2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

Exploring Haptic Augmentation and Language Design for Smartphone-Based Teleoperation

Zac Baylis, Ziling Chen, Rolando Bautista-Montesano, Yeo Jung Yoon, Thomas Bohné, Sławomir K. Tadeja, and John Liu

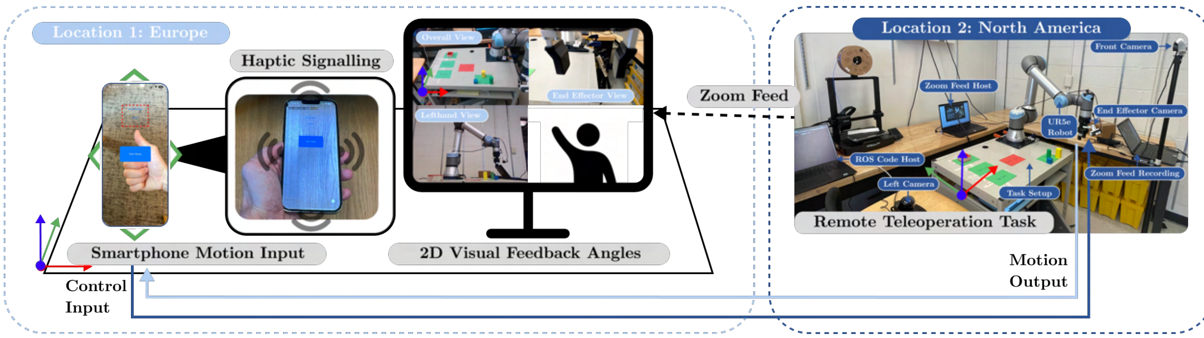

Fig. 1. Our cross-continental teleoperation system enables real-time control of a UR5e robotic arm using an unmodified smartphone for motion input, haptic signalling, and video-based feedback. The system employed a range of haptic cues to communicate five aspects of the physical task: contact, gripper state, alignment, error boundary, and motion initiation. The architecture relies on commodity devices such as smartphones, web cameras, and video streaming, thereby reducing cost and complexity compared to systems requiring specialised hardware, while still supporting precision industrial-style tasks.

***Abstract*— Smartphone-based teleoperation is gaining traction as a versatile remote control solution, using widely available hardware to provide a portable and scalable interface for telerobotics. However, a crucial limitation of such an approach is the lack of effective haptic feedback, which restricts accuracy and increases operator workload. While smartphones offer a low-entry barrier as well as both portability and scalability, current interfaces rely almost exclusively on visual cues. To address this gap, we investigate the use of symbolic haptic feedback delivered through an unmodified mobile device to support remote manipulation tasks. We designed a combined teleoperation task that integrates object sorting and peg-in-hole insertion, embedding five candidate haptic cues (i.e., contact, gripper state, alignment, error boundary, and motion initiation). A within-subjects study with 16 participants compared visual-only and visual-plus-haptic conditions. Results show that haptic augmentation reduced total errors by 42% and significantly lowered perceived workload. Continuous cues for alignment and error boundaries achieved the highest recognition rates of 94% and 81%, respectively, while brief state cues were less reliably interpreted. Post-task interviews highlighted user preference for simple, continuous, and intense signals in visually ambiguous scenarios. Our findings provide new design guidelines for haptic cue prioritisation and encoding strategies.**

I. INTRODUCTION

The widespread digitisation of manufacturing shop floors has driven substantial gains in operational efficiency and adaptability [2]. Simultaneously, the increasing robotisation of production lines has intensified the need for novel approaches to human-robot collaboration (HRC) [3]. Among these, remote teleoperation (i.e., telerobotics) is a critical area, focused on bridging the gap between human operators and distant physical robots. It is a key enabler for addressing tasks beyond the reach of conventional automation, which allows human operators to extend their presence and skills into dangerous or distant environments for critical tasks like remote manipulation, sensing, and maintenance [4].

In this context, mobile devices, such as smartphones, offer accessible, portable, and sensor-rich teleoperation platforms, providing advantages in scalability, cost, and usability [5]. Their usability is further enhanced by the worldwide adoption of mobile devices and thereby also pre-existing and vast familiarity with smartphone-based interaction methods. This baseline familiarity may significantly reduce the potential entry barrier for the operators [6]. Despite these benefits, current smartphone-based teleoperation interfaces typically deliver simplistic or inadequate tactile signals, limiting their effectiveness in accuracy-critical manufacturing scenarios [7]. Specifically, this absence of effective tactile feedback negatively affects the precision and efficiency of teloperated tasks, as well as the operator's cognitive load [8]. While high-fidelity haptic devices are available, they are often costly or impractical for flexible or resource-constrained settings [9].

While the integration of haptic feedback into teleoperation systems holds considerable promise [5], the specific characteristics of tactile information (e.g., contact onset, slippage direction, or force magnitude) that bring the most benefit for encoding with smartphone-available tactile feedback, such as vibrations, remain largely unexplored. Optimizing symbolic haptic cues, often referred to as "*tactons*" (short, symbolic patterns) for smartphone interfaces, offers a novel approach to significantly enhance teleoperation performance

Z. Baylis and T. Bohné are with the University of Cambridge: {zacb2, tmb35}@cam.ac.uk; Z. Chen, R. Bautista-Montesano, Y. Jung Yoon, S.K. Tadeja, and J. Liu are with the Massachusetts Institute of Technology: {zlchen, r_b_m, yoony, sktadeja, johnhliu}@mit.edu. Approval for experimental procedures was granted by the Cambridge Engineering Ethics Committee. This paper is based on the MPhil thesis by Zac Baylis [1].

and improve overall usability [7], [10]. Addressing this gap could help to accelerate the broader adoption of teleoperation in new and complex environments.

Consequently, this paper presents an investigation of haptic signals generated by an unmodified iPhone device to augment smartphone-based teleoperation. Specifically, the contributions of this paper are as follows:

1) A novel, literature-grounded smartphone-based haptics system for industrial telerobotics.
2) In-depth experimental comparison of smartphone-enabled robotic teleoperation with and without vibration-based haptic feedback.
3) Empirically-driven extensions of tactons' design principles to smartphone-based industrial telerobotics.

II. RELATED WORK

To characterise mobile-based teleoperation in manufacturing, we conducted a 20-year (2005-2025) systematic literature review (SLR) following PRISMA guidelines [11]. Searches were run in *Scopus*, *ACM Digital Library*, and *IEEE Xplore*, with manual backfilling via *Google Scholar*. Our SLR included topics such as telerobotics in manufacturing, human–robot control, haptic/tactile feedback, and mobile interfaces. In total, 343 records were retrieved, 18 duplicates were removed, and the remaining 325 were screened at the metadata level. Finally, evaluating the full-text of 62 eligible reports, we included 52 studies.

TABLE I
MOBILE-BASED TELEOPERATION THEMES AND REPRESENTATIVE EXAMPLES IN MANUFACTURING

Application Theme	References
Assistance & Telepresence (15)	[4], [6], [12]–[24]
Operation & Production (15)	[4], [6], [14], [16], [25]–[35]
Maintenance & Repair (12)	[2], [4], [7], [14], [15], [25], [29], [30], [33], [36]–[39]
Handling & Logistics (8)	[16], [22], [28]–[31], [40], [41]

A. Manufacturing Teleoperation Themes

Across the search results, applications of teleoperation within manufacturing fell into four themes: (i) *assistance and telepresence*, (ii) *operation and production*, (iii) *maintenance and repair*, and (iv) *handling and logistics* (see Table I).

Telepresence systems combine audio–visual and, in select cases, tactile channels for remote inspection and guidance [12]. Immersive facilities with virtual reality (VR) or augmented reality (AR) technology overlays frequently enhance situation awareness (e.g., non-smartphone-based systems such as remote robot guidance [13]) and education (e.g., smartphone-based teleoperated labs improving motivation and confidence [6]). However, implementations remain dominated by head-mounted and desktop interfaces, with mobile solutions being in the minority.

Teleoperation supports flexible manufacturing by enabling remote or distributed human-in-the-loop control for assembly, micro-manipulation, welding, and painting [4], [14], [26]. Large initiatives demonstrate work-from-home participation in complex assembly [27]. As above, AR/VR augmentation is common, whereas mobile-based control is comparatively under-represented in production settings.

TABLE II
ADOPTION ISSUES FOR MOBILE-BASED INDUSTRIAL TELEROBOTICS

Challenge	Type	References
Limited tactile feedback	Technical	[7], [10], [39], [49]
Sensor reliability	Technical	[5], [6], [37], [42], [50]
Network reliability	Technical	[22], [29], [37]
Visual feedback dependency	User	[17], [18], [37], [46]
Operator workload	User	[33], [50]
Absence of regulations	Regulatory	[48]

In hazardous or constrained environments, teleoperation mitigates risk and downtime. Examples include AR-assisted disassembly of EV batteries [25] and smartphone-based error recovery [37]. These studies underscore the potential of lightweight devices but also reveal limits tied to sensing, feedback fidelity, and standardisation.

Warehouse teleoperation encompasses the supervision of remote vehicles (e.g., forklifts) and exception handling, including commercial deployments and oversight [40]. Immersive frameworks improve visualisation and precision [28], [31], while smartphone-based control remains less common.

Across themes, the literature emphasises immersive and desktop interfaces, while only a few studies investigate smartphone-based teleoperation despite its ubiquity and cost advantages. Moreover, when smartphones are used, haptic feedback is typically simplistic or absent, leaving open which *tactile events* should be encoded and how to encode them for reliability under cognitive load. This gap motivates our focus on symbolic haptic cue design for smartphone teleoperation.

B. Smartphones as Teleoperation Interfaces

Smartphones combine portability, broad operator familiarity, as well as rich on-board sensing and actuation [5], [42], enabling intuitive gesture-based control and rapid onboarding [37], [43]. Empirical prototypes have demonstrated real-time manipulation and telepresence in manufacturing-adjacent contexts [6], [44], [45]. Alongside these benefits, three constraints classes recur (see Table II): *technical* limits (i.e., sensor reliability in dynamic or occluded scenes; haptic fidelity; network latency) [7], [37]; *user* factors (e.g. visual dependence, depth ambiguity, fatigue) [18], [33], [46], [47]; and *regulatory* gaps (i.e., validation, safety, and interoperability standards) [48]. These constraints converge on the need for principled tactile augmentation to ease visual load and improve precision.

C. Haptic Feedback for Teleoperation

Cutaneous feedback improves situational awareness and can lead to a reduction in force and placement errors in manipulation tasks, particularly when visual depth cues are ambiguous [9], [51], [52]. In smartphone contexts, haptic tactons, offer a practical means to encode salient events (i.e., contact, boundary approach, and alignment) within the constraints of mobile actuators [7], [10]. Prior studies have shown that simple haptic cues can lower response times and error rates despite limited information bandwidth [53]–[55]. However, existing implementations rarely (i) prioritise *which* cues most benefit teleoperation, or (ii) validate *encoding strategies* under realistic cognitive load on smartphones.

We address this gap threefold. First, we derive a functional set of smartphone-compatible cues from the literature. Second, we implement them as symbolic haptic patterns on an iPhone. Finally, we empirically compare visual-only vs. visual-plus-haptics in representative manipulation tasks.

III. HAPTIC ENHANCED TELEOPERATION

A. Identification of Relevant Haptic Domains

A key challenge in smartphone-based teleoperation is the absence of validated haptic cues tailored to precision and safety. To address this, we conducted a second targeted literature review, also following PRISMA [11], which led to the inclusion of 13 directly relevant studies.

Symbolic haptic feedback, previously labelled in literature as "tactons", haptic icons, or cues, uses haptic vibration motors to encode interaction events as abstract tactile signals [10], [56], [57]. Unlike grounded haptic devices, these cues act as lightweight symbolic markers (e.g., object contact, boundary alerts, or orientation guidance). Prior work shows they improve situational awareness and reduce cognitive load despite limited haptic bandwidth [58], [59].

From the literature, four functional roles emerge: (i) *event-specific feedback* (e.g., collisions, grasps) [60], (ii) *contact/force acknowledgement* cues [54], [61], (iii) *alignment/orientation guidance* (e.g., for insertion or assembly) [62], [63], and (iv) *error or safety alerts* (e.g., warning of hazards or boundary violations) [10], [37], [64].

Common encodings include paired pulses for state changes [56], [65], short pulses for contact detection [63], [66], continuous vibrations for alignment [64], and repeated patterns for error alerts [56]. Optimisation strategies such as intensity modulation, temporal patterning, and adaptive cue design improve discriminability [10], [37], [56].

TABLE III

SELECTED HAPTIC CUES, THEIR SIGNAL CHARACTERISTICS, AND ACTIVATION SCENARIOS FOR SMARTPHONE-BASED TELEOPERATION

Domain	Signal Type	Activation Scenario
Contact	Short discrete pulse	Object grasping
Error	Repeated buzz	Zone breached
State	Paired pulse	Gripper opened/closed
Spatial Guidance	Continuous	Correct pose detected
Spatial Guidance	Triple pattern	Movement initiated

On this basis, we selected five cues for empirical testing (see Table III): (i) contact detection, (ii) gripper state, (iii) alignment guidance, (iv) error alerts, and (v) motion initiation. This minimal set strikes a balance between interpretability, cognitive load, and smartphone-imposed constraints.

B. Smartphone-Based Haptic System

We developed our system on top of an iPhone-deployed TeleopLab platform [6]. The base system was reconfigured

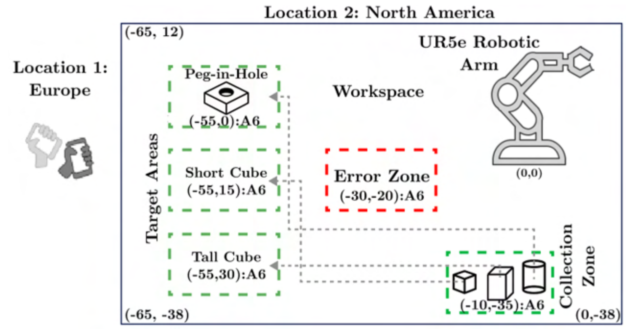

Fig. 3. Task architecture and zone layout (not at scale).

while retaining the overall interaction loop. The remote cell used a UR5e[1] robotic arm with a Robotiq 2F-85 gripper[2] mounted on an 81×56 cm table. The mobile client ran on an unmodified iPhone 13 Pro Max, with visual feedback in the local lab shown on a 15-inch MacBook Air via Zoom. A ROS Noetic (Ubuntu 20.04) stack handled logging, zone evaluation, and the client API.

To balance the field of view and cognitive load, we employed three web cameras, namely, an overview, side, and end-effector perspective (see Fig. 1), mounted on a custom 3D-printed bracket. Joint speeds were capped at 50% of the UR5e's rating, limiting TCP velocity to < 250 mm/s in line with ISO 10218-1 [67] and ISO/TS 15066 [68].

Local haptics were rendered via Apple CoreHaptics using five symbolic vibration patterns. Participants used a portrait-oriented iPhone to control a UR5e robot through direct pose mirroring. ARKit-derived 6-DOF handset poses were mapped to the robot's Cartesian tool center point, with rotational constraints applied to improve stability. Teleoperation was initiated via UDP streaming and on-screen gripper toggles. A ROS node logged state data (pose, collisions, regions) to MySQL and exposed it via a Flask API. The iOS client polled this API at 5 Hz, decoding `RobotPoseDTO` to trigger haptic cues upon entering alignment or error zones. Guidance was gated by a flag to avoid spurious feedback.

IV. USER STUDY

A. Participants

We recruited sixteen participants (7 male, 9 female; age $M = 23.9$, $SD = 1.3$, range 22–27) using an opportunity sampling method. Most reported limited or no prior robotics or teleoperation experience (14/16), though all were familiar with smartphones. With respect to assembly or manufacturing, only 7 reported prior familiarity.

B. Experimental Design and Task

The study adopted a within-subjects design to evaluate the influence of haptic augmentation on smartphone-based teleoperation. The independent variable was the feedback modality with two conditions: (i) *visual-only* baseline, and

[1] Universal Robots, "Robot step file—UR5e/UR7e—e-series," https://www.universal-robots.com [Accessed: 15 August 2025]
[2] Robotiq, https://robotiq.com/ [Accessed: 15 August 2025]

(ii) *visual-plus-haptics* condition in which symbolic haptic cues were delivered through the smartphone interface. This structure enabled direct performance comparisons between modalities while controlling for inter-participant variability.

Visual feedback was delivered via a three-view camera display, while haptic feedback was transmitted via the handset. Continuous vibration-guided alignment during insertion approaches, with rhythmic alerts signalling boundary violations. Discrete tactile cues confirmed grasp or toggle events.

We designed a combined teleoperation to consist of two sequential components: (i) *object sorting* and (ii) *peg-in-hole insertion* (see Fig. 3). The sequence was designed to embed all five symbolic haptic cues into a natural workflow, ensuring that each cue type could be tested in realistic contexts rather than isolated trials.

In the sorting phase, three colour-coded blocks (*blue, green, yellow*) were used to enforce a fixed pick-and-place order and unambiguous mapping between each block and its target location. Colour served only as instructional scaffolding, while the blocks varied in shape and length to introduce manipulation variability and mitigate learning effects. This gross-motor manipulation required repeated grasp–release cycles, systematically triggering *contact detection* cues at the moment of grasp and release, and *gripper state* cues on each open/close action. Sorting reflects common industrial tasks such as bin-picking and palletising, and provides multiple repetitions to assess whether users could reliably interpret discrete event-based cues under load.

The **insertion phase** required participants to guide a block into a narrow pegboard slot. This precision subtask was chosen because insertion is widely regarded as a benchmark for assessing operator awareness and fine motor alignment in teleoperation [69]. Alignment cues were triggered when the gripper entered predefined staging and task zones around the target slot, providing continuous haptic guidance until successful insertion. This allowed for the controlled testing of whether haptic feedback improved spatial awareness beyond that offered by visual information alone.

To probe the effectiveness of **safety-critical alerts**, an A6-sized (105 x 148 mm) error zone was defined at the centre of the workspace, marked physically on the table and implemented virtually in the controller. Entering this region, or moving the end effector off the tabletop, triggered *error cues* that simulated real-world boundary violations. In contrast, alignment zones acted as positive guidance regions, ensuring that participants encountered both supportive and corrective feedback within the same task.

Finally, each trial began with a **motion initiation** cue upon activation of teleoperation. This ensured participants consistently experienced all five cue types within a single, repeatable workflow. By integrating gross and fine manipulation with structured error and alignment zones, the task design directly addressed the study's research questions, i.e., which cues are most recognisable, which encodings are most effective, and whether haptic-enabled feedback improves task performance in realistic manufacturing-style teleoperation.

TABLE IV
HAPTIC ENCODINGS AND TRIGGERS (CONDENSED).

Cue	Encoding (engine; timing) and trigger
Contact	CoreHaptics transient (1.0/1.0) at t=0; +0.8 s after `<Gripper>` close.
Gripper (state)	UIImpact (heavy) ×2 at t=0, 0.18 s on toggle.
Alignment	CoreHaptics continuous (0.7/0.6) up to 120 s; active in A5/A6 if phone downward; 0.2 s tail-off.
Error (boundary)	CoreHaptics transient ×3 at t=0, 0.07, 0.13 s; repeats every 0.5 s in A4 or off-table.
Motion (activ.)	UIImpact (heavy) short–long–short at t=0, 0.09, 0.24 s on *Start Teleop*.

Values in parentheses are intensity/sharpness; A4/A5/A6 are defined zones

C. Experiment Protocol

The experiment took place across the Atlantic Ocean, with participants physically present in Europe and the robotics setup located in North America. Given the intercontinental setting, we conducted a post-experimental latency characterisation to quantify end-to-end system delay. Round-trip control latency, measured via timestamped command–response logging across the UDP pipeline, averaged 300 ms, with a maximum observed delay of 500 ms falling within common ranges for teleoperation [70].

After providing informed consent, participants completed a pre-experiment questionnaire that captured their demographics and prior experience with robotics, teleoperation, and assembly tasks. Next, participants were then familiarised with the smartphone interface, which provided three controls: **Start Teleop** (motion mirroring), **Open/Close Gripper**, and **Reset**. Training involved a simple pick-and-place task, repeated as needed until participants were comfortable, or 10 minutes of training had elapsed. Before continuing, participants completed the *Simulator Sickness Questionnaire* (SSQ) [71] to check for pre-existing discomfort (e.g., nausea, oculomotor disturbance). Only those reporting no significant symptoms proceeded.

The main study comprised the two counterbalanced conditions: visual-only (control) and visual-plus-haptic. Each participant completed the full teleoperation task once under each condition. During each trial, we captured task completion time, defined as total duration from task initiation to successful completion, further decomposed into three sub-blocks to capture performance patterns [72]. We also recorded the error rate defined as the count of error-zone entries (validated via pose data), object misalignments or drops, and unintended collisions (recorded manually).

After each trial, participants completed the *NASA Task Load Index* (NASA-TLX) to track perceived cognitive taskload [73], *System Usability Scale (SUS)* to capture overall usability rating of the interface [74], and *Flow Short Scale (FSS)* to measure perceived flow (deep involvement, concentration and focus on a task) and anxiety during task execution [75]. The experiment concluded with a semi-structured interview and a cue recognition test guided by the *Technology Acceptance Model* (TAM) [76]) covering perceived usefulness, ease of use, and helpfulness of cues.

Finally, a *cue recognition test* assessed interpretability: participants were presented with vibration patterns and asked

TABLE V

DESCRIPTIVE RESULTS OF OBJECTIVE PERFORMANCE MEASURES AND CUE RECOGNITION BY FEEDBACK CONDITION

Performance Measures							
Measure	**Condition**	**N**	**Min**	**Max**	**M**	**SE**	**95%-CI**
Total Errors	Visual-only	16	2	12	5.94	0.82	[4.33, 7.55]
	Visual-plus-haptic	16	0	8	3.44	0.65	[2.16, 4.71]
Task completion time (min)	Visual-only	16	6.05	22.97	10.58	1.21	[8.22, 12.94]
	Visual-plus-haptic	16	5.12	17.37	9.28	0.99	[7.34, 11.21]

Cue Recognition Test					
Measure	**Cue**	**N**	**Accuracy (%)**	**SE**	**95%-CI**
Cue Recognition	Contact (detection)	16	25.00	11.18	[10.18, 49.50]
	Gripper (state)	16	18.75	10.08	[6.59, 43.01]
	Alignment	16	93.75	6.25	[71.67, 98.89]
	Error (boundary)	16	81.25	10.08	[56.99, 93.41]
	Motion (activation)	16	62.50	12.50	[38.64, 81.52]

to identify the corresponding event (e.g., contact, gripper closed), enabling analysis of recognisability and confusion.

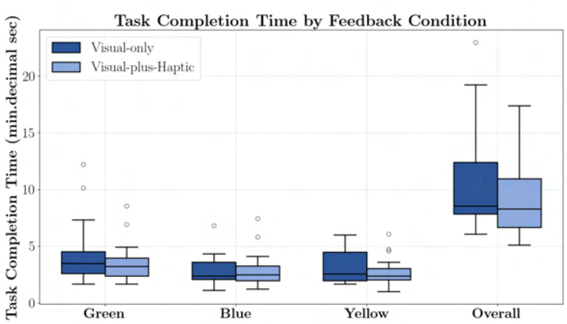

Fig. 4. Task completion time by feedback condition (circle = an outlier).

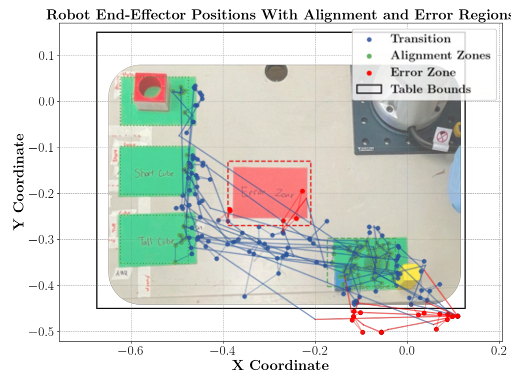

Fig. 5. Robot end effector pose plot example overlaid onto an image of the task setup, recorded during visual-only conditions.

V. USER STUDY RESULTS

Statistical significance was evaluated at a conservative level of $\alpha = 0.05$. A z-score outlier screen ($|z| < 3$) identified no outliers. Descriptive summaries for the objective measures and cue recognition are provided in Table V.

A. Objective Performance

The Shapiro-Wilk and Levene's tests of the natural-log transformed values revealed normality and homogeneity. A paired t-test showed no significant difference between conditions, $t(15) = 1.85$, $p = 0.082$, though visual-plus-haptics trended faster by $\approx 7.8\,\text{s}$ on average. A two-way repeated-measures ANOVA with factors *condition* and *block* (*blue cube, green cylinder, yellow cuboid*) confirmed a main effect of block, $F(2, 30) = 4.80$, $p = 0.016$, with *green* slowest and *blue* fastest, but no condition or interaction effects (see Fig. 4). Here, the block reflects the sequential execution of three distinct block–target pairings included to capture within-task difficulty and learning effects.

Errors comprised collisions, zone infringements, and misalignments (see Fig. 5). Counts were modelled with a log–Poisson general linear model (GLM) after verifying Poisson adequacy at the condition level using Kolmogorov–Smirnov tests. Haptic augmentation significantly reduced total errors ($\chi^2(1) = 10.41$, $p = 0.001$, $\exp(B) = 0.579$) corresponding to an estimated $\approx 42\%$ reduction relative to visual-only (see Table V). A repeated-measures ANOVA (condition $\times$ error type) corroborated main effects of condition ($F(1, 15) = 8.52$, $p = 0.011$) and error type ($F(2, 30) = 13.55$, $p < 0.001$), without interaction, indicating a consistent reduction across error categories.

Recognition was highest for *alignment* (93.75%) and *error* (81.25%), moderate for *motion* (62.50%), and near chance for *contact* (25.00%) and *gripper* (18.75%) (see Table V). Binomial tests against a 20% chance level (five-option identification) confirmed above-chance recognition for *alignment*, *error*, and *motion* ($p < 0.001$), but not for *contact* ($p = 0.402$) or *gripper* ($p = 0.648$). A logistic regression indicated that *gripper* cues were significantly less recognisable ($\beta = -4.17$, $p = 0.001$), consistent with participants' reports.

B. Subjective Evaluations

Weighted NASA-TLX scores were significantly lower with haptics ($(t(15) = 2.93$, $p = 0.010$ Cohen's $d = 0.73$) (see Fig. 6). However, in light of recent NASA-TLX critiques, we advise caution when interpreting the overall weighted score [77]. Median reductions were most evident for mental demand, effort, and frustration, suggesting a practically meaningful decrease in workload with haptic-enabled feedback and a more consistent user experience compared to the high-variance workload perceptions of the non-haptic condition.

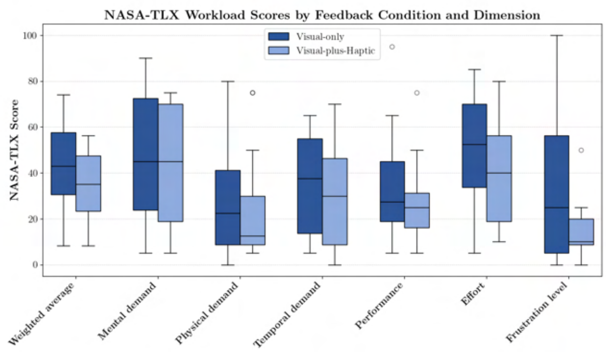

Fig. 6. TLX scores by feedback condition (0 = low demand, 100 = high demand; for performance, 0 = perfect and 100 = failure; circle = outlier)

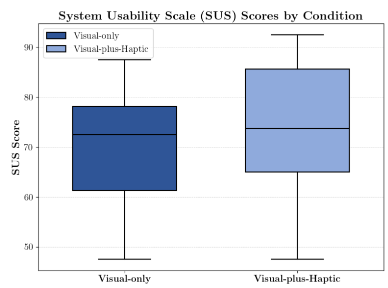

Fig. 7. SUS scores by feedback condition (higher score is better).

Received usability was acceptable under both conditions (visual-only $M = 70.16$, haptics $M = 73.59$). However, the mean difference was not significant ($t(15) = -1.15$, $p = 0.270$) (see Fig. 7).

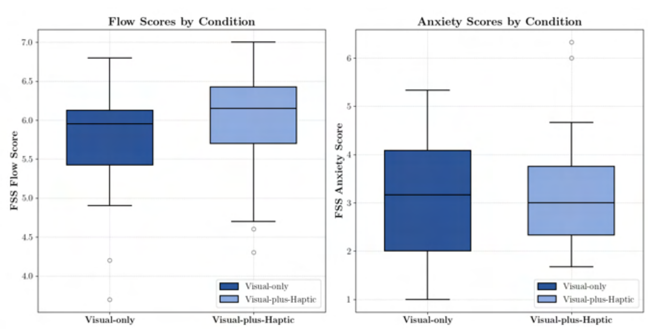

Fig. 8. FSS scores by feedback condition, separated into Flow (left) and Anxiety (right) dimensions (higher flow and lower anxiety are better).

Flow trended higher with haptics ($M = 5.94$ vs. $M = 5.68$) with a moderate effect size, but did not reach significance, $t(15) = -2.03$, $p = 0.060$. Also, anxiety did not differ between the two conditions ($M = 3.29$ vs. $M = 3.19$), $t(15) = -0.31$, $p = 0.758$ (see Fig. 8).

C. Qualitative Feedback

Interview data converged on three themes that contextualise the quantitative findings. First, participants described clear benefits of haptics. Here, 12/16 preferred the haptic-enabled condition, citing improved spatial awareness and more reliable boundary detection, particularly when visual depth cues were ambiguous. Second, cue interpretability varied systematically, specifically, alignment and boundary-error feedback were deemed intuitive and continuous, whereas contact and gripper cues were often ambiguous or confusable, mirroring the recognition results. Third, usability considerations centred on depth perception, camera distortion, and end-effector orientation. Several participants noted small on-screen controls and edge-adjacent placements that contributed to miss-clicks and inflated error counts. Despite these issues, the interface's overall simplicity was valued.

VI. DISCUSSION

A. Implications for Smartphone-Based Teleoperation

The study results demonstrate a clear hierarchy of cue usefulness and interpretability. The *alignment* and *boundary–error* feedback achieved the highest recognition rates (93.75% and 81.25%, respectively), *motion initiation* was moderate (62.50%), and brief, discrete state cues, i.e., *contact, gripper* were near chance (25.00%, and 18.75%, respectively) as shown in Table V. Based on these results, we arrived at three design guidelines offered below.

Interviews converged on the same pattern, i.e., continuous guidance was described as "noticeable" and "confidence building", while rhythmic error alerts were "clear" and "hard to miss", whereas single taps were frequently overlooked or confused. Thereby, what the cue communicates (alignment, boundary) and how it is encoded (continuous or rhythmic versus brief transients) jointly determine effectiveness. This extends previously established tacton design guidance [10] to the smartphone teleoperation with limited actuation bandwidth, salience through temporal extent or repetition is more reliable than isolated pulses for semantically rich events.

Qualitative comments attribute the accuracy (estimated $\approx 42\%$ error reduction) and lower workload benefits to clearer boundary awareness and easier fine alignment under visually ambiguous viewpoints. This matches prior findings that cutaneous cues can reduce slips and improve certainty of state [7], [9] while not necessarily accelerating task execution in repetitive regimes [33].

Three design guidelines follow directly from our study results. First, *prioritise alignment and boundary cues* as primary channels on phones as they carried the largest perceptual and behavioural payoff in our task, and they generalise to many industrial workflows. Second, *encode high-stakes events with continuous or rhythmic tactons*. If state cues are required, increase separability with paired or patterned pulses rather than single transients, and consider intensity and tempo modulation to communicate proximity or severity, consistent with transformational tacton principles [10]. Third, *treating haptics as a remedy for visual ambiguity* is not a replacement for interface ergonomics. The participants highlighted depth perception, orientation management, and small or edge-adjacent buttons as usability impediments. Therefore, while haptic confirmation is helpful, the camera layout and user interface must also be carefully designed.

These findings bridge two strands of work, i.e., teleoperation safety and accuracy with haptic augmentation [7],

[9] and tacton design for mobile and assistive contexts [10]. Unlike grounded kinesthetic devices, smartphones constrain force bandwidth. Nonetheless, by selecting *which* events to represent (e.g., alignment, boundary) and *how* to render them (e.g., continuous, rhythmic), we observe substantial accuracy and workload benefits on commodity hardware.

B. Limitations and Future Work

Generalisability is constrained by the participant profile, laboratory setting, and task duration. Real production environments introduce shift-length fatigue, distractions, and richer error taxonomies. Results also reflect a single smartphone actuator, while device-to-device variance may shift perceptual thresholds. Future work should therefore recruit professional operators, extend trial horizons, and test audio–vibro fusion under variable latency.

VII. CONCLUSION

In this work, we examined whether, and how, symbolic haptics on unmodified smartphones can enhance teleoperation in manufacturing-relevant tasks. Across objective, subjective, and qualitative measures, haptic-enabled feedback produced a robust reduction in errors (about 42%) and a significant decrease in perceived workload. At the same time, task completion time and usability remained comparable to those of the visual-only control. The pattern of cue recognisability and interview reports converged on a simple rule: prioritise *what* matters (e.g., alignment, boundary) and render it with *encodings* that are continuous or rhythmic rather than brief and discrete.

Practically, these findings yield actionable guidance for industrial-related deployments such as implementing alignment and boundary tactons as first-class signals; encoding state changes with paired or patterned pulses if needed; exploiting intensity and tempo to convey confidence or proximity; and addressing visual ambiguity and interface ergonomics alongside haptics. Our results also extend tacton design principles to the smartphone teleoperation context, showing that even with limited actuator bandwidth, meaningful safety and quality gains are achievable on commodity smartphone devices.

Future studies should validate these findings with experienced operators in production settings over longer horizons, systematically optimise encodings for state and contact events, integrate depth- and orientation-aware assistance, and explore multi-sensory fusion under realistic network conditions. Taken together, these steps will refine a principled "haptic language" for mobile teleoperation and accelerate its adoption in smart manufacturing.

REFERENCES

[1] Z. Baylis, "Exploring Augmentations and Haptic Language Design for Smartphone Interfaces," Master's thesis, Cambridge University, 2025.

[2] S. Grabowska, "Smart factories in the age of Industry 4.0," *Management Systems in Production Engineering*, pp. 90–96, 2020.

[3] E. Matheson, R. Minto, E. G. G. Zampieri, M. Faccio, and G. Rosati, "Human–robot collaboration in manufacturing applications: A review," *Robotics*, vol. 8, no. 4, 2019.

[4] F. Chen, B. Gao, M. Selvaggio, Z. Li, and D. Caldwell, "A framework of teleoperated and stereo vision guided mobile manipulation for industrial automation," in *Proc. IEEE Int. Conf. Mechatronics Autom.*, pp. 1641–1648, 2016.

[5] D. Rodriguez, C. Perez, M. Jagersand, and P. Figueroa, "A comparison of smartphone interfaces for teleoperation of robot arms," in *Proc. XLIII Latin Amer. Comput. Conf. (CLEI)*, pp. 1–8, 2017.

[6] Z. Chen, Z. Zhao, and J. Liu, "Application of phone-based robotic arm teleoperation in remote hands-on labs for engineering education," in *ASEE Annual Conference & Exposition*, June 2025.

[7] C. Pacchierotti, A. Tirmizi, G. Bianchini, and D. Prattichizzo, "Enhancing the performance of passive teleoperation systems via cutaneous feedback," *IEEE Trans. Haptics*, pp. 397–409, 2015.

[8] J. G. Wildenbeest, D. A. Abbink, C. J. Heemskerk, F. C. T. van der Helm, and H. Boessenkool, "The impact of haptic feedback quality on the performance of teleoperated assembly tasks," *IEEE Trans. Haptics*, vol. 6, no. 2, pp. 242–252, 2013.

[9] M. Tavakoli, R. V. Patel, M. Moallem, and A. Aziminejad, *Haptics for Teleoperated Surgical Robotic Systems*. World Scientific, 2008.

[10] S. A. Brewster and L. M. Brown, "Tactons: structured tactile messages for non-visual information display," in *Proc. Conf. Australasian User Interface (AUIC)*, pp. 15–23, 2004.

[11] A. Liberati, D. G. Altman, *et al.*, "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration," *BMJ*, 2009.

[12] X. Shen and S. Shirmohammadi, "Telepresence," in *Encyclopedia of Multimedia* (B. Furht, ed.), Boston, MA, USA: Springer, 2008.

[13] L. Peppoloni, F. Brizzi, *et al.*, "Augmented reality-aided telepresence system for robot manipulation in industrial manufacturing," in *21st ACM Symp. Virtual Reality Softw. Technol*, pp. 237–240, 2015.

[14] T. Luu, Q. Nguyen, T. Tran, *et al.*, "Effectiveness of IoT and VR for real-time teleoperation of industrial robots," Dec. 2023.

[15] D. Aschenbrenner, M. E. Latoschik, and K. Schilling, "Industrial maintenance with augmented reality: two case studies," in *22nd ACM Symp. Virtual Reality Softw. Technol. (VRST)*, pp. 341–342, 2016.

[16] Y. Wu, P. Balatti, M. Lorenzini, F. Zhao, W. Kim, and A. Ajoudani, "A teleoperation interface for loco-manipulation control of mobile collaborative robotic assistant," *IEEE Robotics and Automation Letters*, vol. 4, pp. 3593–3600, Oct. 2019.

[17] A. Hong, H. H. Bülthoff, and H. I. Son, "A visual and force feedback for multi-robot teleoperation in outdoor environments: a preliminary result," in *IEEE Int. Conf. Robot. Autom.*, pp. 1471–1478, 2013.

[18] J. Fu, A. Rota, S. Li, J. Zhao, Q. Liu, E. Iovene, G. Ferrigno, and E. D. Momi, "Recent advancements in augmented reality for robotic applications: a survey," *Actuators*, p. 323, 2023.

[19] A. Aly, S. Griffiths, V. Nitsch, K. Pastra, and T. Taniguchi, "Towards intelligent social robots: from naive robots to robot sapiens," in *IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, (Spain), Oct. 2018.

[20] Y. Wu, P. Balatti, M. Lorenzini, F. Zhao, W. Kim, and A. Ajoudani, "A teleoperation interface for loco-manipulation control of mobile collaborative robotic assistant," *IEEE Robotics and Automation Letters*, vol. 4, pp. 3593–3600, Oct. 2019.

[21] J. A. Frank and V. Kapila, "Towards teleoperation-based interactive learning of robot kinematics using a mobile augmented reality interface on a tablet," in *Indian Control Conf.*, pp. 385–392, 2016.

[22] H. Cho, H. Kim, *et al.*, "Semi-autonomous haptic teleoperation of multiple omni-directional mobile robots," in *2013 13th Int. Conf. Control Autom. Syst. (ICCAS)*, pp. 320–324, 2013.

[23] C. Stellmacher, F. Mathis, Y. Weiss, M. B. Loerakker, N. Wagener, and J. Schöning, "Exploring mobile devices as haptic interfaces for mixed reality," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2024.

[24] A. Farooq, H. Z. Tan, A. Weill-Duflos, J. R. Cooperstock, and R. Raisamo, "Embedded haptic waveguides to improve tactile feedback: designing a custom 3D-printed surface to enhance signal mediation," in *IEEE Sensors 2020*, pp. 1–4, 2020.

[25] F. Zhao, W. Deng, and D. T. Pham, "A robotic teleoperation system with integrated augmented reality and digital twin technologies for disassembling end-of-life batteries," *Batteries*, vol. 10, p. 382, 2024.

[26] J. Schilp, M. Ehrenstrasser, S. Clarke, B. Petzold, and M. F. Zaeh, "Smart sensor application in teleoperated micro-assembly systems," in *Proc. SPIE Intell. Manuf.*, pp. 38–49, 2004.

[27] H. Evarts, "NSF funds a robot teleoperation system that could transform the manufacturing workforce." Columbia Eng. News, Dec. 2020.

[28] N. Chen, K. P. Tee, C. M. Chew, and R. Yan, "Intuitive interaction for

robotic grasping," in *Proc. Workshop at SIGGRAPH Asia: Workshop on Sketches of a Robot's Hand (WASA)*, pp. 105–111, 2012.
[29] F. A. Chicaiza, E. Slawiñski, and V. Mut, "Dual-arm mobile manipulator teleoperation with coupled rate-position mapping under time-varying delays," *IEEE Access*, pp. 103463–103487, 2025.
[30] J. A. Marvel, R. Bostelman, and J. Falco, "Multi-robot assembly strategies and metrics," *ACM Comput. Surv.*, pp. 1–32, 2018.
[31] W. Xu, M. Klein, T. Huf, *et al.*, "Virtual reality-based human-Robot interaction for remote pick-and-place tasks," in *Companion Proc. 2024 ACM/IEEE Int. Conf. Human-Robot Interact.*, p. 1148, 2024.
[32] H. Xu, M. Chen, G. Li, L. Wei, S. Peng, H. Xu, and Q. Li, "An immersive virtual reality bimanual telerobotic system with haptic feedback." arXiv:2501.00822, 2025.
[33] D. Whitney, E. Rosen, E. Phillips, G. Konidaris, and S. Tellex, "Comparing robot grasping teleoperation across desktop and virtual reality with ROS reality," in *Robotics Research (Springer Proc. Adv. Robot.)* (N. Amato, G. Hager, S. Thomas, and M. Torres-Torriti, eds.), (Cham, Switzerland), pp. 335–350, Springer, 2020.
[34] F. Liu, A. Lelevé, D. Eberard, and T. Redarce, "A dual-user teleoperation system with Online Authority Adjustment for haptic training," in *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 1168–1171, 2015.
[35] V. Nitsch and B. Färber, "A meta-analysis of the effects of haptic interfaces on task performance with teleoperation systems," *IEEE Trans. Haptics*, pp. 387–398, 2013.
[36] V. Zhidchenko, E. Startcev, and H. Handroos, "Blindfolded operation as a method of haptic feedback design for mobile machinery," in *Proc. EuroHaptics 2024*, pp. 323–337, 2025.
[37] A. Werner and W. Melek, "5G virtual reality manipulator teleoperation using a mobile phone." arXiv:2405.07128, 2024.
[38] M. Esfandiari, J. W. Kim, B. Zhao, G. Amirkhani, *et al.*, "Cooperative vs. teleoperation control of the steady hand eye robot with adaptive sclera force control: a comparative study," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, pp. 8209–8215, 2024.
[39] T. Takami, I. Mizoguchi, and H. Kajimoto, "TelEdge: haptic telecommunication of a smartphone by electro-tactile stimulation through the edges." arXiv:2411.05154, 2024.
[40] Symbotic, "Teleoperations: remote controlled assistance for robot efficiency." `https://www.symbotic.com/`, Apr. 2025.
[41] J. Louca, K. Eder, J. Vrublevskis, and A. Tzemanaki, "Impact of haptic feedback in high latency teleoperation for space applications," *ACM Trans. Hum. Robot Interact. (THRI)*, vol. 13, pp. 1–21, Mar. 2024.
[42] A. M. Walker and D. P. Miller, "Tele-operated robot control using attitude-aware smartphones," in *Proc. 7th ACM/IEEE Int. Conf. Human–Robot Interact. (HRI)*, pp. 269–270, 2012.
[43] A. J. Hung, J. Chen, A. Shah, and I. S. Gill, "Telementoring and telesurgery for minimally invasive procedures," *Journal of Urology*, vol. 199, no. 2, pp. 355–369, 2018.
[44] A. E. Ainasoja, S. Pertuz, and J.-K. Kämäräinen, "Smartphone teleoperation for self-balancing telepresence robots," in *Proc. SciTePress Int. Conf. – Robotics*, pp. 74–81, 2019.
[45] L. Jibawi, S. Said, K. Tadakuma, and J. Berengueres, "Smartphone-based home robotics," *CoRR*, vol. abs/1803.02122, 2018.
[46] J. Jankowski and A. Grabowski, "Usability evaluation of VR interface for mobile robot teleoperation," *International Journal of Human–Computer Interaction*, vol. 31, no. 12, pp. 882–889, 2015.
[47] S. Lu, M. Y. Zhang, T. Ersal, and X. J. Yang, "Effects of a delay compensation aid on teleoperation of unmanned ground vehicles," in *Companion Proc. ACM/IEEE Int. Conf. Human-Robot Interact.*, pp. 179–180, 2018.
[48] S. Lu, R. Zhong, and W. Shi, "Teleoperation technologies for enhancing connected and autonomous vehicles," in *Proc. IEEE Int. Conf. Mobile Ad Hoc Smart Syst. (MASS)*, pp. 435–443, 2022.
[49] D. Leonardis, M. Gabardi, S. Marcheschi, M. Barsotti, F. Porcini, D. Chiaradia, and A. Frisoli, "Hand teleoperation with combined kinaesthetic and tactile feedback: A full upper limb exoskeleton interface enhanced by tactile linear actuators," *Robotics*, vol. 13, no. 8, 2024.
[50] D. B. Priya and M. Subramaniyam, "Fatigue due to smartphone use? Investigating research trends and methods for analysing fatigue caused by extensive smartphone usage: a review," *Work*, vol. 72, no. 2, pp. 637–650, 2022.
[51] S. Choi and K. J. Kuchenbecker, "Vibrotactile display: perception, technology, and applications," *Proc. IEEE*, pp. 2093–2104, Sept. 2013.
[52] L. Marchal-Crespo, M. van Raai, G. Rauter, P. Wolf, and R. Riener, "The effect of haptic guidance and visual feedback on learning a complex tennis task," *Exp. Brain Res.*, pp. 277–291, 2013.
[53] S. Buratti, D. Deiana, A. Noccaro, *et al.*, "Effect of vibrotactile feedback on the control of the interaction force of a supernumerary robotic arm," *Machines*, p. 1085, 2023.
[54] P. Galambos, "Vibrotactile Feedback for Haptics and Telemanipulation: Survey, Concept and Experiment," *Acta Polytech. Hung.*, 2012.
[55] Y. Gong, H. M. Husin, E. Erol, V. Ortenzi, and K. J. Kuchenbecker, "AiroTouch: enhancing telerobotic assembly through naturalistic haptic feedback of tool vibrations," *Front. Robot. AI*, 2024.
[56] E. Hoggan and S. A. Brewster, "Designing audio and tactile crossmodal icons for mobile devices," in *Proc. 9th Int. Conf. Multimodal Interfaces (ICMI)*, pp. 162–169, 2007.
[57] H. Qian, R. Kuber, and A. Sears, "Towards identifying distinguishable tactons for use with mobile devices," in *11th Int. ACM SIGACCESS Conf. Computers and Accessibility*, pp. 257–258, 2009.
[58] C. Bermejo and P. Hui, "A survey on haptic technologies for mobile augmented reality," *ACM Comput. Surv.*, pp. 1–35, 2022.
[59] M. Kurzweg, Y. Weiß, M. O. Ernst, A. Schmidt, and K. Wolf, "A survey on haptic feedback through sensory illusions in interactive systems," *ACM Comput. Surv.*, pp. 1–39, 2024.
[60] R. Sakuma, K. Narumi, Y. Kawahara, and T. Hiraki, "TactPrint: 3D printing lattice-based tactile displays with optimized and local vibration," in *ACM CHI Extended Abstracts*, pp. 1–4, 2024.
[61] K. Y. Choi, N. ElHaouij, J. Lee, R. W. Picard, and H. Ishii, "Design and evaluation of a clippable and personalizable pneumatic-haptic feedback device for breathing guidance," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, pp. 1–36, 2022.
[62] D. Trinitatova and D. Tsetserukou, "Study of the effectiveness of a wearable haptic interface with cutaneous and vibrotactile feedback for VR-based teleoperation," *IEEE Transactions on Haptics*, vol. 16, pp. 463–469, Oct.–Dec. 2023.
[63] E. Pescara, A. Stubenbord, T. Röddiger, L. Fang, and M. Beigl, "Where should I look? Comparing reference frames for spatial tactile cues," in *Int. Symp. Wearable Comput.*, pp. 68–72, 2021.
[64] B. Marques, S. Silva, P. Dias, and B. Santos, "Which notification is better? Comparing visual, audio and tactile cues for asynchronous mixed reality remote collaboration: a user study," in *21st Int. Conf. Mobile Ubiquitous Multimedia*, pp. 276–278, 2022.
[65] K. E. MacLean, "Designing with haptic feedback," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, pp. 783–788, 2000.
[66] C. M. Reed, H. Z. Tan, and L. A. Jones, "Haptic communication of language," *IEEE Trans. Haptics*, pp. 134–153, 2023.
[67] International Organization for Standardization, "ISO 10218-1:2025, Robotics—Safety requirements for industrial robots," 2025.
[68] International Organization for Standardization, "ISO/TS 15066:2016, Robots and robotic devices—Collaborative robots," 2016.
[69] A. Sivakumar, K. Shaw, D. Pathak, *et al.*, "Robotic telekinesis: learning a robotic hand imitator by watching humans on YouTube," in *Proc. Robotics: Science and Systems (RSS)*, 2022.
[70] I. T. Feldstein and S. R. Ellis, "A simple video-based technique for measuring latency in virtual reality or teleoperation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 9, pp. 3611–3625, 2021.
[71] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness," *Int. J. Aviat. Psychol.*, pp. 203–220, 1993.
[72] B. Rummel, "About time: A practitioner's guide to task completion time analysis," *J. Usability Studies*, pp. 124–134, 2020.
[73] S. Hart and L. Staveland, "Development of NASA-TLX (Task Load Index): results of empirical and theoretical research," in *Human Mental Workload*, Adv. Psych., pp. 139–183, North Holland, 1988.
[74] J. Brooke, "SUS: A "quick and dirty" usability scale," in *Usability Evaluation in Industry* (P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland, eds.), pp. 189–194, UK: Taylor & Francis, 1996.
[75] F. Rheinberg, R. Vollmeyer, S. Engster, and D. Sreeramoju, "Fss - flow short scale (english version)," 09 2023.
[76] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, 1989.
[77] M. L. Bolton, E. Biltekoff, *et al.*, "The mathematical meaninglessness of the NASA task load index: a level of measurement analysis," *IEEE Trans. Hum. Mach. Syst.*, vol. 53, no. 3, pp. 590–599, 2023.