

## Article

# Operator Expertise in Bilateral Teleoperation: Performance, Manipulation, and Gaze Metrics <sup>†</sup>

Harun Tugal <sup>1,\*</sup>, Ihsan Tugal <sup>2</sup>, Fumiaki Abe <sup>3,‡</sup>, Masaki Sakamoto <sup>3,‡</sup>, Shu Shirai <sup>3,‡</sup>, Ipek Caliskanelli <sup>1</sup> and Robert Skilton <sup>1</sup>

<sup>1</sup> UK Atomic Energy Authority (UKAEA), Remote Applications in Challenging Environments (RACE), Culham Campus, Abingdon OX14 3DB, Oxfordshire, UK; ipek.caliskanelli@ukaea.uk (I.C.); robert.skilton@ukaea.uk (R.S.)

<sup>2</sup> Software Engineering Department, Mus Alparslan University, 49250 Mus, Turkey; i.tugal@alparslan.edu.tr

<sup>3</sup> Tokyo Electric Power Company (TEPCO), 1-3 Uchisaiwai-cho 1-chome, Chiyoda-ku, Tokyo 100-8560, Japan; abe.fumiaki@jaea.go.jp (F.A.); sakamoto-masaki@tepsys.co.jp (M.S.); shirai.shu@tepcoco.jp (S.S.)

\* Correspondence: harun.tugal@ukaea.uk

<sup>†</sup> This paper is an extended version of our paper published in Tugal, H.; Abe, F.; Sakamoto, M.; Shirai, S.; Caliskanelli, I.; Skilton, R. Factors Influencing Operator Expertise in Bilateral Telerobotic Operations: A User Study. In Proceedings of the IEEE 18th International Conference on Control, Automation, Robotics and Vision (ICARCV), Dubai, United Arab Emirates, 12–15 December 2024; pp. 697–703.

<sup>‡</sup> Secondees at RACE/UKAEA, Culham Campus.

**Abstract:** This paper presents a comprehensive user study aimed at assessing and differentiating operator expertise within bilateral teleoperation systems. The primary objective is to identify key performance metrics that effectively distinguish novice from expert users. Unlike prior approaches that focus primarily on psychological evaluations, this study emphasizes direct performance analysis across a range of telerobotic tasks. Ten participants (six novices and four experts) were assessed based on task completion time and difficulty, error rates, manipulator motion characteristics, gaze behaviour, and subjective feedback via questionnaires. The results show that experienced operators outperformed novices by completing tasks faster, making fewer errors, and demonstrating smoother manipulator control, as reflected by reduced jerks and higher spatial precision. Also, experts maintained consistent performance even as task complexity increased, whereas novices experienced a sharp decline, particularly at higher difficulty levels. Questionnaire responses further revealed that novices experienced higher mental and physical demands, especially in unfamiliar tasks, while experts demonstrated higher concentration and arousal levels. Additionally, the study introduces gaze transition entropy (GTE) and stationary gaze entropy (SGE) metrics to quantify visual attention strategies, with experts exhibiting more focused, goal-oriented gaze patterns, while novices showed more erratic and inefficient behaviour. These findings highlight both quantitative and qualitative measures as critical for evaluating operator performance and informing future teleoperation training programs.

**Keywords:** bilateral teleoperation; human–robot interaction; gaze tracking; operator performance measurements



Academic Editors: Konstantinos E. Psannis, Sotirios Kontogiannis and George Kokkonis

Received: 2 April 2025

Revised: 30 April 2025

Accepted: 6 May 2025

Published: 9 May 2025

**Citation:** Tugal, H.; Tugal, I.; Abe, F.; Sakamoto, M.; Shirai, S.; Caliskanelli, I.; Skilton, R. Operator Expertise in Bilateral Teleoperation: Performance, Manipulation, and Gaze Metrics. *Electronics* **2025**, *14*, 1923. <https://doi.org/10.3390/electronics14101923>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Telerobotic systems enable humans to manipulate objects remotely, with bilateral systems enhancing precision and control through force feedback, making them essential for tasks requiring dexterity and safety. These systems are widely used in industries such as nuclear decommissioning, subsea exploration, defense, and robotic surgery [1–3]. While

the performance evaluation of telerobotic systems often relies on both quantitative and qualitative metrics, user studies play a particularly crucial role due to the human-in-the-loop nature of these systems [4].

In domains like robotic surgery, expertise assessment has been extensively studied, with standardized automated performance metrics (APMs) derived from kinematic data, system events, and instrument grip forces, providing a structured framework for evaluating skill levels [5,6]. However, for bilateral telerobotic systems in other safety-critical fields—such as nuclear decommissioning—there is no universally accepted set of quantitative benchmarks for defining operator expertise. For example, becoming a fully qualified remote handling operator for the Joint European Torus (JET) can take up to two years [7], yet the criteria for assessing proficiency remain largely experience-based rather than systematically quantified. This motivates the need for research into advanced operator classification strategies, particularly those that can support both training validation and human–robot collaboration in high-stakes environments.

In addition to skill classification, device sensitivity is particularly relevant in precision fields such as medicine and micro-assembly. High-sensitivity systems—those that can accurately capture and respond to subtle human inputs—enable safer and more effective telemanipulation in tasks where fine motor control is critical. As systems become more sophisticated, aligning operator expertise with robot responsiveness becomes essential to ensure safe and optimal task execution.

A key aspect of expertise assessment in telerobotics is motor performance, which plays a crucial role in determining operator skill. Motor performance has traditionally been measured using objective metrics such as task completion time, path length, and the number of corrective movements, providing insights into dexterity and precision in high-stakes domains like surgery, aerospace, and industrial control [8]. However, existing measures often fail to capture the underlying cognitive processes that differentiate expert and novice operators.

To address this, human–machine interface (HMI) research has explored how users interact with control devices such as computer mice, keyboards, and joysticks. Before the advent of gaze-tracking technologies [9], user attention in 2D interfaces was inferred through input methods such as mouse and joystick movements. Usability studies in these contexts measured indicators like travel time between interface elements, task completion times, and error rates, offering valuable insights into human performance—particularly for untrained users [10,11]. Applying similar techniques to teleoperation could provide a more comprehensive understanding of how expertise influences both motor control and cognitive strategies.

This study, building on previous insights from [12], shifts the focus from system design to a deeper investigation of manual dexterity, motor performance, and visual attention in high-precision teleoperation tasks. The main contribution of this paper is the introduction of a novel, integrated evaluation framework that combines classical performance metrics with gaze-based measures to objectively differentiate operator expertise. A key component of this work is the utilization of two gaze metrics—gaze transition entropy (GTE) and stationary gaze entropy (SGE)—which effectively quantify visual attention strategies in bilateral teleoperation. By analyzing how expertise influences task execution and gaze behaviour, the study moves beyond the widely accepted notion that “experts perform better” to explain how and why experts excel across multiple dimensions of teleoperation performance. The findings demonstrate that expertise significantly impacts decision-making, motor control, and gaze coordination, with experts exhibiting more structured gaze patterns and smoother manipulator control. Statistical analyses reveal significant performance differences between the two groups (i.e., novice and expert), underscoring the potential of gaze-based metrics

for evaluating operator proficiency and informing future teleoperation training programs. This contribution provides a new, data-driven approach for evaluating skill acquisition and guiding interface design in the context of bilateral teleoperation.

## 2. Related Work

Time is one of the most commonly analysed metrics, as it directly reflects the efficiency of task execution. Early foundational work by Welford [13] introduced the measurement of reaction times and task durations in motor performance studies. Fitts's law [14], which models the relationship between movement time, target distance, and target size, is widely used to predict the time required to perform tasks involving pointing or reaching movements. The law has been a cornerstone in understanding human motor control and is particularly relevant in robotic teleoperation and haptic systems [15]. Hannaford et al. [16] extended these concepts in the field of robotics, emphasizing that minimizing task execution time is critical for improving teleoperation performance, especially when the operator is working under time constraints or limited visual feedback.

Recent work continues to highlight the importance of time as a performance metric, particularly in applications requiring precision, such as robotic surgery. Tugal et al. [17] examined the role of task completion time in robotic-assisted interventions, showing that faster completion times are often associated with more skilled operators in minimally invasive surgery. Additionally, studies in [18,19] applied Fitts's law to robotic teleoperation and haptic interfaces, demonstrating how movement time is influenced by target size and distance, directly affecting task efficiency in complex environments like remote manipulation tasks.

Path length, which measures the distance travelled by the hand or tool during task execution, is another valuable metric. It captures the efficiency of movement and is often used to detect unnecessary or sub-optimal actions. Hwang et al. [20] showed that shorter path lengths tend to correlate with greater user proficiency in surgical operations. More recently, studies in [21,22] have demonstrated that path length is a reliable predictor of motor skill in robotic surgery and teleoperation. These studies emphasize the importance of smooth and controlled movements for optimal performance. Movement frequency and the number of corrective actions further provide insights into the user's motor control and cognitive load. The study in [22] showed that novice users tend to make more frequent movements, indicative of less precise control compared to experts. More recent studies, such as those in [23], suggest that the number of corrective movements is a key indicator of motor learning and skill acquisition in tasks involving robotic manipulation.

Eye tracking is widely utilized across various fields to analyse human behaviour and cognitive processes. In robotic surgery, researchers have used eye tracking to assess surgeons' workload, gaze patterns, and visual attention distribution during laparoscopic operations [24–26]. Studies have shown that expert surgeons tend to focus more on task-relevant areas with longer fixation durations and shorter saccade durations compared to novices [27]. Similarly, in aviation, eye tracking has been instrumental in highlighting differences in monitoring behaviour between experienced pilots and novices [28,29]. Eye tracking has also been applied in the field of driving to evaluate hazard perception and driver behaviour [30,31]. These studies underscore the versatility and importance of eye tracking as a tool for understanding human cognition and performance across various domains.

Beyond basic motion analysis, research in precision domains has highlighted the importance of expert classification using soft computing and sensor integration. For instance, approaches such as those in [32] leverage data-driven classification techniques to detect and assess defects in biomedical components, providing inspiration for similarly

robust methods of classifying human performance in robotic systems. Although their focus is on material delamination, the study offers a conceptual parallel in terms of using sensitive detection frameworks and classifier robustness in critical environments.

In line with these trends, operator classification in human–robot interaction must also consider system-level sensitivity. Robotic systems used in surgical or hazardous applications are expected to respond to micro-level hand motions with minimal delay and high fidelity. Such requirements have led to increased research into adaptive, sensor-rich interfaces that complement skilled operator input with high-resolution response fidelity. Sensitivity in this context is not only a hardware attribute but also a crucial factor in interpreting user intent and ensuring safe task execution.

Entropy is a measure of randomness, uncertainty, or disorder in a system and is a fundamental concept across various fields, including physics, information theory, and statistics [33]. In information theory, entropy reflects the amount of information within a message, or more specifically, the level of uncertainty associated with that message. The entropy of a sequence  $x$  is calculated using Shannon's equation:

$$H(x) = - \sum_{i=1}^n p(x_i) \log_b p(x_i), \quad (1)$$

where  $n$  is number of states,  $p(x_i)$  is the probability of each state in  $x$ , and  $\log_b$  is the logarithm with base  $b$ . A high entropy value indicates a greater degree of variation or unpredictability in the information, while a low entropy value signifies more predictability and less uncertainty [34].

In the context of gaze behaviour, gaze entropy measures the complexity and unpredictability of how an individual scans a visual scene [35]. It can be used to assess focus, task engagement, or cognitive load. For example, the entropy of eye movements may indicate how much difficulty a person is experiencing while processing information or how easily they are distracted. In usability studies, eye movement entropy helps to evaluate how users interact with websites, software, or devices, shedding light on how intuitively they navigate an interface.

There are two primary types of gaze entropy: stationary gaze entropy (SGE) and gaze transition entropy (GTE):

- Stationary gaze entropy (SGE) measures the overall distribution of eye movements over time. It assumes that each saccade (rapid eye movement) is independent and reflects the predictability of gaze points. Fixation coordinates are used to compute SGE values, and a duration threshold may be applied to determine how long the eye must remain fixed on a point to count as a fixation. Lower entropy corresponds to more predictable gaze patterns, while higher entropy indicates more irregular and varied gazes. This metric is particularly useful for assessing the regularity of eye movements when the observer is focused on a single point for extended periods.
- Gaze transition entropy (GTE) measures the complexity of transitions between gaze points. It uses a transition matrix to describe how quickly the eyes move between predefined spatial regions and how information is distributed across these transitions. GTE is commonly employed in eye-tracking studies to analyse the path that the eyes follow within a scene and the amount of information these transitions convey.

Together, these measures provide valuable insights into gaze behaviour, revealing patterns related to attention, cognitive effort, and visual processing.

Studies [35,36] provide methods for calculating entropy for both SGE and GTE. Briefly, entropy is calculated after discretizing the visual space into specific sections using spatial grids, a common method in gaze-tracking studies. By grouping fixations into predefined

regions, it becomes possible to examine the predictability and complexity of gaze patterns. The number of these state spaces affects spatial specificity; the more state spaces, the higher the specificity, which directly influences the maximum observable entropy. Higher entropy values indicate more complex and unpredictable gaze patterns, while lower entropy reflects more focused and predictable viewing behaviour.

Eye movements are influenced by both bottom-up and top-down processes. Bottom-up processes are driven by external visual stimuli, such as colour or brightness, while top-down processes are influenced by internal factors like task demands, experience, and goals. These dynamics suggest that gaze control is a predictive process shaped by the interaction of the visual environment and cognitive factors. Gaze entropy is, therefore, a useful metric for assessing attention, cognitive load, and the interaction between a user and visual interface.

In remote robotic operations, situation awareness (SA) and workload are essential for ensuring operational safety and optimizing human performance [37]. These concepts are widely acknowledged across various industries, including healthcare [38,39], transportation [40], aviation [41], and telerobotics [42,43]. Generally, in such studies, self-reporting methodologies such as the Situation Awareness Rating Technique (SART) [44] and NASA Task Load Index (TLX) [45] are utilized. For an instance comparison of the validity of these approaches, see [46].

Overall, the integration of multiple performance metrics—situation awareness, time, path length, movement frequency, and gaze—offers a comprehensive understanding of motor performance. As system complexity increases in fields such as medicine and precision manufacturing, aligning operator skill with robot sensitivity and feedback becomes essential. Thus, classification techniques that incorporate both motion and perceptual data are likely to shape the future of adaptive training systems and interface design.

### 3. Methods

#### 3.1. Participants

The study participants were employees and secondees from the Department of Remote Applications in Challenging Environments (RACE) within the UK Atomic Energy Authority (UKAEA). The group mainly consisted of engineers, technicians, and operators who had prior experience and familiarity with teleoperation.

A total of 10 users (1 female and 9 male) participated in the study. They were categorized into two groups: novice and expert. The novice participants had an average of 9 months of experience in teleoperation, while the experts had around 5 years of experience.

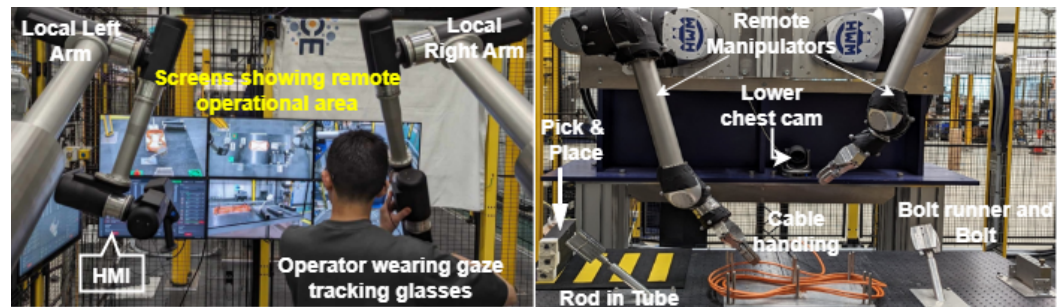
The procedures avoided invasive or potentially dangerous methods. Data were stored and analysed anonymously. All participants provided written informed consent.

#### 3.2. Experimental Setup

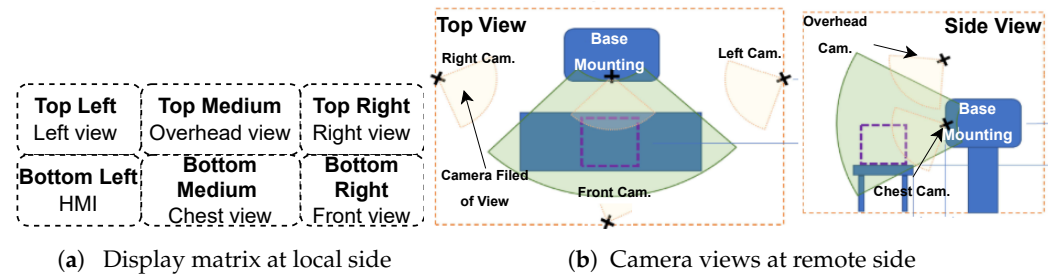
The experimental consisted of a dual-hand Telbot bilateral teleoperation system (see Figure 1), Tobii gaze-tracking glasses, and questions for the participants after completing the experiment.

The Telbot system is a bilateral telerobotic system equipped with remote manipulators that offer seven degrees of freedom and can carry loads of up to 20 kg at their end effectors. These manipulators are controlled by local robots that are kinematically similar, each featuring six degrees of freedom. Human operators manage these local robots, ensuring precise and responsive control. Five cameras are positioned on the remote side, as shown in Figure 2b, capturing images that are then projected onto monitors in front of the operators, including the HMI of the telerobotic system, as shown in Figure 1 (left) and Figure 2a.

We captured the robotic system’s internal sensory information using the OPC Unified Architecture (OPC-UA) protocol, with a sampling frequency of  $f_s = 1$  kHz. Eye movements were recorded using Tobii Pro Glasses 3, a wearable mobile eye-tracking system that samples eye positions at 50 Hz. The recordings were analysed using Tobii Pro Lab software Version:1.194. To minimize experimenter effects, such as “eye-tracker awareness”, operators received no instructions other than to perform their tasks as usual. Also, throughout the study, data on operators’ task duration and errors were recorded. Subsequently, operators were asked to complete questionnaires regarding telerobotic handling qualities.



**Figure 1.** Telbot dual-hand bilateral telerobotic system: the left view depicts the local side, while the right view shows the remote side, featuring various tasks used in the experiments.



**Figure 2.** Display matrix at the local side and illustrative top–down and side views of remote manipulators cell showing camera positions.

### 3.3. Experimental Procedure: Tasks

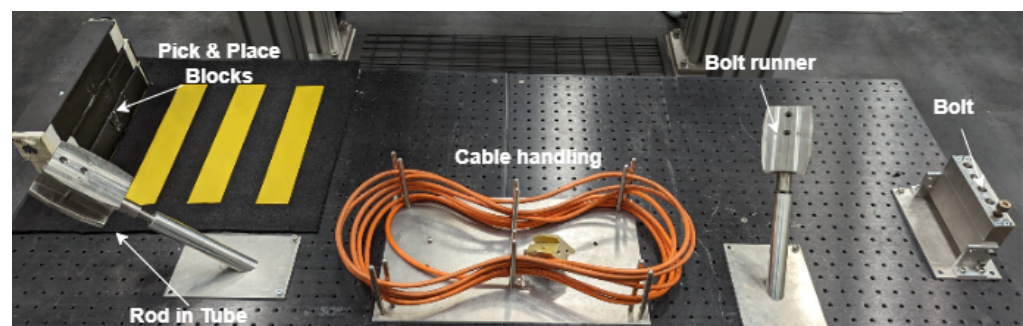
Operators were asked to complete, with a random order, six distinct tasks, with a primary focus on accuracy while also considering time. The first five tasks—pick and place, rod in tube, bolting, cable handling, and wire loop—were analysed in terms of manipulator motion, task completion time, and gaze behaviour. The sixth task was designed specifically to evaluate performances using Fitts’s Law, capturing key constraints and parameters relevant to remote robotic operations. While these tasks do not exactly replicate real-world operations, they incorporate the critical elements required for meaningful analysis. During the experiments, to reduce the impact of fatigue, operators were permitted to take rest breaks between tasks at their discretion.

- Pick and place: This task centres around manipulating blocks of similar size and visual appearance yet composed of distinct materials (i.e., different weights such as 50 g, 2 kg, and 6 kg). Initially, these blocks are stacked at a designated starting point. The primary goal is to correctly position the blocks in their designated spots from the stacking location, taking into account their individual weights. Following the determination of the placement sequence, the user is then required to return the blocks to the original stacking location.
- Rod in tube: The arrangement comprises a rod and a tube, as seen in Figure 3, where the length of the rod surpasses 100 mm and the tube length extends beyond 80 mm. Participants are tasked with accomplishing this assignment utilizing their right-hand

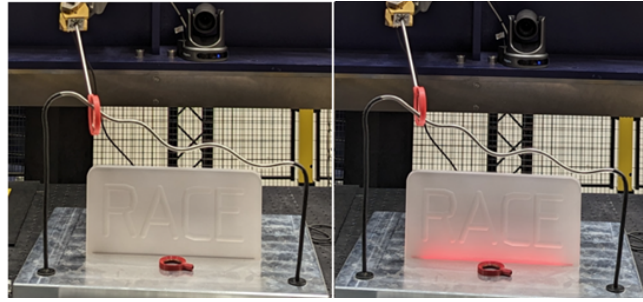
arm/device, all while avoiding the jamming or wedging of the rod and refraining from exerting undue force on either the rod or the tube. The plate that holds the tube will be firmly affixed to the surface, positioning the tube at a  $90^\circ$  phase angle relative to the robot's base.

- **Bolting:** This test involves two blocks connected by a dowel, with the upper block designed to accommodate a bolt and the lower block featuring a tapped hole. A single M10 remote-handling-style bolt is used for this specific task. Participants are instructed to fully tighten the bolt, ensuring that excessive torque is avoided and cross-threading is prevented. After completing the tightening phase, participants must then disengage and re-engage the bolt, carefully undoing it and returning it to its initial position.
- **Cable handling:** This task replicates remote cable handling activities, emphasizing the need for the precise and direct manipulation of cables using grippers to prevent any damage. The evaluation involves a 10 m length of standard multi-core electrical cable with a 7 mm diameter, including a remote-handleable connector at one end (refer to Figure 3). The cable is initially wound onto a fixture, and participants must use both manipulators (left and right) to carefully unwind it, passing it between hands as needed. The cable must then be placed on the table in a structured manner that ensures that it remains untangled, within reach and view of the telerobotic system, and free of loops or overhangs. This arrangement allows for efficient rewiring using the telerobotic devices without the risk of entanglement or loss of control.
- **Wire loop:** In this task, users guide a metal loop (probe) along a winding wire path without touching the wire (see Figure 4) [47]. If contact occurs, an electric circuit triggers light and sound (a buzzing noise), indicating an error. Participants are instructed to navigate the probe back and forth, minimizing contact to assess the system's positional accuracy and sensitivity.
- **Multi-rod-in-tube:** This task quantifies performances based on task difficulty and operator experience, similarly to Fitts's law. Participants insert a 12 mm dowel into holes of varying diameters (13.65 mm, 12.5 mm, 11.35 mm, and 10.2 mm) and distances (100 mm, 300 mm, 500 mm, and 700 mm), reorienting the rod between trials at tilt angles of  $45^\circ$  and  $60^\circ$  (see Figure 5). The goal is to complete as many insertions as possible within a set time (e.g., 1 min), with difficulty increasing as the hole size decreases and distances expand.

While the tasks are ordered based on the operators' perceived difficulty, no quantitative comparison between them can be made using Fitts's law. Therefore, the final task was specifically included in the experiment for this purpose.



**Figure 3.** Pick-and-place, rod-in-tube insertion, cable handling, and bolting tasks performed using remote manipulators.



**Figure 4.** The wire loop game; when the probe (metal loop) makes contact with the wire, the light-emitting device will illuminate (as depicted in the image on the right), accompanied by a buzzing sound from the sound-emitting device.



**Figure 5.** Multi-rod-in-tube task: the task difficulty varies based on the hole diameters (13.65 mm, 12.5 mm, 11.35 mm, and 10.2 mm) and the distance between the holes (100 mm, 300 mm, 500 mm, and 700 mm).

### 3.4. Manipulators' Motion

Throughout the tasks, the remote manipulators began from identical initial positions. Employing the recorded joint angles, we calculated the total path length ( $\Delta_e$ ) covered by the remote manipulators' end-effectors and the average manipulability ( $\bar{\mu}$ ), and we assessed the trajectory's smoothness using the jerk. The total path length is determined as follows:

$$\Delta_e = \sum_{k=1}^{n-1} \sqrt{(x_{k+1} - x_k)^2 + (y_{k+1} - y_k)^2 + (z_{k+1} - z_k)^2},$$

where  $x$ ,  $y$ , and  $z$  are end-effector positions with respect to the base, and  $n$  is the maximum number of recorded samples.

Dexterity plays a crucial role in remote handling, enabling the serial manipulator to execute complex tasks without encountering joint limits. The manipulability index,  $\mu$ , serves as a proxy for measuring the dexterity of the feasible configurations of the manipulator. For non-redundant manipulators, it can be expressed as follows:

$$\mu = \sqrt{\det(J(q)J(q)^T)} = \prod_i \sigma_i,$$

where  $\sigma_i$  denotes the singular values for the Jacobian matrix ( $J(q)$ ).

Agile and smooth point-to-point movements are crucial for operational safety. By examining the jerk of the end-effector, one can assess the smoothness of the tip trajectories in the operational space [48]. The jerk can be derived through the Jacobian matrix and its time derivatives:

$$\hat{j}_e = \ddot{J}(q)\dot{q} + 2\dot{J}(q)\ddot{q} + J(q)\ddot{\ddot{q}}.$$

### 3.5. Gaze Tracking

In this study, gaze tracking was employed to capture the focal points of trained operators as they navigated through multiple screens to perform specific tasks. These tasks required compensating for the lack of depth perception while interacting with buttons and tools. We analysed their gaze patterns using heat maps and entropy measures to quantify the predictability and complexity of their visual behaviour. The experiment involved six different viewing angles projected onto a display matrix, allowing the operator to carry out

remote operations. Each of these angles represented a state, resulting in a six-state space. By identifying where the operators' fixations occurred within this space, we calculated the probability values  $p_i$ , which were then used to determine the stationary gaze entropy based on (1). Table 1 shows the naming of each state and an example of the probability distribution of fixations. Figure 2a illustrates the display matrix, representing the state space of visual field regions.

**Table 1.** Distribution of illustrative fixations in visual field regions.

State Space ( $i$ )	Number of Fixations	Proportion of Fixations ( $p_i$ )
1. Top Left (TL)	21	0.21
2. Top Medium (TM)	15	0.15
3. Top Right (TR)	14	0.14
4. Bottom Left (BL)	23	0.23
5. Bottom Medium (BM)	6	0.06
6. Bottom Right (BR)	21	0.21
Total	100	1

Gaze transition entropy (GTE) was calculated to assess the complexity and unpredictability of transitions between different gaze points. The formula used is as follows:

$$H = - \sum_{i=1}^n p_i \sum_{j=1}^n p(i,j) \log_2 p(i,j),$$

where  $H$  represents the uncertainty of the state sequence  $x$  given that the previous state is known,  $p_i$  denotes the stationary distribution for state  $i$ , and  $p(i,j)$  is the transition probability from state  $i$  to state  $j$ . This calculation allows us to assess the complexity and unpredictability of transitions between gaze points during task execution.

The visual space was divided into six discrete regions (state spaces), corresponding to the viewing angles used during the remote operation. By discretizing the visual environment, we were able to calculate probability distributions for fixations in each region. The entropy for both SGE and GTE was computed based on the frequency of fixations and transitions within these regions.

SGE and GTE offer complementary perspectives on gaze behaviour. SGE focuses on the distribution of fixations across different regions, providing insights into how predictable an operator's gaze points are. GTE, on the other hand, emphasizes the transitions between these regions, measuring the complexity of gaze movements. Together, these measures help us understand the multifaceted nature of visual attention during task execution, with different trends emerging based on task complexity and operator gaze behaviour.

#### 4. Main Results

The duration of task completions and any errors encountered during pick-and-place and wire loop tasks were investigated. Additionally, remote manipulator motions, including manipulability, jerk, and total path length, were analysed across all experimental tasks.

Meaningful differences in task duration and remote manipulator motion across expertise levels were assessed through statistical analyses on all groups. Normality tests were performed on the data groups, and for those failing the initial test, Box–Cox transformation was applied (with the same  $\lambda$  used for transformation across compared groups). Subsequently, all groups passed the normality tests at a significance level of  $p = 0.05$ .

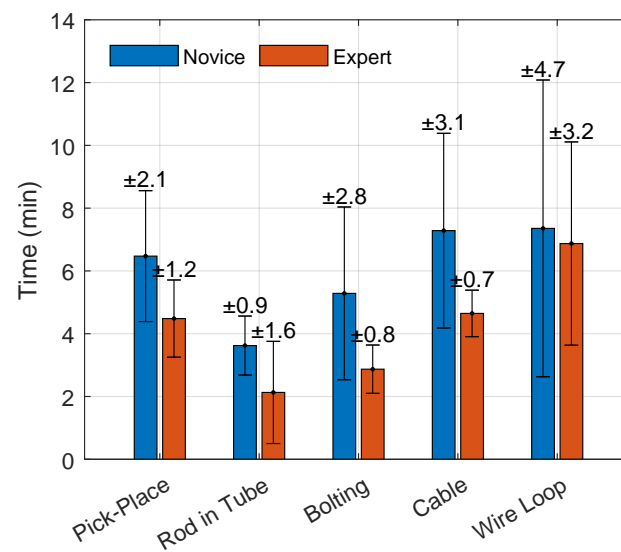
The influence of expertise in dual comparisons, such as task completion duration and expert–novice correlations, was analysed using Welch's t-test (implemented in Matlab

R2023a using *ttest2()*). A significance threshold of  $p = 0.05$  was consistently employed for all statistical tests in the paper.

#### 4.1. Duration of Task Completion and Error Analyses

Typically, it is expected that task completion time will decrease with increasing experience. However, it is crucial to note that task completion duration alone does not offer a comprehensive measure of performance. For example, experienced operators often prioritize error prevention over speed, resulting in a more balanced assessment of their proficiency.

Figure 6 depicts the average task durations for each group, highlighting the notable trend that experienced users tend to complete tasks more swiftly. However, it is evident that there is considerable variability among users, which is underscored by the substantial standard deviation shown in the figure.



**Figure 6.** Average task durations for each group.

The analyses indicate a statistically significant difference ( $p = 0.0132$ ) in task completion durations between experts and novices. More specifically, experienced users consistently complete all five tasks 2.6 min faster compared to novice users.

The average errors (standard deviation) committed by each group were analysed in two tasks: pick and place and wire loop. In the wire loop task, the recorded errors indicate instances where participants made contact between the probe and the wire. For the pick-and-place task, the numbers represent how often blocks were inaccurately positioned, reflecting difficulty in discerning the weight differences.

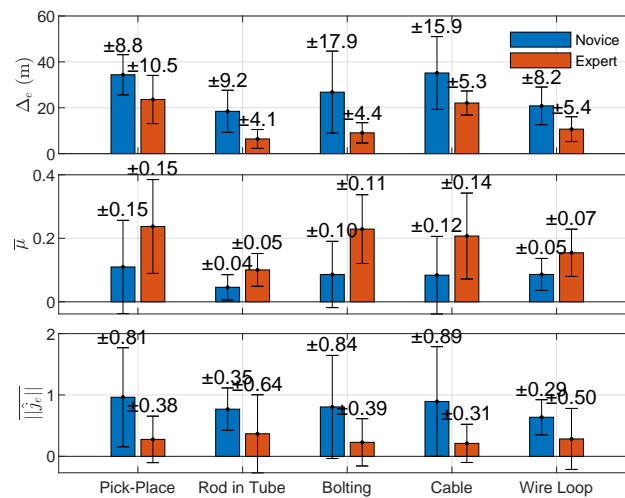
In the wire loop task, expert users not only completed the task more rapidly but also made fewer mistakes compared to novice users, as detailed in Table 2. Conversely, in the pick-and-place task, expert users exhibited a higher frequency of errors. Specifically, they encountered difficulty distinguishing the weights of the light and medium blocks. This difference may be attributed, as mentioned by experienced operators during interviews, to the extensive experience they have with the MASCOT system [49,50], which reflects less electromechanical impedance relative to the operators compared to the system under consideration.

**Table 2.** Average (standard deviation) error made by each group in two tasks; pick and place (the number of times blocks were inaccurately positioned) and wire loop (number of instances where the probe made contact with the wire).

Subject Group	Tasks	
	Pick and Place	Wire Loop
Novice	0.33 ( $\pm 0.51$ )	22.33 ( $\pm 3.61$ )
Expert	0.75 ( $\pm 0.50$ )	7.5 ( $\pm 5.44$ )

#### 4.2. Motion of the Remote Manipulators

The average calculated total path length, manipulability, and jerk for each task is illustrated in Figure 7.



**Figure 7.** Average total path length, manipulability, and jerk at teaching task among the groups.

Expert users clearly perform fewer motions with the remote manipulators, evidenced by a statistically significant difference ( $p = 2.1415 \times 10^{-5}$ ) in remote manipulator displacement when compared to novice operators.

Furthermore, expert operators tend to position remote manipulators closer to the centre of the workspace compared to novice operators. This is reflected in a statistically significant difference ( $p = 0.000172$ ) in remote manipulator's average manipulability between expert and novice operators.

Moreover, not only do expert operators control remote manipulators with less displacement and optimal postures but they also execute smoother movements. This is supported by a statistically significant difference ( $p = 0.000365$ ) in the remote manipulator's total jerk when comparing expert and novice operators.

#### 4.3. Penalty Method

In the multi-rod-in-tube task, the distances between pairing tubes and their diameter size varied systematically. In this way, a difficult index (ID) can be calculated as follows:

$$ID = \log_2\left(\frac{d}{\omega} + 1\right),$$

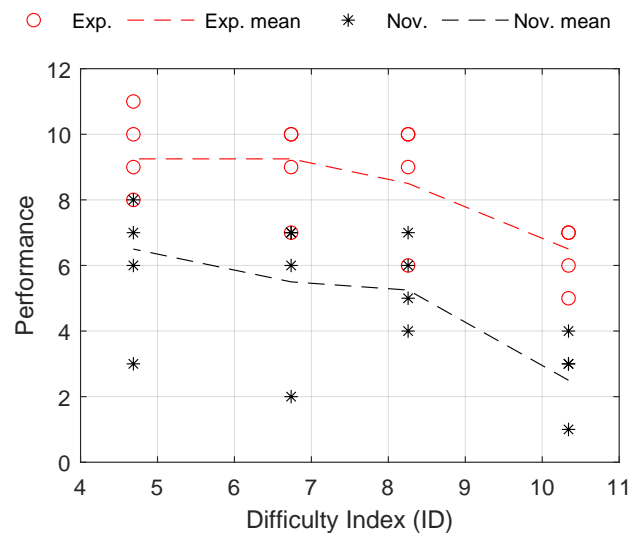
where  $d$  denotes the distance between the pairing tubes with same diameter, and  $\omega$  is the width between the tube's and rod's diameters.

Figure 8 compares the performance of expert and novice operators in the multi-rod-in-tube task as a function of the task difficulty index (ID). Higher performances are shown on the y-axis, with the difficulty increasing along the x-axis (ID 4 to 11).

Expert operators (red circles) maintained high performance at lower difficulty levels (ID 4–6) but experienced a slight decline as difficulty increased beyond ID 6. While their performance dropped with more complex tasks, they remained relatively consistent compared to novices.

Novice operators (black asterisks), on the other hand, showed a sharp decline in performance, particularly after ID 7. The fitted model (black dashed line) highlights a steady decrease as tasks became more complex, indicating greater difficulty in managing challenging tasks.

A noticeable gap emerged between experts and novices at higher difficulty levels (ID 9–10), with novices struggling significantly more. The results suggest that while experts adapt better to increasing difficulty, novice performance deteriorates rapidly, indicating a need for additional training or task refinement for novices at these complexity levels.



**Figure 8.** Comparison of performance versus difficulty index for expert and novice operators, along with the corresponding models, during the multi-rod-in-tube task.

#### 4.4. Gaze Tracking

The gaze heat map for the tasks, illustrated in Figure 9, offered insights into the distinct approaches employed by expert and novice operators. Previous studies have suggested that fixation duration, representing the total time spent in fixations, reflects the information processing load and tends to increase with workload [24]. Here, similarly to [24], the absolute fixation duration time is scaled to a percentage of the exercise duration as

$$FD(\%) = \frac{\text{Sum of fixation duration}}{\text{Exercise duration}} \times 100.$$

In the pick-and-place task, expert operators demonstrated a focused strategy, precisely placing each block using both overhead (top middle in the display matrix) and chest (bottom middle) cameras. Novice operators, on the other hand, predominantly relied on the chest camera (see Figure 9).

For the rod-in-tube task, novice operators tended to inspect the rod angle by utilizing both the overhead and chest cameras to align it with the tube. In contrast, expert operators efficiently maintained the rod's position for pulling in/out, relying solely on the overhead camera. Novice operators placed greater emphasis on the front camera for pulling in/out

the rod, while expert operators used it less frequently, relying on their expertise to complete the task smoothly.



**Figure 9.** Operator gaze heat map during the experiments: top expert and bottom novice.

In the cable handling task, expert operators leaned on the overhead camera for uncoiling, leveraging their familiarity with the task. Novice operators, however, tended to utilize both overhead and chest cameras for uncoiling, suggesting a need to check more cameras during the task.

For the wire loop task, expert operators relied heavily on both overhead and chest cameras, with relatively fewer views from the left and right cameras. Novice operators, while also using the overhead camera, needed to check the right and left cameras more frequently than their expert counterparts, potentially leading to additional time spent on camera checks to complete the task.

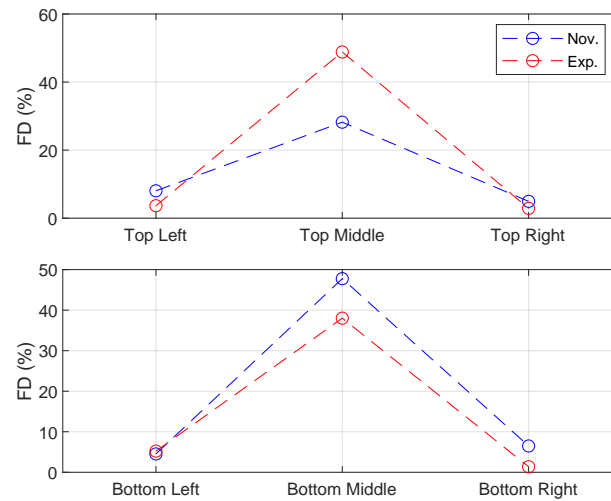
Figure 10 shows the fixation duration percentage of novice and expert operators with respect to various viewpoints. Novice operators mainly focus on the camera with a similar viewpoint to the users, while experts smoothly navigate through multiple angles. These findings highlight the different visual strategies used by expert and novice operators in bilateral telerobotic operations. Expert operators compensate for the lack of 3D perception by scanning multiple viewing angles continuously, while novices tend to focus mainly on the monitor displaying the same viewpoint. Intensive training and good hand–eye coordination are considered crucial for effectively scanning multiple viewing angles.

By analysing both the GTE and SGE, additional differences between expert and novice operators were also observed. Experts demonstrated more focused and stable gaze patterns, while novices exhibited more scattered and inconsistent eye movements. The entropy values varied across tasks, influenced by both task complexity and the participants' experience levels (see Figure 11a,b). These findings suggest that experts not only direct their gaze more efficiently but also exhibit lower gaze entropy, indicating better control- and task-oriented focus during remote operations.

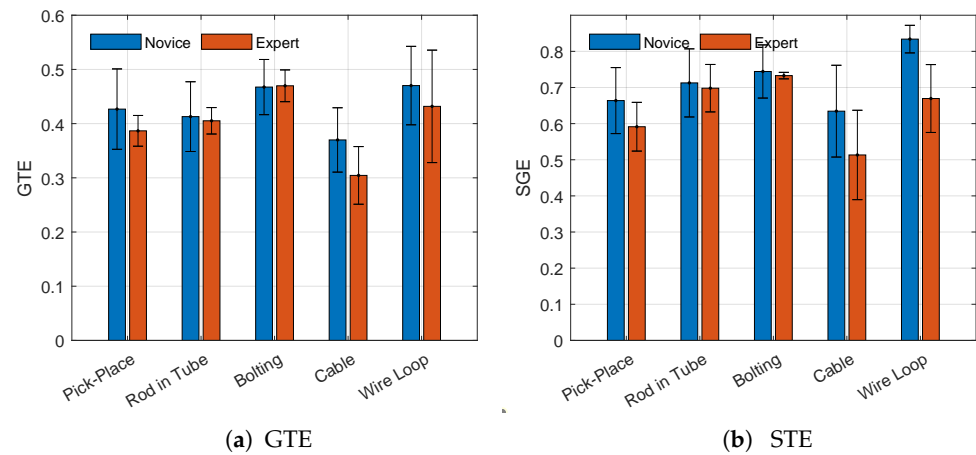
The GTE quantifies the variability or complexity in the sequence of gaze shifts between different points of interest. Higher GTE values indicate more erratic or inconsistent gaze patterns. In this study, experts had a lower average GTE (1.998) compared to novices (2.147), suggesting that experts exhibited more stable and predictable gaze movements, whereas novices showed more irregular and less controlled transitions between gaze points (see Table 3). This difference highlights the efficiency of expert gaze control during task execution.

**Table 3.** Average GTE values for novice and expert users, aggregated over ten participants across five different tasks.

Subject Group	Tasks					GTE Sum
	Pick and Place	Rod in Tube	Bolting	Cable Handling	Wire Loop	
Novice	0.4267	0.4129	0.4674	0.3699	0.4702	2.1472
Expert	0.3866	0.4052	0.4698	0.3044	0.4318	1.998



**Figure 10.** Fixation duration percentage of novice and expert operators on various viewpoints.



**Figure 11.** The novice and expert users’ gaze transition and stationary gaze entropies across five tasks.

As illustrated in Figure 11a, novice operators generally showed higher GTE values, indicating more unpredictable and inefficient gaze movements, particularly in tasks such as cable handling and wire loop. These tasks posed greater challenges for novices, resulting in higher gaze entropy and more erratic visual scanning. Conversely, experts displayed lower entropy, indicative of more focused and goal-directed gaze strategies, which implies a more efficient processing of visual information during these complex tasks.

The SGE measures the duration of fixations and the distribution of gaze points. A higher SGE value indicates more varied and dispersed attention, meaning that the individual shifts focus frequently or has highly variable fixation durations. In this study, experts had a lower average SGE (3.205) compared to novices (3.589), as shown in Table 4, suggesting that experts maintained a more focused and steady gaze, concentrating on

fewer points for longer periods. This indicates that experts are less prone to distractions, allowing them to maintain sustained attention on critical areas during task execution.

As visualized in Figure 11b, novices generally exhibited higher SGE values, indicating more scattered and inconsistent fixation behaviour. This was particularly notable in tasks like cable handling and wire loop, where experts demonstrated significantly lower SGE, reflecting their ability to focus on critical areas with longer fixation durations. The lower entropy in experts points to their superior ability to sustain attention on important regions during task execution, while novices distribute their attention more unevenly, resulting in greater variability in their gaze patterns and reduced task efficiency.

**Table 4.** Average SGE values for novice and expert users, aggregated over ten participants across five different tasks.

Subject Group	Tasks					SGE Sum
	Pick and Place	Rod in Tube	Bolting	Cable Handling	Wire Loop	
Novice	0.6637	0.7126	0.7442	0.6344	0.8338	3.5891
Expert	0.5914	0.6980	0.7329	0.5132	0.6694	3.2051

Experts demonstrate more focused, consistent, and goal-oriented gaze patterns, whereas novices tend to exhibit more random and erratic eye movements. The lower GTE and SGE values for experts suggest that they manage their gaze more efficiently and with greater control during task execution. This is particularly evident in tasks like “Wire Loop” and “Cable Handling”, where experts show significantly lower entropy, while novices display more irregular gaze behaviour.

These findings highlight that expertise significantly influences gaze control, with task complexity also playing a role in gaze patterns. The differences in entropy values suggest that gaze entropy could serve as a useful metric for distinguishing levels of expertise, offering potential solutions for optimizing operator performance based on gaze behaviour analysis.

#### 4.5. Impact of Gaze Entropy Metrics on Skill Classification

To evaluate the added value of incorporating gaze entropy metrics—GTE and SGE—in distinguishing operator skill levels, we conducted a comparative analysis using Hedges’  $g$  effect sizes across both individual metrics and composite scores [51].

The composite effect size based on motor performance metrics alone (task time, path length, jerk, and manipulability) was Hedges’  $g = 2.175$ . When gaze entropy metrics were included, the composite effect size increased to  $g = 2.547$ , demonstrating a measurable improvement in discriminatory power. This result supports the hypothesis that gaze-based metrics contribute additional, complementary information beyond traditional motor indicators.

Among individual metrics, the SGE during the wire loop task yielded the highest effect size ( $g = 2.286$ ), indicating that gaze regularity is a strong predictor of expertise in complex teleoperation scenarios. These results highlight the relevance of visual attention measures for expert classification and support their integration into future operator evaluation and training frameworks.

#### 4.6. Questionnaires

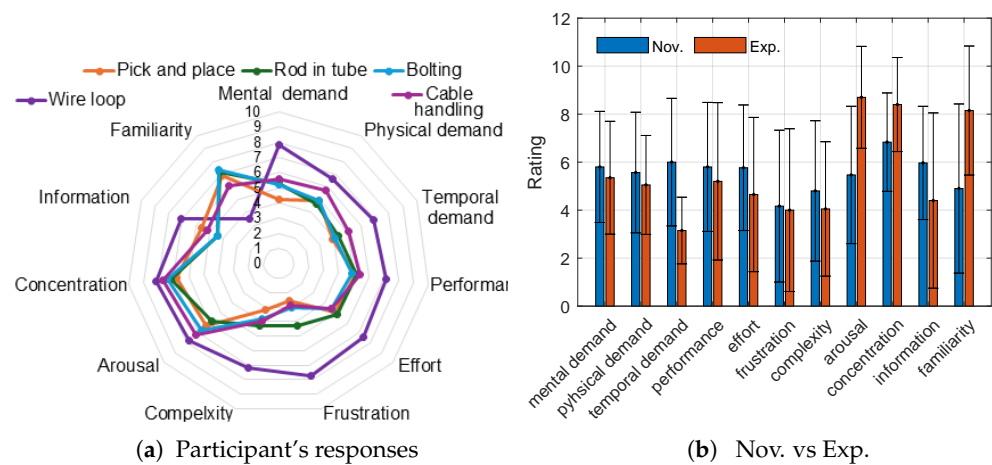
After completing the experiments, participants were asked to complete an 11-question survey (similar to SART and NASA TLX questionnaires) about their impressions for each

task performed. These questionnaires assessed various categories, with participants providing ratings on a scale from 1 to 10 for the following:

- Mental, physical, and temporal demands; performance; effort; frustration; complexity; arousal; concentration of attention; information quantity; and familiarity.

Figure 12a graphically represents the participants' responses. Across all tasks, participants consistently demonstrated high levels of concentration and arousal.

With the exception of the wire loop game, participants exhibited familiarity with the tasks. As a result, the mental, physical, and temporal demands were generally at a moderate level. It is noteworthy that task familiarity, regardless of complexity, influenced the amount of effort participants needed to exert to complete the task. The wire loop game stood out as the least familiar task for participants, resulting in elevated levels of mental, physical, and temporal demands, as well as increased effort and frustration.



**Figure 12.** Operators' response to the questions.

Figure 12b displays the user responses to the questionnaires categorized by their experience level. Overall, experts showed higher levels of arousal ( $p = 0.0001$ ), concentration ( $p = 0.0084$ ), and familiarity with the tasks ( $p = 0.0034$ ), while novices reported higher temporal demand ( $p = 0.0002$ ).

During the trials, it was observed that the majority of operators successfully completed tasks without errors, such as dropping blocks or jamming the rod. However, operators did not receive post-trial feedback on their performance, except for the wire loop game, where they could observe their mistakes. For instance, feedback on whether they managed to sort blocks according to their weights was omitted. In the questionnaires, most operators reported performing well during the trial, indicating a high level of self-assessment skill for remote telerobotic operations. Furthermore, the importance of training emerged in the questionnaires, with operators noting that they required more effort to complete tasks that they were less familiar with.

## 5. Discussion

This study provides a comprehensive analysis of operator expertise in bilateral telerobotic systems by evaluating both objective performance metrics and subjective user feedback. The findings highlight key parameters that differentiate expert operators from novices, offering valuable insights for training and system optimization.

One of the most significant distinctions between experts and novices lies in their ability to efficiently complete tasks while minimizing unnecessary motion. Performance metrics such as task completion time, total path length, jerk, and remote manipulator manipulability clearly demonstrated that experts consistently outperformed novices. These

differences suggest that expertise is characterized by greater motor efficiency and refined control strategies, which are essential for optimizing teleoperation performance.

Another critical aspect of expertise is the ability to compensate for perceptual limitations inherent in telerobotic systems. Experts demonstrated a superior ability to scan multiple viewpoints, allowing them to better interpret spatial relationships despite the lack of depth perception. Novices, by contrast, often relied on a single display, which may have contributed to their reduced situational awareness and less efficient task execution.

The introduction of the penalty method provided a novel perspective on performance relative to task difficulty. While experts maintained consistent performance across increasing difficulty indices, novices exhibited a sharp decline in effectiveness as complexity increased. This highlights a key challenge in teleoperation training—helping novice operators build adaptability and resilience when faced with more demanding tasks. Additionally, based on a predictive model of operator performance with respect to varying difficulties, the operator's experience level can be quantitatively estimated, providing a useful tool for automated skill assessment and training personalization.

Gaze entropy analysis, particularly through GTE and SGE, revealed additional differences in cognitive processing strategies. Experts displayed lower entropy values, reflecting structured and purposeful gaze behaviour, whereas novices exhibited higher entropy, indicative of erratic and inefficient visual scanning. This pattern was especially evident in tasks such as cable handling and the wire loop challenge, where experts' lower gaze entropy suggested superior attentional control and task-specific visual strategies.

Subjective questionnaire responses further reinforced these findings, highlighting disparities in mental and physical workloads between experts and novices. Novices reported higher levels of temporal demand and frustration, particularly in unfamiliar tasks, whereas experts exhibited greater arousal, concentration, and familiarity. The alignment between subjective feedback and objective performance metrics emphasizes the role of experience in managing both physical and cognitive demands in teleoperation scenarios.

Beyond bilateral teleoperation, the proposed method offers broader applicability in evaluating the effectiveness of human–robot interaction in collaborative tasks. For instance, it can be applied to assess human-guided robotic systems in scenarios such as collaborative object manipulation, where coordination and shared control are critical (see, for instance, [52]). By applying metrics such as gaze entropy and motion smoothness, the approach presented in this study could help quantify the efficiency and fluency of human–robot collaboration, offering a more comprehensive view of user adaptation and system responsiveness.

These findings underscore the value of the proposed metrics not only for operator benchmarking in remote manipulation but also for advancing the design and validation of intelligent, human-in-the-loop robotic systems in broader domains.

#### *Potential Limitations*

This study presents findings that should be considered alongside certain limitations related to participant experience, experimental setup, and the scope of evaluated metrics.

One key limitation is the relatively small number of expert operators available for participation. The participant pool primarily consisted of RACE operators and staff with varying levels of teleoperation experience, which may not fully capture the diversity of expertise found in broader industrial or field settings. A larger and more varied sample, including operators from different domains, could enhance the generalizability of the results.

Additionally, the experimental setup was conducted in a controlled training facility, where the robotic arms were separated by a fence and covered by a curtain. While this setup aimed to simulate real-world conditions, it does not fully replicate the operational

complexity of actual teleoperation control rooms, which often involve additional supervision protocols, communication constraints, and environmental stressors. These factors could significantly influence teleoperator performance and workload, aspects not fully captured in this study.

Finally, the study focused on a limited set of performance and physiological metrics. While the inclusion of gaze entropy measures provided novel insights into visual attention strategies, a broader range of metrics could offer a more comprehensive assessment of teleoperation under varying workload levels. Additional physiological indicators, such as cardiovascular responses and muscle activity through electromyography (EMG), could provide further insights into the cognitive and physical demands of teleoperation. Future studies should consider incorporating these factors to develop a more holistic understanding of operator performance.

## 6. Conclusions

This study underscores the profound impact of expertise on teleoperation performance, particularly in gaze control, task adaptability, and perceived workload. By integrating gaze entropy analysis with task difficulty metrics and subjective feedback, a more holistic approach to evaluating operator performance emerges. These insights can inform the development of targeted training programs aimed at improving novice adaptability and efficiency in complex tasks.

Future research could expand upon these findings by incorporating a larger and more diverse participant pool and extending the range of tasks to further validate gaze entropy as a robust metric for assessing operator skill. Additionally, refining training protocols (e.g., focusing on enhancing manipulator control efficiency and encouraging multi-viewpoint visual strategies) based on these insights may enhance skill acquisition and operational efficiency in teleoperation and other remote robotic applications.

**Author Contributions:** H.T., F.A., M.S. and S.S., conceived and designed the study and performed experimental testing and data collection. H.T., I.T., F.A. and M.S. conducted the data analysis, and H.T. wrote the article. I.C. and R.S. participated in proofreading. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the LongOps programme through UKRI (Project Reference 107463), the Nuclear Decommissioning Authority (NDA), and Tokyo Electric Power Company (TEPCO). It was also supported by the UKAEA/EPSC Fusion Grant 2022/27 (EP/W006839/1), which enabled the utilisation of related work for the decommissioning of fusion devices. The views and opinions expressed herein do not necessarily reflect those of the funding organisations.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Acknowledgments:** The authors would like to thank the anonymous expert operators and operators who underwent training for the teleoperation and volunteered to participate in the experiments.

**Conflicts of Interest:** Author Fumiaki Abe, Masaki Sakamoto, Shu Shirai were employed by the company Tokyo Electric Power Company (TEPCO). The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## References

1. Gong, X.; Wang, L.; Mou, Y.; Wang, H.; Wei, X.; Zheng, W.; Yin, L. Improved four-channel PBTDPA control strategy using force feedback bilateral teleoperation system. *Int. J. Control. Autom. Syst.* **2022**, *20*, 1002–1017. [\[CrossRef\]](#)
2. Su, H.; Qi, W.; Yang, C.; Sandoval, J.; Ferrigno, G.; Momi, E.D. Deep neural network approach in robot tool dynamics identification for bilateral teleoperation. *IEEE Robot. Autom. Lett.* **2020**, *5*, 2943–2949. [\[CrossRef\]](#)
3. Riaziat, N.D.; Erin, O.; Krieger, A.; Brown, J.D. Investigating haptic feedback in vision-deficient millirobot telemanipulation. *IEEE Robot. Autom. Lett.* **2024**, *9*, 6178–6185. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Samur, E. *Performance Metrics for Haptic Interfaces*; Springer Science & Business Media: London, UK, 2012.
5. Chen, J.; Cheng, N.; Cacciamani, G.; Oh, P.; Lin-Brandt, M.; Remulla, D.; Gill, I.S.; Hung, A.J. Objective Assessment of Robotic Surgical Technical Skill: A Systematic Review. *J. Urol.* **2019**, *201*, 461–469. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Kutana, S.; Bitner, D.P.; Addison, P.; Chung, P.J.; Talamini, M.A.; Filicori, F. Objective assessment of robotic surgical skills: Review of literature and future directions. *Surg. Endosc.* **2022**, *36*, 3698–3707. [\[CrossRef\]](#)
7. Collins, S.; Wilkinson, J.; Thomas, J. Remote handling operator training at JET. In Proceedings of the 11th International Symposium on Fusion Nuclear Technology, Barcelona, Spain, 16–20 September 2013.
8. Norton, A.; Ober, W.; Baraniecki, L.; McCann, E.; Scholtz, J.; Shane, D.; Skinner, A.; Watson, R.; Yanco, H. Analysis of human–robot interaction at the DARPA Robotics Challenge Finals. *Int. J. Robot. Res.* **2017**, *36*, 483–513. [\[CrossRef\]](#)
9. Liu, X.; Zhang, Y.; Jiang, X.; Zheng, B. Human eyes move to the target earlier when performing an aiming task with increasing difficulties. *Int. J. Hum.–Comput. Interact.* **2023**, *39*, 1341–1346. [\[CrossRef\]](#)
10. Albert, B.; Tullis, T. *Measuring the User Experience: Collecting, Analyzing, and Presenting ux Metrics*; Interactive Technologies; Elsevier Science: Amsterdam, The Netherlands, 2022.
11. Hauser, K.; Watson, E.N.; Bae, J.; Bankston, J.; Behnke, S.; Borgia, B.; Catalano, M.G.; Dafarra, S.; van Erp, J.B.F.; Ferris, T.; et al. Analysis and perspectives on the ana avatar xprize competition. *Int. J. Soc. Robot.* **2024**, *17*, 473–504. [\[CrossRef\]](#)
12. Tugal, H.; Abe, F.; Sakamoto, M.; Shirai, S.; Caliskanelli, I.; Skilton, R. Factors Influencing Operator Expertise in Bilateral Telerobotic Operations : A User Study. In Proceedings of the IEEE 18th International Conference on Control, Automation, Robotics and Vision (ICARCV), Dubai, United Arab Emirates, 12–15 December 2024; pp. 697–703.
13. Welford, A.T. The measurement of sensory-motor performance: Survey and reappraisal of twelve years' progress. *Ergonomics* **1960**, *3*, 189–230. [\[CrossRef\]](#)
14. Fitts, P.M. The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* **1954**, *47*, 381–391.
15. Xiao, H.; Sun, Y.; Duan, Z.; Huo, Y.; Liu, J.; Luo, M.; Li, Y.; Zhang, Y. A study of model iterations of fitts' law and its application to human–computer interactions. *Appl. Sci.* **2024**, *14*, 7386. [\[CrossRef\]](#)
16. Hannaford, B.; Wood, L.; Mcaffee, D.A.; Zak, H. Performance evaluation of a six-axis generalized force-reflecting teleoperator. *IEEE Trans. Syst. Man Cybern.* **1991**, *21*, 620–633. [\[CrossRef\]](#)
17. Tugal, H.; Gautier, B.; Tang, B.; Nabi, G.; Erden, M.S. Hand-impedance measurements with robots during laparoscopy. *Robot. Autom. Syst.* **2022**, *154*, 104130. [\[CrossRef\]](#)
18. Wang, Z.; Fey, A.M. Human-centric predictive model of task difficulty for human-in-the-loop control tasks. *PLoS ONE* **2018**, *13*, e0195053. [\[CrossRef\]](#)
19. Kourtesis, P.; Vizcay, S.; Marchal, M.; Pacchierotti, C.; Argelaguet, F. Action-specific perception & performance on a fitts's law task in virtual reality: The role of haptic feedback. *IEEE Trans. Vis. Comput. Graph.* **2022**, *28*, 3715–3726. [\[CrossRef\]](#)
20. Hwang, H.; Lim, J.; Kinnaird, C.; Nagy, A.G.; Panton, O.N.M.; Hodgson, A.J.; Qayumi, K.A. Correlating motor performance with surgical error in laparoscopic cholecystectomy. *Surg. Endosc. Other Interv. Tech.* **2006**, *20*, 651–655. [\[CrossRef\]](#)
21. Nisky, I.; Okamura, A.M.; Hsieh, M.H. Effects of robotic manipulators on movements of novices and surgeons. *Surg. Endosc.* **2014**, *28*, 2145–2158. [\[CrossRef\]](#)
22. Aghazadeh, F.; Zheng, B.; Tavakoli, M.; Rouhani, H. Motion smoothness-based assessment of surgical expertise: The importance of selecting proper metrics. *Sensors* **2023**, *23*, 3146. [\[CrossRef\]](#)
23. Aghazadeh, F.; Zheng, B.; Tavakoli, M.; Rouhani, H. Surgical tooltip motion metrics assessment using virtual marker: An objective approach to skill assessment for minimally invasive surgery. *Int. J. Comput. Assist. Radiol. Surg.* **2023**, *18*, 2191–2202. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Wu, C.; Cha, J.; Sulek, J.; Zhou, T.; Sundaram, C.P.; Wachs, J.; Yu, D. Eye-tracking metrics predict perceived workload in robotic surgical skills training. *Hum. Factors* **2020**, *62*, 1365–1386. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Khan, R.S.A.; Tien, G.; Atkins, M.S.; Zheng, B.; Panton, O.N.M.; Meneghetti, A.T. Analysis of eye gaze: Do novice surgeons look at the same location as expert surgeons during a laparoscopic operation? *Surg. Endosc.* **2012**, *26*, 3536–3540. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Li, S.; Duffy, M.C.; Lajoie, S.P.; Zheng, J.; Lachapelle, K. Using eye tracking to examine expert-novice differences during simulated surgical training: A case study. *Comput. Hum. Behav.* **2023**, *144*, 107720. [\[CrossRef\]](#)

27. Li, Y.; Reed, A.; Kavoussi, N.; Wu, J.Y. Eye gaze metrics for skill assessment and feedback in kidney stone surgery. *Int. J. Comput. Assist. Radiol. Surg.* **2023**, *18*, 1127–1134. [[CrossRef](#)]
28. Bruder, C.; Hasse, C. Differences between experts and novices in the monitoring of automated systems. *Int. J. Ind. Ergon.* **2019**, *72*, 1–11. [[CrossRef](#)]
29. Ayala, N.; Zafar, A.; Kearns, S.; Irving, E.; Cao, S.; Niechwiej-Szwedo, E. The effects of task difficulty on gaze behaviour during landing with visual flight rules in low-time pilots. *J. Eye Mov. Res.* **2023**, *16*, 1–16. [[CrossRef](#)]
30. Cao, S.; Samuel, S.; Murzello, Y.; Ding, W.; Zhang, X.; Niu, J. Hazard perception in driving: A systematic literature review. *Transp. Res. Rec.* **2022**, *2676*, 666–690. [[CrossRef](#)]
31. Arias-Portela, C.Y.; Mora-Vargas, J.; Caro, M.P. Situational awareness assessment of drivers boosted by eye-tracking metrics: A literature review. *Appl. Sci.* **2024**, *14*, 1611. [[CrossRef](#)]
32. Versaci, M.; Laganà, F.; Manin, L.; Angiulli, G. Soft computing and eddy currents to estimate and classify delaminations in biomedical device CFRP plates. *J. Electr. Eng.* **2025**, *76*, 72–79. [[CrossRef](#)]
33. Shannon, C.E. The mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
34. Tuğal, İ.; Karıcı, A. Comparisons of Karıcı and Shannon entropies and their effects on centrality of social networks. *Phys. A Stat. Mech. Its Appl.* **2019**, *523*, 352–363. [[CrossRef](#)]
35. Krejtz, K.; Duchowski, A.; Szmids, T.; Krejtz, I.; González Perilli, F.; Pires, A.; Vilaro, A.; Villalobos, N. Gaze transition entropy. *ACM Trans. Appl. Percept.* **2015**, *13*, 1–20. [[CrossRef](#)]
36. Shiferaw, B.; Downey, L.; Crewther, D. A review of gaze entropy as a measure of visual scanning efficiency. *Neurosci. Biobehav. Rev.* **2019**, *96*, 353–366. [[CrossRef](#)]
37. Riley, J.M.; Kaber, D.B.; Draper, J.V. Situation awareness and attention allocation measures for quantifying telepresence experiences in teleoperation. *Hum. Factors Ergon. Manuf. Serv. Ind.* **2004**, *14*, 51–67. [[CrossRef](#)]
38. Stubbings, L.; Chaboyer, W.; McMurray, A. Nurses' use of situation awareness in decision-making: An integrative review. *J. Adv. Nurs.* **2012**, *68*, 1443–1453. [[CrossRef](#)]
39. Huggins, A.; Claudio, D. A performance comparison between the subjective workload analysis technique and the NASA-TLX in a healthcare setting. *IIEE Trans. Healthc. Syst. Eng.* **2018**, *8*, 59–71. [[CrossRef](#)]
40. von Janczewski, N.; Kraus, J.; Engeln, A.; Baumann, M. A subjective one-item measure based on NASA-TLX to assess cognitive workload in driver-vehicle interaction. *Transp. Res. Part F Traffic Psychol. Behav.* **2022**, *86*, 210–225. [[CrossRef](#)]
41. Nguyen, T.; Lim, C.P.; Nguyen, N.D.; Gordon-Brown, L.; Nahavandi, S. A review of situation awareness assessment approaches in aviation environments. *IEEE Syst. J.* **2019**, *13*, 3590–3603. [[CrossRef](#)]
42. Kaber, D.B.; Onal, E.; Endsley, M.R. Design of automation for telerobots and the effect on performance, operator situation awareness, and subjective workload. *Hum. Factors Ergon. Manuf. Serv. Ind.* **2000**, *10*, 409–430. [[CrossRef](#)]
43. Yang, E.; Dorneich, M.C. The emotional, cognitive, physiological, and performance effects of variable time delay in robotic teleoperation. *Int. J. Soc. Robot.* **2017**, *9*, 491–508. [[CrossRef](#)]
44. Bolton, M.; Biltekoff, E.; Humphrey, L. The level of measurement of subjective situation awareness and its dimensions in the situation awareness rating technique (SART). *IEEE Trans.-Hum.-Mach. Syst.* **2022**, *52*, 1147–1154. [[CrossRef](#)]
45. Hart, S.G. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2006**, *50*, 904–908. [[CrossRef](#)]
46. Braarud, P.Ø. Investigating the validity of subjective workload rating (NASA TLX) and subjective situation awareness rating (SART) for cognitively complex human-machine work. *Int. J. Ind. Ergon.* **2021**, *86*, 103233. [[CrossRef](#)]
47. Read, J.C.; Begum, S.F.; McDonald, A.; Trowbridge, J. The binocular advantage in visuomotor tasks involving tools. *i-Perception* **2013**, *4*, 101–110. [[CrossRef](#)] [[PubMed](#)]
48. Fang, Y.; Qi, J.; Hu, J.; Wang, W.; Peng, Y. An approach for jerk-continuous trajectory generation of robotic manipulators with kinematical constraints. *Mech. Mach. Theory* **2020**, *153*, 103957. [[CrossRef](#)]
49. Hamilton, D.; Preece, G. Development of the MASCOT Telem manipulator Control System; Technical Report 01. 2014. Available online: <https://scipub.euro-fusion.org/wp-content/uploads/2014/11/EFD01006.pdf> (accessed on 12 August 2024).
50. Tugal, H.; Abe, F.; Sakamoto, M.; Shirai, S.; Caliskanelli, I.; Liu, W.; Marin-Reyes, H.; Zhang, K.; Skilton, R. Nuclear bilateral telerobotic systems: Performance comparison and future implications. *Robotica* **2024**, *43*, 242–255. [[CrossRef](#)]
51. Maher, J.M.; Markey, J.C.; Ebert-May, D. The other half of the story: Effect size analysis in quantitative research. *CBE Life Sci. Educ.* **2013**, *12*, 345–351. [[CrossRef](#)]
52. Xing, X.; Burdet, E.; Si, W.; Yang, C.; Li, Y. Impedance learning for human-guided robots in contact with unknown environments. *IEEE Trans. Robot.* **2023**, *39*, 3705–3721. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.