warning systems are usually based on auditory and visual warnings in general aviation. However, Geehan [4] has presented findings that most pilots find haptic warnings to be more effective than auditory or visual alerts. Despite this fact, auditory and visual warning systems are more common due to their lower price. Although the presented device is applied to aviation, its use in other sectors using a joystick for any scenario involving guidance methods, such as control of work machines or telemanipulation is possible.

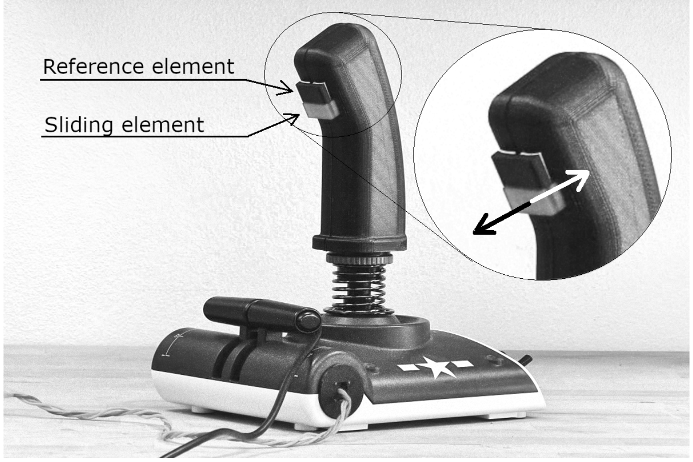

Fig. 1. Joystick with sliding and reference element used in the experiment. The reference element defines the position of the sliding element when the joystick is in the target position. The arrows represent possible sliding element movement directions.

In the previous research, the authors tested the system with tactile feedback on a flight simulator, and an in-flight test [5]. These tests did not show the expected improvement in pilot-aircraft interaction but revealed the necessity for defining a training process. Piloting relies on muscle memory, and the introduction of haptic feedback during rapid learning phases led to significant performance changes and an increase in measured parameter variances. Moreover, conducting training in flight tests would entail high costs when compared to laboratory investigations. Recognizing this, the primary focus of the research presented in this article is to explore and understand the learning curve associated with haptic feedback guidance. The learning effect is investigated on a guidance task without visual feedback. The reason is that pilots use mostly only peripheral vision for aircraft control feedback. Vision should stay reserved for other control and navigation tasks. The long-term goal of our haptic feedback research is to decrease the visual workload in aircraft control. The learning process in using haptic feedback is tracked over time through

Pavel Zikmund and Michaela Horpatzká are with the Brno University of Technology, Brno, 616 69, Czech Republic (email zikmund@fme.vutbr.cz) Miroslav Macík is with the Czech Technical University in Prague, Prague, 121 35, Czech Republic.

Manuscript submitted April 28, 2023; revised November 19, 2023 and February 6, 2024

participants' performance. Understanding the learning curves will enable us to determine the necessary training required before conducting further tests, which should confirm or reject the benefit of haptic feedback in aircraft control.

II. RELATED WORK

The section summarizes the prior research and important aspects of haptic perception with a focus on tactile guidance on the learning effect. The interaction method described in this paper aims to be used in general aviation, where often one aircraft is used by several pilots with only moderate experience. Such a system requires fast learning, good memorability, adequate efficiency in its task, and minimizing user errors. These requirements are in accordance with the definition of classical usability by J. Nielsen [6]. Haptic perception involves both cutaneous and kinesthetic stimuli [7], [8], where most approaches to supplying various kinds of information to users rely on cutaneous stimuli. Systems involving haptic guidance are often used in application domains where the use of vision is limited or where vision is saturated by other tasks like aviation.

Aviation is a domain where haptic methods are frequently used, as the demands on vision are high, and divided attention could be the source of human errors. Haptic feedback has been a natural part of aircraft control since the beginning of heavier-than-air flight. The forces in aircraft main controls correspond to aerodynamic forces. The stiffness of controls increases with airspeed, and aerodynamic effects cause stick vibrations at low airspeed close to stall conditions when turbulent air hits the surface of the elevator. Modern aircraft with power steering simulate some natural effects by employing artificial systems. For instance, the stick shaker provides stall warnings [9] by vibrating the control column. More sophisticated systems aimed to provide even more detailed information about the angle-of-attack by modulating the vibrations [10]. The purpose this system is similar to ours, however our method is based on shape-change.

Several studies tried to employ haptic interaction as an additional channel to provide information in the aviation domain. For example, Van Erp [11] studied a tactile display consisting of 64 vibrotactile elements to help a pilot with guidance and control tasks. He found that the localized vibration on the pilot's body was easily coupled with spatial information, such as direction to a waypoint or threat. Cardin et al. [12] presented a system comprising eight vibrotactile actuators attached to the pilot's body. The system stimulates a pilot to catch his attention and provides information about the aircraft's attitude without needing to read the flight instruments. This research shows the benefits of haptic warning and guidance in comparison to only visual guidance during long-term flights. Fellah and Guiatni [13] proposed a tactile display to help with keeping an aircraft within safe limits and provide situational awareness support. The tactile display consisting of low-cost tactile actuators was designed to substitute the saturated visual channel. Vibration motors were placed on the pilot's body (abdomen, back, left, and right sides). The authors concluded that tactile feedback is suitable for feeding information about the aircraft's state. These sources highlight the opportunities and possibilities of using tactile feedback in aviation and illustrate how rich information can be provided through the haptic channel. On the contrary, they are inconvenient for general aviation as they require attaching tactile actuators directly to the pilot's body. Also, the mentioned studies do not focus on the influence of the participant's experience and details on how they learned to use the tested systems.

As described above, the application of the interaction method described in this study requires fast training and good memorability. The following paragraph focuses on these aspects in methods that involve haptic interaction. One of the frequently monitored parameters influencing performance during training is the time intervals between each training session. Wang et al. [14] investigated the training time interval duration influence on a tactile orientation discrimination task. Two compared groups trained at one-day and one-week intervals. The training intervals affected only the early stage of learning up to the third session. Both groups reached the same level after five sessions. Such results indicate that training intervals might be flexible if enough sessions are planned. Tactile perceptual learning has also been studied in the context of Braille script reading by Kass et al. [15]. They found that tactile learning is more intense between sessions than within a session. These findings correspond to previously published results by Karni and Sagi [16], [17] for visual learning. Other research papers are dedicated to auditory and visual perceptual learning. The concept of slow learning between sessions and fast learning within the first session were distinguished by Atienza et al. [18], Qu et al. [19], and Molloy et al. [20]. Ashley and Pearson [21] presented the importance of consolidation between sessions. Repeated within-day testing or overtraining leads to detrimental effects on perceptual learning. Consolidation of learned information during sleep has the power to prevent such deficits in learning. The time intervals between learning sessions have a minor or no effect on performance after a small number of sessions [14]. Performance improvement is more evident between learning sessions than within a session [15], except for the fast increase in performance within the first learning session [18]–[20]. These facts were reflected in the design of the experiment described in section III-B.

In order to evaluate the performance and improvement of pilots during tactile guidance training, various parameters might be used to measure the effectiveness of the training. These parameters include measures of accuracy, reaction time, workload, and situational awareness. In this paragraph, we focus on reaction times, while the measurement of guidance accuracy is discussed in the next paragraph. Reaction times in visual and tactile tasks were measured by Kim et al. [22]. The average tactile reaction time was 0.241 seconds, while the visual reaction time was 0.329 seconds. Workload and situational awareness were addressed by Elliot et al. [23]. They tested visual and tactile navigation displays in a navigation and guidance task in a strenuous outdoor environment. The research concluded that the visual display supported global awareness, while the tactile display supported local guidance, leading to a lower mental workload rating. The positive effect of tactile guidance on workload reduction was also reported by

De Stigter, Mulder, and Van Paassen [24] in the study on application to a haptic flight director. An improvement in situational awareness was reported also in [13]. The aim of this study is to evaluate the effectiveness of tactile guidance methods, so we need to emphasize important measurable parameters along with results achieved by comparable methods. Humans can manifest faster reaction times to tactile stimuli than to visual ones [22], which further motivates the application of our method in the aviation domain. Tactile interaction might be better suited for local guidance, contributing to lower overall mental workload rating in combined tasks [23], [24], possibly leading to better situational awareness [13].

The effect of tactile feedback on the accuracy of guidance needs to be considered from several perspectives. Haptic guidance might improve the guidance accuracy as stated in [24] and by Nieuwenhuizen and Bulthoff [25] in the context of haptic shared control for personal aerial vehicles. On the other side, Voudouris et al. [26] state that the perception of tactile stimuli presented on a moving hand is systematically suppressed, which could be attributed to the limited capacity of the brain to process task-irrelevant sensory information. Authors investigated whether humans are able to enhance in parallel movement relevant tactile signals when performing goal-directed reaching movement. Conducted experiments suggest that participants were able to flexibly modulate tactile sensitivity by suppressing movement-irrelevant and enhancing movement-relevant signals in parallel when performing goal-reaching tasks. Juravle and Spence [27] investigated sensory suppression in complex motor tasks like juggling. The experiment required participants to detect gaps in the continuous signal provided by different modalities (haptic, auditory). The authors state that participants were significantly less sensitive to detecting a gap in tactile stimulation while juggling. The results demonstrate movement-related tactile sensory suppression related to the decision component in tactile suppression. Humans could trigger tactile suppression in the brain before the motor command. Tactile suppression might be a risk factor for the application of our method in aviation as humans may tend to suppress the tactile stimuli when focusing on another demanding task.

Other articles focus on the change in sensitivity to tactile stimuli. Chamnongthai et al. [28] propose a method to tackle the effect of temporal decrease of human force-detection capabilities in finger holder setups. They investigated the impact of Stochastic resonance on the user's haptic performance. The results show that human fingertip sensitivity significantly increases when Stochastic resonance is applied. Bensmaia et al. [29] investigate the effects of extended suprathreshold vibratory stimulation on the sensitivity of three types of neural afferents (slowly adapting Type 1, rapidly adapting, and Pacinian). The results show that prolonged suprathreshold stimulation can result in substantial desensitization of all neural afferent types. Temporal change in sensitivity to tactile stimuli – tactile adaptation should be reflected in the design and experiment as users will be in contact with the sliding element for long periods. Additionally, this factor is expected to have adverse effects on performance over time, potentially affecting within-session performance.

Interaction methods that employ tactile guidance have applications in many domains, including medicine, aviation, or even navigation of those with vision impairments. The frequent motivation factor is that the use of vision is limited in the particular domain, either by objective factors (visually impaired, low light environment) or by the overload of vision by another task (i.e., medicine, aviation). Many approaches related to tactile guidance rely on complex apparatuses directly connected to some part of the human body, e.g., [13]. Unlike that, we decided to follow the *come as you are* design constraint proposed by Triesch and Malsburg [30], which means that the users do not need to wear special equipment such as vests or gloves to use the system. This should enhance the acceptance of our method in the general aviation domain.

The analysis of the related work supports the need to investigate the learning effect of the proposed guidance method. Haptic feedback may help with faster skill acquisition and improve performance in path-following tasks. On the other hand, Sullivan, Pandey, Byrne, and O'Malley [31] observed a slight decrease in accuracy in setup with haptic feedback in research focused on movement smoothness while performing a mirror-tracking (path-following) task. Also, it is necessary to reflect on the effect of desensitization caused by physiological limitations of human neural afferents involved in the tactile sense as described by Bensmaia et al. [29] and by Chamnongthai et al. [28]. The experiment should focus on tactile guidance accuracy and reaction time and how these measures are affected by training. As suggested by Ashley and Pearson [21], the experiment should consider the effect of overtraining. In our experiment, we will measure performance changes within test sessions as well as between test sessions.

III. METHOD

A. Hardware

The hardware is based on the Mad Catz Pacific AV8R joystick, as shown in Fig 1. The shape of the joystick mechanism was modified to decrease the force peak needed to move out of the central position. The original handle was replaced by a handle with a sliding element. The sliding element performs a translational movement in and out of the joystick handle and is located under the operator's index finger. Just above the sliding element is the reference element. The reference element is rigid and represents the neutral position of the haptic guidance system. The operator's finger can be placed on the interface between these two elements, and information is obtained by comparing the position of the sliding element relative to the reference element. There are three main clues. If the sliding element is aligned with the reference (the operator can't feel any distinction between the elements), that means the target position is obtained and no action from the operator is required. In case the sliding element moves backward compared to a reference, that means the joystick should be pulled back. If the sliding element moves forward compared to a reference, the joystick should be moved forward as well (in this case, the operator feels a force of the sliding element moving towards his finger). If the required joystick movement is within one-quarter of the

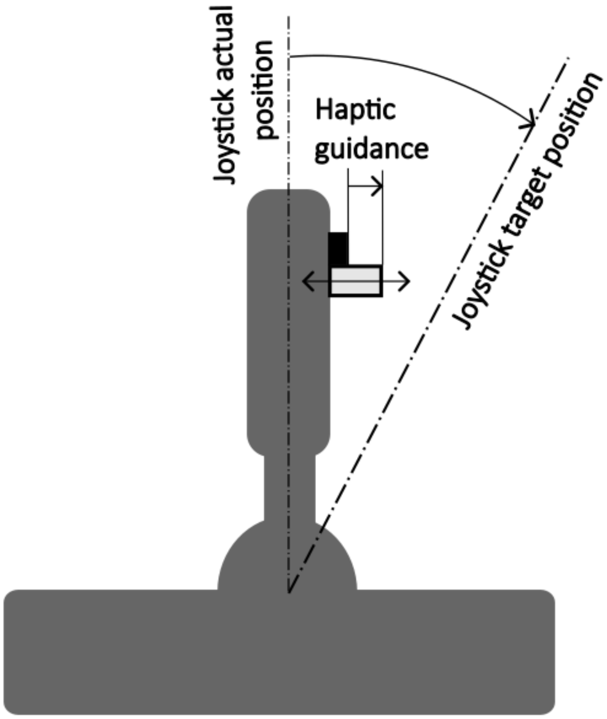

Fig. 2. The Haptic guidance dimension is proportional to the demanded trajectory of the joystick, ranging from the actual position to the target position. The Haptic guidance dimension reacts to continuous changes in both the actual and target positions.

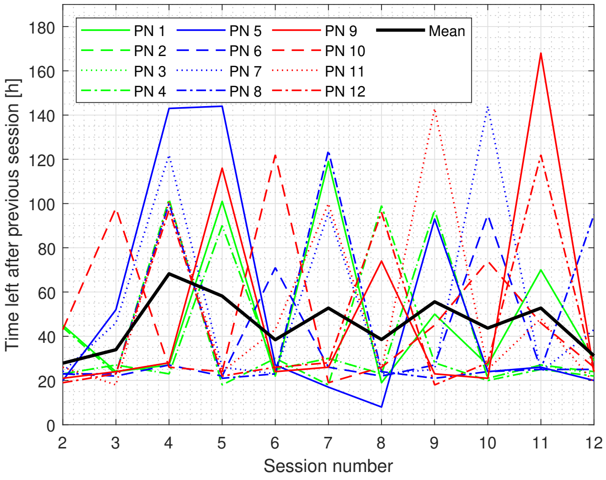

Fig. 3. The time between every two following sessions for all participants and the mean value.

joystick's range, the size of the position difference between the sliding element and the reference front surfaces is proportional to the joystick deflection required to reach the target position, as shown in Fig. 2. If higher deflection is required, the sliding element's deflection saturates at its maximum inner or outer position. The operator applies force only to the joystick handle. The sliding element then reacts to the actual position and indicates the distance to the target position. No force acting on the sliding element is required. The sliding element is powered by two SG90 digital servomotors, each providing a maximum force of 20 N, and its movement range is 8 mm. A vibration mode actuated by short-period front-back sliding element movements has not been used in this experiment.

B. *Test procedure*

Twelve undergraduate and graduate volunteer students were recruited to participate in this experiment. Participant eligibility was verified, and written informed consent was obtained from each participant. The group contained two females and ten males aged 19 to 26 (mean 21.67, SD 2.23). The participants were recruited from the student population in contrast to the previous simulator and flight tests where professional pilots were recruited [5]. The guidance task described in the following paragraphs did not involve aircraft dynamics. Therefore, no pilot skills were required to participate in this experiment. The subjects repeated the experiment in 12 sessions with a break between sessions of at least 8 hours. The mean time between sessions was 45 h (SD 36 h, and median 26 h) as shown in Fig. 3. The difference between the mean and median values is caused by weekend breaks between sessions. The maximum time between sessions was seven days due to an illness in the case of Participant no. 9.

Each session included two different tasks with no prior training. Both tasks have been done only once in the same order within each session. Subjects were guided only by the sliding element without any visual feedback. The computer screen provided information only about the beginning and end of each task but did not provide information during the experiments. As shown in Fig. 4, the first task was to guide the joystick to 20 randomly generated positions. Each position was generated as a constant random position with a uniform distribution over the joystick front-back travel range with a random duration uniformly distributed between 3 and 6 seconds. Both parameters were randomized across both sessions and participants. All participants used their dominant hand.

A continuously changing target position characterizes the second task shown in Fig. 5. Twelve different variants of 60-second courses of joystick movement were prerecorded before the experiment. The angular rotation speed of the joystick was (Mean 4.61 deg/s, SD 5.02 deg/s, and peak values of 30 deg/s). Each subject started with a different variant according to Latin square order. Each session contained one of the twelve variants. This helped eliminate the effect of the difficulty of variants from the learning effect. The subjects completed a short questionnaire after each session. They assessed their workload in both tasks. The Bedford workload rating scale published by Roscoe [32] was used. The questionnaire had a range from 1 to 10, where 1 means an insignificant workload and 10 means that is not possible to complete the task. Each subject received a USB stick as a reward after finishing all twelve testing sessions.

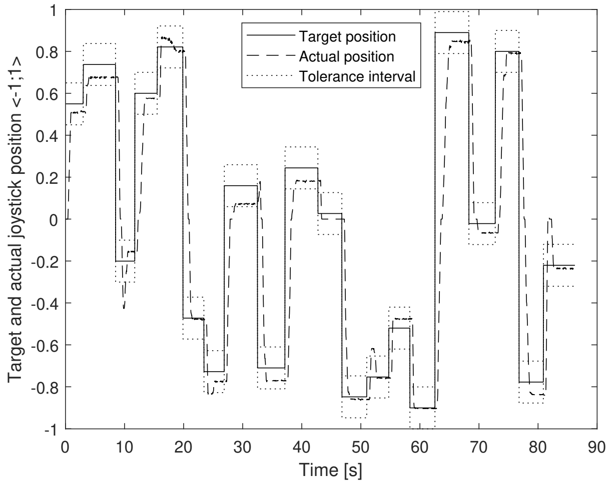

Fig. 4. One set of measured data from Task 1, illustrating the actual and target positions of the joystick over time. The task involves twenty random target positions.

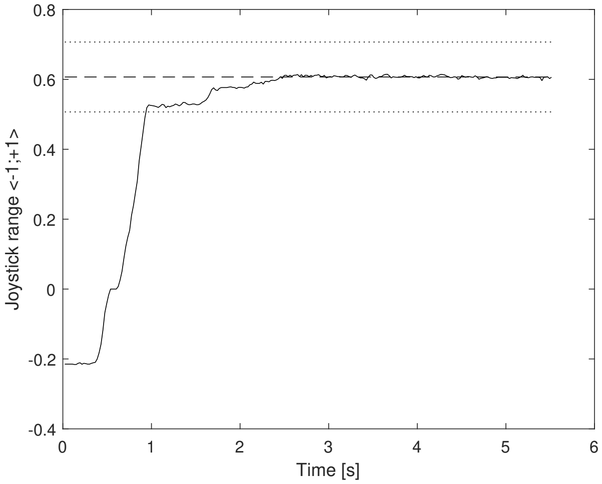

Fig. 6. This is one of the correct response records. The correct response was defined as reaching the target position tolerance interval without identified overshoot or non-minimum phase response.

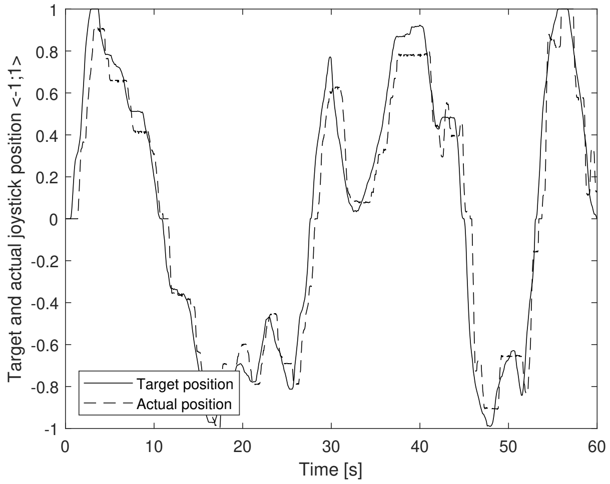

Fig. 5. One set of measured data from Task 2, illustrating the actual and target positions of the joystick over time.

C. Metrics

The learning effect was evaluated as an improvement of some quantitative criteria over time. The two main criteria were the time to reach the target position (TRTP) and the average error (AE). TRTP was measured from the time of the target position generation to the first reach of the target position tolerance interval. The tolerance interval was set at +-5 % of the joystick range around the target position. AE between the target and the actual joystick positions was measured from the first achievement of the target interval till the new target position was generated. These criteria were analyzed both between-session and within-session. Additional evaluated criteria were reaction delay (RD) and self-assessed workload. RD represents the time interval between the generation of the new target position and the beginning of the response. That means RD is a part of TRTP. TRTP and RD criteria were evaluated only for the first task. The second task was evaluated by the AE between the target and actual joystick positions.

The learning effect was assessed within-session for AE and TRTP in Task 1. For statistical analysis, twenty individual attempts were combined into ten levels. This was done by averaging each pair of consecutive attempts to create a new value for each level. This approach was necessary because Mauchly's test of Sphericity could not be applied when the number of repeated measurements exceeded the number of participants.

The mean values of all criteria were calculated for each session and each participant. The mean values of all parameters were analyzed by one-way repeated measures ANOVA to investigate the influence of the learning effect. In addition, Mauchly's test of sphericity indicates the assumption of sphericity and if Mauchly's test has been violated, Greenhouse-Geisser correction was used. The amount of $p < 0.05$ is considered as a significant difference. Finally, Post hoc analysis was done with Tukey's HSD Test. This test revealed homogeneous groups which helped to define the learning curve character along sessions.

D. Response characteristics

Most typical response characteristics were defined and identified among all the measured responses in Task 1. The correct response pattern is shown in Fig. 6. It was defined by reaching the target interval and remaining in it till the new random target position was generated. Two other conditions defining correct responses were that the overshoot and non-minimum phase response did not occur.

The overshoot refers to response characteristics crossing the target interval and returning back, remaining within the interval for the last second of the target position duration. The overshoot characteristic is shown in Fig. 7.

The non-minimum phase response is characterised by an incorrect initial decision regarding the response direction, as shown in Fig. 8.

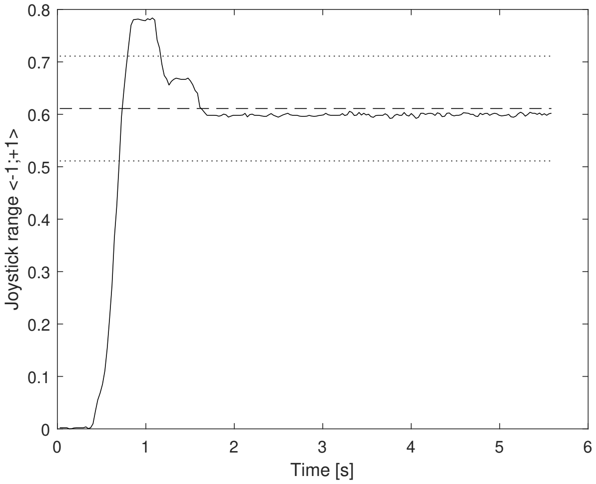

Fig. 7. A sample record of the response with overshoot. The overshoot was defined by crossing the tolerance interval and returning to it.

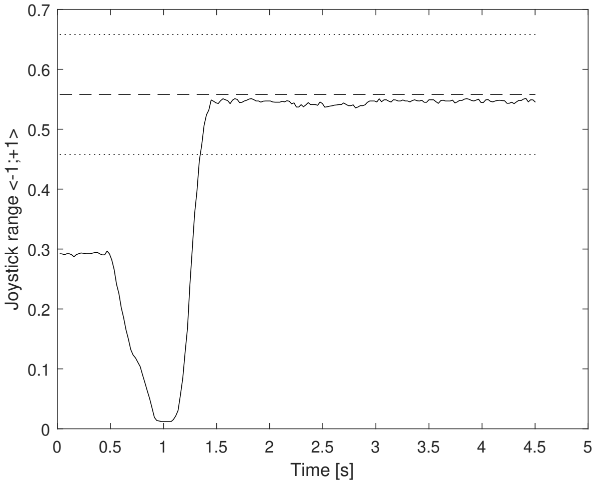

Fig. 8. A sample record of a non-minimum phase response. The non-minimum phase response includes all attempts where participants initially move in the opposite direction of the generated target position.

IV. Results

The chapter is divided into two parts corresponding to Task 1 and Task 2. The first part begins with the results of response characteristics, followed by a statistical evaluation for both tasks.

A. Task 1

1) Response characteristics: Twenty-six responses of all 2,880 attempts (0.9 %) did not reach the target interval till the last second of the interval duration. In this case, the time to reach the target position and the error between the target and actual positions were not defined in these attempts. The improvement of correct response attempts in Task 1 is shown in Fig. 9. The number improved from 63.3 % in the first session to over 90 % in the last three sessions. Other characteristic responses were overshoot and non-minimum phase responses.

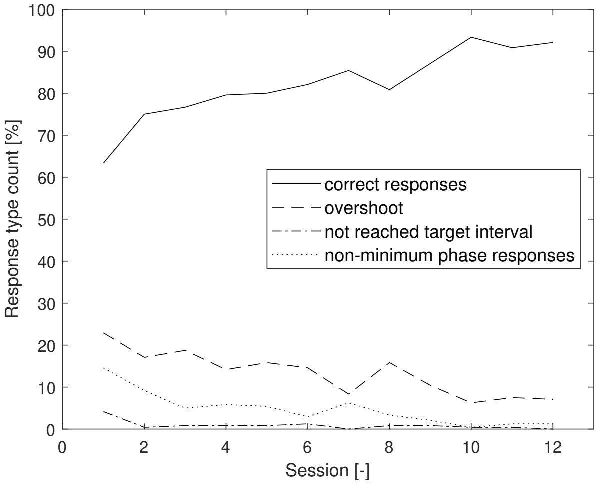

Fig. 9. The count of all participants' correct responses, responses with overshoot, responses without the reaching tolerance interval of the target position, and non-minimum phase responses over 12 sessions in Task 1.

Three hundred and eighty-one attempts (13.22 %) include an overshoot of the target interval. The overshoot influenced the error between the actual and target positions in a negative way. The count of overshoot response improves from 22.9 % in the first session to 7.0 % in the last session. The time to reach a target position was not influenced or penalized by an overshoot. The non-minimum phase response includes a representation of one hundred and thirty-eight attempts (4.79 %). The non-minimum phase response negatively influenced the TRTP. The count of non-minimum responses improves from 14.58 % in the first session to equal or less than 1.25 % in the last three sessions.

2) Time to reach target position: TRTP Fig. 10 represents the mean values of each participant for each session. The TRTP showed statistical significance between sessions in repeated measures ANOVA with Greenhouse-Geisser correction (F (4.292, 47.21) = 2.782, p = 0.0341). Post hoc analysis was performed using Tukey's HSD Test, which revealed two homogeneous groups. These groups showed a significant improvement only in time between the first two sessions. The sessions from the second to the last can be grouped together as one homogeneous group, with a slight increase in TRTP value observed during the last two sessions.

The means of each participant's TRTP evaluated within-session are displayed in Fig. 11. The TRTP showed statistical significance within-session in repeated measures ANOVA with Greenhouse-Geisser correction (F (3.974, 43.719) = 5.161, p = 0.0017). Post hoc analysis revealed three homogeneous groups. The distribution of the homogeneous groups does not indicate any within-session learning effect of TRTP. Homogeneous groups revealed the best performance in the middle of Task 1 between attempts no. 9-14. The slowest response was achieved in an attempt no. 4-8 and 15-16.

3) Average error between target and actual joystick positions: The AE between the target and actual position of the joystick after the first achievement of the target interval Fig. 12 represents the mean values of the parameter for each

This article has been accepted for publication in IEEE Transactions on Haptics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TOH.2024.3368663

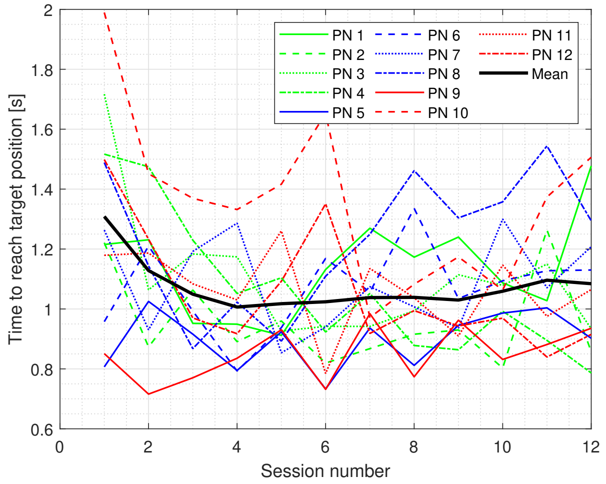

Fig. 10. Time to reach target position in Task 1 for all participants for each session. The time was measured from the generation of the new target position and the first reach of the tolerance interval.

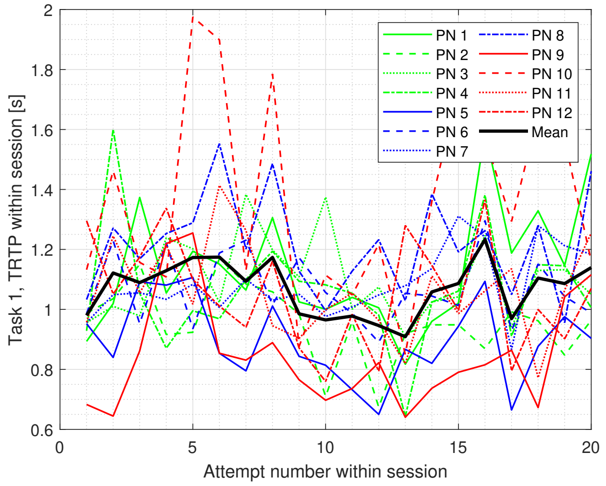

Fig. 11. Average time to reach target position in Task 1 for all participants within-session.

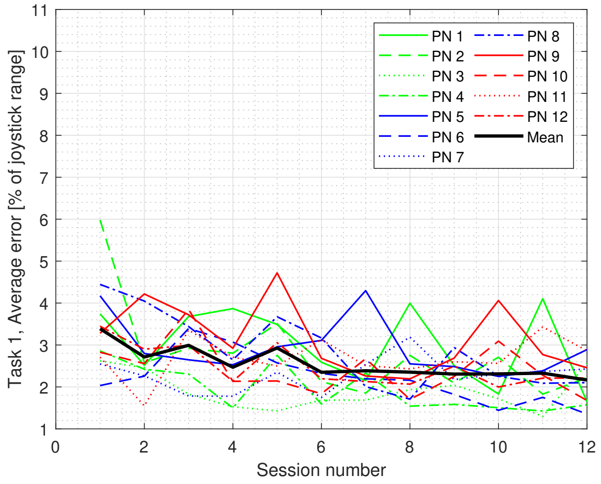

Fig. 12. Task 1: Improvement in guiding accuracy over 12 sessions. The accuracy is defined as the mean difference between the target and actual joystick positions after the first reach of the tolerance interval.

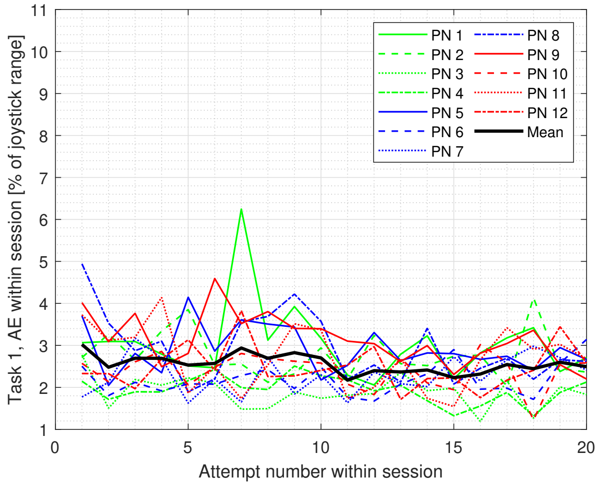

Fig. 13. Task 1: Average error between the target and actual joystick positions within-session over 20 attempts.

participant for each session. The AE showed statistical significance between sessions in repeated measures ANOVA with Greenhouse-Geisser correction (F (5.200, 57.206) = 4.256, p = 0.0021). Post hoc analysis revealed three homogeneous groups. Significant improvement is observed in the first six sessions, however sessions 3 and 5 indicate a reversal. An improvement cannot be distinguished from the sixth to the eleventh session. The mean value decreased from 3.39 % (SD = 1.08 %) in the first session to 2.16 % (SD = 0.51 %) in the last session.

AE was analyzed within-session, as well as TRTP. The means of each participant are shown Fig. 13. Again, twenty single attempts were merged into ten levels. The AE showed statistical significance in within-session repeated measures ANOVA (F (9, 99) = 2.462, p = 0.0141). Post hoc analysis was done with Tukey's HSD Test. However, despite the statistical significance in the ANOVA, post hoc revealed only a single homogeneous group consisting of all levels. The within-session learning effect on AE was not observed.

4) Reaction delay: RD represents the time interval between the generation of the new target position and the beginning of the response. Repeated measures ANOVA did not show any significant difference in RD along sessions (F (11, 121) = 1.265, p = 0.253). Mauchly's Sphericity Test was not violated in this case. The total mean value of RD is 0.4459 s (SD = 0.054 s). The fastest participant's mean reaction delay was only 0.369 s, whereas the slowest participant's mean reaction delay was 0.484 s.

5) Self-assessed workload: Workload assessments in both tasks were included in the questionnaire after each session. Participants ranked their workload using the Bedford workload scale. The maximum value on the scale used by the participant was 9 which means an extremely high workload with no spare capacity. This value was used only by Participant 1. He was

This article has been accepted for publication in IEEE Transactions on Haptics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TOH.2024.3368663

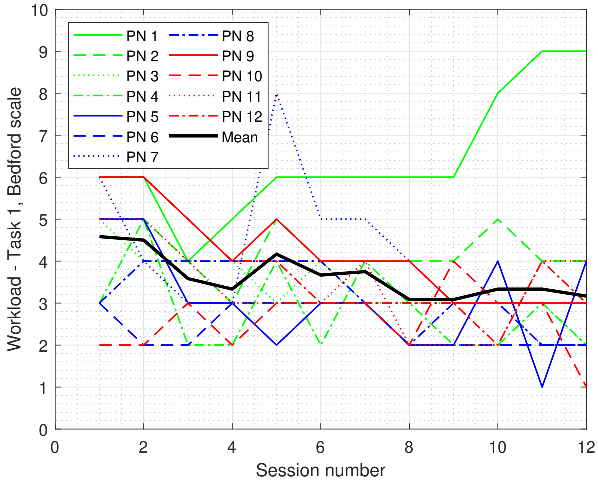

Fig. 14. Self-assessed workload in Task 1 over 12 sessions. Participant 1 was removed from ANOVA analysis as an outlier.

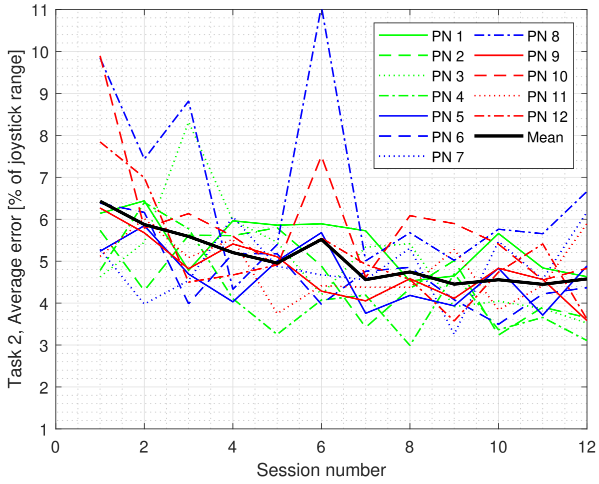

Fig. 15. Task 2: Average error between the target and actual joystick positions for each participant.

excluded from the evaluation as an outlier. He was the only one who assessed the workload as growing over the course of learning time as shown in Fig. 14. Possible reasons are stated in the discussion.

The last session was excluded from the ANOVA because of the same reason as merging attempts to levels in within-session analyses. The workload showed statistical significance between sessions in repeated measures ANOVA (F (10, 100) = 5.176, $p < 0.0001$). Post hoc analysis revealed three homogeneous groups. These groups displayed an improving trend until the eighth session. The mean values start around levels 4 and 5 on the Bedford scale. These levels are defined as ”insufficient spare capacity for easy attention to additional tasks” and ”reduced spare capacity, additional tasks cannot be given the desired amount of attention.” The mean values at the end of the training were 2.81 (SD = 1.03). Values 2 and 3 are characterized as ”Workload low” and ”Enough spare capacity for all desirable additional tasks.”

B. Task 2

1) Average error between target and actual joystick positions: The average error between the target and actual joystick positions throughout the second task in Fig. 15 represents the mean values for each participant. The AE in Task 2 showed statistical significance between sessions in repeated measures ANOVA with Greenhouse-Geisser correction (F (4.0673, 44.741) = 5.222, $p < 0.0001$). Post hoc analysis revealed three homogeneous groups. The homogeneous groups revealed no improvement after the seventh session. AE in Task 2 decreased from 6.43 % (SD = 1.83 %) in the first session to 4.58 % (SD = 1.16 %) in the last session. The reversal in the sixth session is given mostly by the coincident weak performance of Participants 8 and 10.

2) Self-assessed workload: The workload values on the Bedford's scale assessed for the second task differ slightly from the first task. No outliers were identified in the workload data for the second task. Each participant's means shows Fig. 16. The workload in Task 2 showed statistical significance between sessions in repeated measures ANOVA with Greenhouse-Geisser correction (F (4.305, 47.355) = 7.082, $p < 0.0001$). Post hoc analysis revealed five homogeneous groups. These groups demonstrate an improving trend throughout all twelve sessions except few reversals. The mean values start around levels 4 and 5 on the Bedford scale and reach 2.33 (SD = 1.15) at the end of the training.

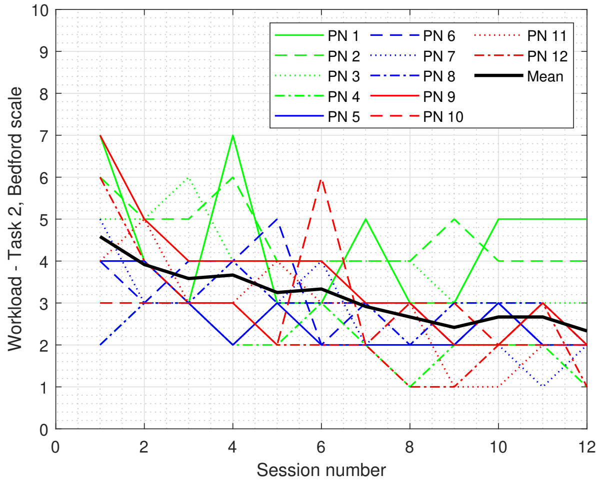

Fig. 16. Self-assessed workload in Task 2 over 12 sessions.

V. DISCUSSION

The measurement of the learning effect was the primary objective of the experiment. The specific research question was to assess the necessary training duration for the guidance method using the joystick with the sliding element. The effect is evaluated by considering the performance of the participants along with the session number. The average error (AE) and the time to reach the target position (TRTP) in Task 1 were evaluated both between-session and within-session.

This article has been accepted for publication in IEEE Transactions on Haptics. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TOH.2024.3368663

Building upon prior research by Zikmund et al. [2], which consisted of a single session, this study extended the investigation to 12 sessions. Two parameters that were measured in both experiments in the first session can be compared. The mean value of TRTP in the previous research was 1.548 s (SD = 0.48 s), while a recent result was 1.31 s (SD = 0.35 s). There was one difference in the Task 1 definition, which consisted of 30 attempts in the previous research. This difference might be due to the improvement of the guidance function of the sliding element during its development. Another parameter measured in the previous research was AE in Task 2, where the result was observed as 6.67 % (SD = 1.12 %), which is comparable to the value of 6.42 % (SD = 1.77 %) measured in the first session of the current study Task 2.

The parameter AE can also be compared between both tasks in the current study. Both tasks exhibit differences in their absolute values. One reason for this difference is that the AE measurement started after the first achievement of the target interval in the first task. AE in the second task was continuously measured throughout the entire guidance duration. The second reason is that guidance to a static position allows for a more precise reaction compared to continuously changing the target position in Task 2.

Within-session measurement of AE and TRTP parameters did not reveal a learning effect during the session. Within-session performance might be influenced by desensitization, as described by Bensmaia et al. [29]. Their results exhibit transitional characteristics that do not correspond to the within-session characteristics analysed in this study. A longer duration of Task 1 would be necessary to study this effect and distinguish between desensitization and within-session learning effects. Within-session results correspond to the findings by Kass et al. [15], who found more intense learning between-session in the case of tactile learning. Fast learning within the first session was not observed because within-session learning was analyzed for all sessions together.

Another aspect of tactile feedback applied to moving hand is suppressed or enhanced perception, as described by Voudouris et al. [26] and Juravle and Spence [27]. In our case, the movement of the hand and the sliding element are interconnected, forming a closed control loop. Thus, we hypothesize that, in this case, there is an enhancement of sensitivity as opposed to the suppression of sensitivity during task-irrelevant movement.

Reaction delay (RD) is not significantly affected by the training. We can decompose RD into perception and decision time. The perception of tactile input was estimated at 0.241 s by a method suggested by Kim et al. [22]. Thus, we can expect that the difference of approximately 0.2 s corresponds to decision time. At this point, the participant decides the direction of the reaction. The success rate in the decision is shown in Fig. 9. Even though the decision time did not change, the rate of correct decisions, described by the non-minimum phase response count, increased significantly.

The next parameter for defining the learning effect was workload. Participant number 1 was excluded as an outlier in the case of workload in Task 1. The Participant's performance did not stand out of the general trend, but his self-assessed workload in Task 1 did. The Participant considers this self-assessment as a sign of perfectionism. This raises the question: why does his self-assessed workload in Task 2 indicate only a slight increase in the last three sessions? A possible answer could be found in two aspects. The first is the magnitude of sliding element deflections from the reference element. The first task appears to be easier in guiding participants to static positions. After reaching the position, participants stopped moving and waited for a new target position to be generated. However, haptic guidance sometimes started with full sliding element deflection, which meant high joystick deflection was required. In contrast, the second task required continuous effort, but the sliding element deflections were possible to keep low throughout Task 2. The second aspect is the order of the tasks. Task 1 was always the first, followed by Task 2. Participants started Task 2 after refreshing their skills in Task 1. We assume that this fact explains why the workload was assessed as lower for Task 2, while the AE was measured as lower in Task 1 due to the different difficulty levels of both tasks.

To answer the principal research question of necessary training duration, one needs to focus on courses of TRTP and RD over repeated sessions. While the TRTP and RD did not show significant learning effect, AE is the most important parameter to consider in training duration. Significant improvement was also observed in the self-assessed workload. However, it cannot be measured as exactly as the AE while its assessment is subjective. Considering AE, the learning effect breaks in the 6th session in Task 1 and the 7th session in Task 2. Participants' improvement after these sessions indicates only insignificant progress. This study did not evaluate the effect of interval duration between sessions.

The training interval duration between sessions affects the learning effect in the early stage, according to Wang et al. [14]. The seven sessions required to reach the trained skills from our results are much greater than the three sessions in Wang's study, where the time between sessions affected training performance. Additionally, the peak values in the intervals between sessions shown in Fig. 3 do not correspond to any peaks in TRTP, AE, or workload performance across sessions. The stated interval duration between training sessions should provide sufficient time for consolidation to prevent overtraining. Except in one case, all intervals between sessions included at least one night. However, due to the limited number of sessions and participants, a more detailed analysis of the effects of the intervals between sessions was not possible.

There is one effect that should not be neglected in the learning effect evaluation. Figures 10 – 15 demonstrate individual differences between participants. That led us to an individual training proposal. Individual training should be defined based on the improvement and achievement of the required performance in both tasks. The required performance might be set as values where the observed parameters converged in this experiment. We propose the following target values to complete the training based on the results presented in Tab. I. The proposed values apply only to the used hardware and participants with an average performance. Therefore, the following criterion for stating that the learning process is completed could be evaluated as a percentage of performance

TABLE I
TARGET PARAMETER VALUES AFTER TRAINING

	Task 1	Task 2
AE [% of joystick range]		
	2.4	4.7
Workload [Bedford scale]		
	3	3

TABLE II
PERCENTAGE IMPROVEMENT WITH RESPECT TO FIRST SESSION PERFORMANCE

Session	AE 1 [%]	AE 2 [%]
1	0	0
6	-30.7	-14.2
7	-29.6	-29.1
12	-36.1	-28.8

improvement. Tab. II shows the relative mean improvement in AE for the sixth, seventh and twelfth sessions compared to performance in the first session. We propose a value of 30 % for relative improvement to indicate that training is completed. The value of 30 % for relative improvement corresponds to other research in tactile learning, as previously mentioned by Kass et al. [15] and Wang et al. [14].

Some other limitations of the work might be found in the following aspects. Our experiment involved only the dominant hand, and participants largely used the same joystick handle grasping technique. We assume that different grasping techniques could affect the performance of haptic guidance. However, this effect could potentially be stronger if vibrations were used, as suggested by Harris et al. [33]. Their study pointed to the fact that human tactile learning is topographically distributed. Trained skills in vibration discrimination did not transfer to other fingers; however, skills in pressure and roughness discrimination are transferred to other fingers. The short duration of each session limited our ability to fully analyze the within-session learning effect, as well as the effect of neural afferents. The application of our results to longer-duration guidance tasks may be influenced by this limitation. Therefore, we propose conducting another experiment after the training phase to analyze the effect of haptic perception over time, which would be separated from the learning curve effect. Tactile suppression may also impact participants' performance when focusing on other demanding tasks. Therefore, haptic guidance should be tested in parallel task scenarios and during high-workload pilot situations.

The definition of the learning effect for the tactile guidance method allows for planning a new set of experiments for investigating the effect on flight performance/safety. Future aerospace engineering research should explore crossover research that compares the efficacy of visual and tactile guidance and their interactions. While visual guidance is more effective than tactile guidance in simple tasks, a comparative study of these methods is needed for situations where the visual modality is overloaded, such as in aircraft control during an emergency situation. The next research step involves the implementation of aircraft dynamics. At this stage, haptic guidance should be tested with trained professional pilots to confirm the potential benefits in aircraft control. Moreover, the proposed tactile guidance method could have applications in contexts that require divided attention and the simultaneous use of multiple senses. For example, this method could be used for remote operation of machinery and medical devices, navigation of people with vision impairments, and in the automotive industry.

VI. CONCLUSION

We have presented experimental results that led to the definition of the learning effect for the tactile guidance task. Human participants tested the proposed tactile guidance method in a set of position-targeting and trajectory-following tasks. The participant's performance progress between sessions shows an improving trend, particularly in the first seven sessions. The average error between the actual and target positions and the self-assessed workload are parameters significantly influenced by the training. On the other hand, reaction delay was not significantly influenced by training and time to reach the target position improvement was identified only between the first two sessions.

Achievable performance in tactile guidance has been presented. The guidance accuracy expressed in the average error between the target and actual position is less than 5 % of the joystick range in the trajectory-following tack. This value means a competitive result in comparison to other tactile guidance methods [34], [35]. These results are valid only for the proposed haptic feedback device. However, the presentation of hardware setup and corresponding performance might be useful for designing novel tactile guidance methods. The performance in haptic guidance is individual. In order to apply the results to define an individual training setup, we propose the additional criterion for defining trained skills as an improvement of 30 % over consecutive sessions in the average error between the target and actual positions between the initial and trained skills.

ACKNOWLEDGMENTS

This research has been supported by TACR, project no. TJ01000122 "Haptic feedback of assistant systems for flight safety improvement", by project No. FSI-S-23-8163 funded by The Ministry of Education, Youth and Sport (MEYS, MŠMT in Czech) institutional support and partially by the Research Center for Informatics project CZ.02.1.01/0.0/0.0/16 019/0000765. The authors thank Assoc. Prof. Zdeněk Karpíšek of the Institute of mathematics at Brno University of Technology for his support in the analysis of results and to anonymous reviewers for their conductive comments.

REFERENCES

[1] G. de Rooij, D. Van Baelen, C. Borst, M. M. van Paassen, and M. Mulder, "Supplementing haptic feedback through the visual display of flight envelope boundaries," in *AIAA Scitech 2020 Forum*, 2020, p. 0373.

[2] P. Zikmund, M. Macik, L. Dubnicky, and M. Horpatzska, "Comparison of joystick guidance methods," in *2019 10th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE, 2019, pp. 265–270.

[3] NTSB, "Most wanted list of transportation safety improvements 2017-2018," 2017. [Online]. Available: https://www.ntsb.gov/Advocacy/mwl/Documents/2017-18/MWL-Brochure2017-18.pdf

[4] J. G. Geehan, "Flight testing angle-of-attack warning combinations on part 23 aircraft," 2017.

[5] P. Zikmund, L. Dubnicky, M. Macik, I. Jebacek *et al.*, "Pilot-aircraft haptic feedback tests," *Aircraft Engineering and Aerospace Technology*, vol. 92, 2020. [Online]. Available: https://www.emerald.com/insight/content/doi/10.1108/AEAT-12-2019-0265/full/html

[6] J. Nielsen, "The usability engineering life cycle," *Computer*, vol. 25, no. 3, pp. 12–22, 1992.

[7] E. P. Scilingo, M. Bianchi, G. Grioli, and A. Bicchi, "Rendering softness: Integration of kinesthetic and cutaneous information in a haptic device," *IEEE Transactions on Haptics*, vol. 3, no. 2, pp. 109–118, 2010.

[8] A. Caspo, G. Wersényi, and M. Jeon, "A survey on hardware and software solutions for multimodal wearable assistive devices targeting the visually impaired," *Acta Polytechnica Hungarica*, vol. 13, no. 5, p. 39, 2016.

[9] L. E. R. Langley and C. Bethwaite, "Notes on research into some aspects of stall-warning devices." *College of Aeronautics Report 72*, 1953.

[10] J. P. Trant Jr, "Preliminary investigation of a stick shaker as a lift-margin indicator," Tech. Rep., 1955.

[11] J. B. van Erp, *Tactile displays for navigation and orientation: perception and behaviour*. Utrecht University, 2007.

[12] S. Cardin, F. Vexo, and D. Thalmann, "Vibro-tactile interface for enhancing piloting abilities during long term flight," *Journal of Robotics and Mechatronics*, vol. 18, no. ARTICLE, pp. 381–391, 2006.

[13] K. Fellah and M. Guiatni, "Tactile display design for flight envelope protection and situational awareness," *IEEE transactions on haptics*, vol. 12, no. 1, pp. 87–98, 2018.

[14] W. Wang, J. Yang, Y. Yu, Q. Wu, J. Yu, S. Takahashi, Y. Ejima, and J. Wu, "Tactile angle discriminability improvement: roles of training time intervals and different types of training tasks," *Journal of neurophysiology*, vol. 122, no. 5, pp. 1918–1927, 2019.

[15] A. L. Kaas, V. van de Ven, J. Reithler, and R. Goebel, "Tactile perceptual learning: learning curves and transfer to the contralateral finger," *Experimental brain research*, vol. 224, no. 3, pp. 477–488, 2013.

[16] Karni, Avi and Sagi, Dov, "Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity." *Proceedings of the National Academy of Sciences*, vol. 88, no. 11, pp. 4966–4970, 1991.

[17] Karni, A and Sagi, D, "The time course of learning a visual skill," *Nature*, vol. 365, no. 6443, pp. 250–252, 1993.

[18] M. Atienza, J. L. Cantero, and E. Dominguez-Marin, "The time course of neural changes underlying auditory perceptual learning," *Learning & Memory*, vol. 9, no. 3, pp. 138–150, 2002.

[19] Z. Qu, Y. Song, and Y. Ding, "Erp evidence for distinct mechanisms of fast and slow visual perceptual learning," *Neuropsychologia*, vol. 48, no. 6, pp. 1869–1874, 2010.

[20] K. Molloy, D. R. Moore, E. Sohoglu, and S. Amitay, "Less is more: latent learning is maximized by shorter training sessions in auditory perceptual learning," *PloS one*, vol. 7, no. 5, p. e36929, 2012.

[21] S. Ashley and J. Pearson, "When more equals less: overtraining inhibits perceptual learning owing to lack of wakeful consolidation," *Proceedings of the Royal Society B: Biological Sciences*, vol. 279, no. 1745, pp. 4143–4147, 2012.

[22] J. Kim, E. Francisco, J. Holden, R. Lensch, B. Kirsch, R. Dennis, and M. Tommerdahl, "Visual vs. tactile reaction testing demonstrates problems with online cognitive testing," *The Journal of Science and Medicine*, vol. 2, no. 2, pp. 1–10, 2020.

[23] L. R. Elliott, J. van Erp, E. S. Redden, and M. Duistermaat, "Field-based validation of a tactile navigation device," *IEEE transactions on haptics*, vol. 3, no. 2, pp. 78–87, 2010.

[24] S. De Stigter, M. Mulder, and M. Van Paassen, "Design and evaluation of a haptic flight director," *Journal of guidance, control, and dynamics*, vol. 30, no. 1, pp. 35–46, 2007.

[25] F. M. Nieuwenhuizen and H. H. Bülthoff, *Evaluation of Haptic Shared Control and a Highway-in-the-Sky Display for Personal Aerial Vehicles*, 2014. [Online]. Available: https://arc.aiaa.org/doi/abs/10.2514/6.2014-0808

[26] D. Voudouris and K. Fiehler, "Enhancement and suppression of tactile signals during reaching." *Journal of Experimental Psychology: Human Perception and Performance*, vol. 43, no. 6, p. 1238, 2017.

[27] G. Juravle and C. Spence, "Juggling reveals a decisional component to tactile suppression," *Experimental Brain Research*, vol. 213, no. 1, pp. 87–97, 2011.

[28] K. Chamnongthai, T. Endo, S. Nisar, F. Matsuno, K. Fujimoto, and M. Kosaka, "Fingertip force learning with enhanced haptic sensation using stochastic resonance," in *2019 IEEE World Haptics Conference (WHC)*. IEEE, 2019, pp. 539–544.

[29] S. J. Bensmaia, Y.-Y. Leung, S. S. Hsiao, and K. O. Johnson, "Vibratory adaptation of cutaneous mechanoreceptive afferents," *Journal of neurophysiology*, vol. 94, no. 5, pp. 3023–3036, 2005.

[30] J. Triesch and C. Von Der Malsburg, "Robotic gesture recognition by cue combination," in *Informatik'98: Informatik zwischen Bild und Sprache 28. Jahrestagung der Gesellschaft für Informatik Magdeburg, 21.–25. September 1998*. Springer, 1998, pp. 223–232.

[31] J. L. Sullivan, S. Pandey, M. D. Byrne, and M. K. O'Malley, "Haptic feedback based on movement smoothness improves performance in a perceptual-motor task," *IEEE Transactions on Haptics*, vol. 15, no. 2, pp. 382–391, 2021.

[32] A. H. Roscoe, "Assessing pilot workload in flight," Royal Aircraft Estabilishment Bedford (United Kingdom) Bedford United Kingdom, Tech. Rep., 1984.

[33] J. A. Harris, I. M. Harris, and M. E. Diamond, "The topography of tactile learning in humans," *Journal of Neuroscience*, vol. 21, no. 3, pp. 1056–1061, 2001.

[34] A. A. Stanley and K. J. Kuchenbecker, "Evaluation of tactile feedback methods for wrist rotation guidance," *IEEE Transactions on Haptics*, vol. 5, no. 3, pp. 240–251, 2012.

[35] C. Rognon, V. Ramachandran, A. R. Wu, A. J. Ijspeert, and D. Floreano, "Haptic feedback perception and learning with cable-driven guidance in exosuit teleoperation of a simulated drone," *IEEE transactions on haptics*, vol. 12, no. 3, pp. 375–385, 2019.

Pavel Zikmund received his M.S. and Ph.D. degrees in mechanical engineering from Brno University of Technology (BUT), the Czech Republic in 2006 and 2013. He is currently an assistant professor at the Institute of Aerospace Engineering, BUT. His research interest includes the design and validation of haptic interfaces for pilot-aircraft interaction, human factors in aviation, and flight testing. He is a lecturer of Flight mechanics - Performance and Control & stability courses.

Michaela Horpatzká received the B.S. degree in mechanical engineering from Brno University of Technology (BUT), Czech Republic in 2016 and M.S. degree in aircraft design in 2019 from BUT, Czech Republic. She is currently working toward the Ph.D. degree in Institute of Aerospace Engineering, BUT. Her research interests include human factors, human machine interaction, aircraft systems and reliability.

Miroslav Macík received the M.S. and Ph.D. degrees in Computer Science from the Czech Technical University in Prague in 2009 and 2016. He is an Assistant Professor at the Department of Computer Graphics and Interaction at the Faculty of Electrical Engineering, Czech Technical University in Prague. His research focuses on multimodal interaction, haptic interaction, and support of spatial orientation. Furthermore, his research interests include the development of applications and interaction methods for older adults with vision impairments.