Comparing Alternate Modes of Teleoperation for Constrained Tasks

Christopher E. Mower[1], Wolfgang Merkt[1], Aled Davies[2], and Sethu Vijayakumar[1]

***Abstract*— Teleoperation of heavy machinery in industry often requires operators to be in close proximity to the plant and issue commands on a per-actuator level using multiple joystick input devices. However, this is non-intuitive and makes achieving desired job properties a challenging task requiring operators to complete costly and extensive training. Despite this, operator fatigue is common with implications for personal safety, project timeliness, cost, and job quality. While full automation is not yet achievable due to unpredictability and the dynamic nature of the environment and task, shared control paradigms allow operators to issue high-level commands in an intuitive, task-informed control space while having the robot optimize for achieving desired job properties.**

In this paper, we compare a number of modes of teleoperation, exploring both the number of dimensions of the control input as well as the most intuitive control spaces. Our experimental evaluations of the performance metrics were based on quantifying the difficulty of tasks based on the well known Fitts' law as well as a measure of how well constraints affecting the task performance were met. Our experiments show that higher performance is achieved when humans submit commands in low-dimensional task spaces as opposed to joint space manipulations.

I. Introduction

Teleoperation of industrial manipulators is generally repetitive and requires high levels of concentration and manual dexterity. Excessive cognitive loads invariably lead to fatigue that can become dangerous. This danger is prevalent in the construction sector seen by having the highest levels of incidents involving fatalities per annum in the United Kingdom [1].

An example of such a task is concrete spraying as shown in Fig. 1. Here, a skilled human operator manipulates the device via some interface to spray a lining of wet concrete onto a excavated tunnel surface. The tunnel surface is often unstructured, due to excavation, and the operators visibility is restricted by high amounts of dust. Despite these restrictions, operators are required to manipulate these devices so that they (1) ensure job-site safety, (2) achieve high task performance, and (3) minimize costs [2]. Simultaneously accounting for safety and task performance in teleoperation tasks is required in a variety of applications: nuclear waste disposal [3], space robotics [4], and subsea [5].

Devices, such as the concrete sprayer systems, are generally controlled by a human operator on a per-actuator level and consequently these control architectures do not lend themselves to easy operation since they force the operator to submit commands directly in the joint space. Since humans typically model tasks in the three-dimensional Cartesian space they must naturally learn an inverse kinematic mapping imposing high monetary costs and time for specialized training. In the literature, methods such as the inverse-kinematics method [6] and optimization techniques [7] have been developed that allow control commands to be submitted in alternative control spaces. A number of works leveraging these advancements have developed assistive techniques for teleoperation, for example, virtual fixtures [8], shared control frameworks that merge human input and autonomy [3], and human supervisory capabilities implementing sliding autonomy [9]. However, for assistive techniques, in general, the issue of sub-task allocation between the human and the device remains one of the main challenges [10].

Fig. 1: Shotcrete application in a freshly excavated tunnel using a 5-DoF concrete spraying unit. Image provided by Costain Laing O'Rourke Joint Venture.

In this work, we investigate the allocation of sub-tasks in the context of control spaces for target acquisition tasks. That is, we ask which control spaces should a human operator submit commands in order to achieve high task performance? Our main source of inspiration comes from the concrete spraying example [11] however our investigation can be generalized to other tasks such as paint spraying [12], semi-autonomous grasping [13], and wiping [14]. We have designed an experiment incorporating the well-known Fitts' law, that quantifies task difficulty, in order to compare and contrast different modes of teleoperation. This paper identifies intuitive and effective control spaces for teleoperation and shared control that will allow for more grounded formulations. Our main contributions are listed as follows:

1) Analysis of the sub-task allocation for target acquisition tasks through an empirically driven investigation.

[1]School of Informatics, University of Edinburgh, UK. Email: chris.mower@ed.ac.uk

[2]Infrastructure division, Costain Group PLC, UK.

This research is supported by the Engineering and Physical Sciences Research Council (EPSRC, grant reference EP/L016834/1) and EU H2020 project Memory of Motion (MEMMO, project ID: 780684). Christopher E. Mower is supported by EPSRC and Costain Group PLC.

2) An extensive data set of results that contains objective and subjective metrics.
3) A comparison between two subsets of our participants based on the participants personal habits that we identified to effect performance.
4) Two generalized performance metrics for Fitts' law relating to teleoperated target acquisition tasks.

The resulting data support several conclusions about control, user preferences, and how habitual traits may effect task performance. Our analysis indicates that there is a relation between control space-dimensionality and performance. It has been seen in our experiments that lower-dimensional task spaces, in general, elicit highest performance and these are generally preferred by users. Users who were identified to play video games on a regular basis were generally able to complete tasks quicker than those who do not. However, those who do not play video games were seen to maintain other performance metrics to a higher standard. An accompanying video is available at `https://youtu.be/SWWtT1vHnsM`.

II. Related Work

Prior work has explored techniques that guide robots motions via synthesized constraints, i.e., potential fields and virtual fixtures. Potential fields, originally developed by Khatib [15], guide a user towards or away from a goal or obstacle [16]. Virtual fixtures create guidance and/or forbidden regions in task space [8]. Later work [17] has developed methods for on-the-fly virtual fixture synthesis with real-time sensory data.

The work of Dragan and Srinivasa [18], [19] formalizes assistive teleoperation under a framework of policy blending. The system, grounded in inverse reinforcement learning, attempts to predict the intentions of the human operator in order to arbitrate the operator and the predicted robot control policies. Other learning-based methods have also been developed by Abi-Farraj et al. [20] and attempt to refine unskilled operator input based on skilled operator input learned by exploiting *learning from demonstration* techniques.

From the concrete spraying literature, a task space control framework was developed by Honegger and Codourey [21] based on the inverse-kinematics method. The system allows the operator to control the end-effector of the concrete spraying unit using a 6-DoF space mouse. Later work by Girmscheid and Moser [11] developed a sliding autonomy approach for a high-DoF concrete spraying unit. They define three levels of autonomy: (1) a manual mode, consisting of no autonomy and implementing common joint-level control, (2) a semi-automated mode that, using a pre-collected laser scan of a tunnel cavity and computed geometry, allows the operator to command only a target position along the wall whilst the system accounts for motion constraints, and (3) an automated mode that plans and executes an entire trajectory. The fully automated mode, they note, is only in development for tunnel-boring machine projects where the excavated tunnel surface is regular and very smooth. Modes of teleoperation compared in this work are inspired by these control techniques developed for concrete spraying applications.

A number of strategies with varying levels of assistance were compared by You and Hauser [22] for a reach-to-target task that must deal with collision avoidance, dynamic constraints, and erroneous input. Their user study showed participants preferred those strategies with higher levels of autonomy. Kim et al. [23] compared manual and autonomous control modes for a pick-and-place task. The results of their experiments concluded that in-fact participants preferred less autonomy. A study by Leeper et al. [24] compares a number of strategies for assisted and non-assisted remote robot grasping. These studies compare various methods for varying levels of autonomy and interface designs. The conflicting results for pick-and-place tasks provides motivation for our work on related tasks (i.e., target acquisition tasks). An investigation and analysis to provide an explanation for these contradictory results is warranted, however this is out-of-scope of this work.

To the best of our knowledge, there has not been work in educing control spaces that are intuitive and elicit high task performance for unskilled operators. In this work, we use the well-known Fitts' law to quantify task difficulty in order to systematically compare the performance of several modes of teleoperation.

III. Methodology

A. General formulation

We assume a teleoperation setting akin to Fig. 1, i.e., a human operates a manipulation system via some interface to achieve some task. Let us model the process of teleoperation by the functional

$$x_{t+1} = f(x_t, h_t) \tag{1}$$

where $x_t \in \mathcal{C}^n$ represents joint configuration for an n-DoF manipulator, $h_t \in \mathcal{H}^m$ are inputs from an m-dimensional interface, and t denotes time. Note, we assume $\mathcal{C}^n$ and $\mathcal{H}^m$ are closed sub-spaces, with control inputs scaled to the range $[-1, 1]$. Assuming a known initial state x_0, all future state can thus be computed using (1). It is also assumed that the function f can be computed fast enough for real-time actuation.

The chosen value for m coupled with the space in which commands h_t are submitted categorizes what we define as the *mode of teleoperation*.

B. Hypotheses

In this section we state our hypotheses. As mentioned in Section I, humans typically model tasks in three-dimensional Cartesian spaces as opposed to the manipulator joint configuration space and thus, intuitively, we expect that modes of teleoperation that allow control commands to be submitted in task spaces will achieve higher performance. Therefore, we base our hypotheses upon this intuition.

Hypothesis 1 (H1): *Modes of teleoperation with lower dimensions will see higher performance.*

Hypothesis 2 (H2): *Modes of teleoperation that submit commands in the manipulator task space will see higher performance.*

Hypothesis 3 (H3): *Participants will rate higher those modes of teleoperation with fewer dimensions.*

Hypothesis 4 (H4): *Participants will rate higher those modes of teleoperation that submit commands in the manipulator task space.*

C. Modes of teleoperation

1) Full joint mode (FJ): Industrial manipulators are operated on a per-actuator level. That is, there exists a one-to-one mapping between each joystick axis and manipulator actuator. Thus, this mode of teleoperation can be modeled using

$$f(x_t, h_t) := x_t + \Delta t A h_t \tag{2}$$

where A is a diagonal matrix such that the n entries are maximum joint velocities, and Δt is the control loop time step. Joint limits are handled by checking (2) prior to sending actuation signals to the joints and clipping the computed values if necessary.

2) Reduced joint mode (RJ): Here, we relieve the control of a number of joints from the human and assign these as a task constraint. Let $x_t = [x_t^{(1)}, x_t^{(2)}]^T$ where $x_t^{(1)} \in \mathcal{C}^m$ denote the joints under human control, and $x_t^{(2)} \in \mathcal{C}^{n-m}$ denote the autonomous joints. As an example and inspired by the concrete spraying task, joints $x_t^{(2)}$ could be assigned the task of end-effector alignment with the surface normal with respect to some pre-collected model of the environment e. We model the mode using

$$f(x_t, h_t) := \left[f^{(1)}\left(x_t^{(1)}, h_t\right), f^{(2)}\left(x_t^{(2)}; e\right) \right]^T \tag{3}$$

where $f^{(1)}$ represents the next joint state for joints $1{:}m$ and similarly for $f^{(2)}$. The human controlled joints are computed in the same way as in the full joint control mode and thus, $f^{(1)}$ takes the same form as (2). The functional form for $f^{(2)}$ is found using an unconstrained optimization expressed by

$$f^{(2)}\left(x_t^{(2)}; e\right) := \underset{x^{(2)} \in \mathcal{C}^{n-m}}{\arg\min} \left\| \phi\left(x^{(2)}\right) - y^*(e) \right\|^2 \tag{4}$$

where $\phi(\cdot)$ is a mapping from the configuration space to the k-dimensional task space, and $\|\cdot\|$ denotes the Euclidean norm, y^* represents the task space goal.

3) Full task mode (FT): An alternative class of modes is defined when the human instead submits commands to the manipulator task space rather than the joint space. For spraying, the manipulator task space comprises three translational and two rotational dimensions of the end-effector. This mode is thus expressed by

$$f(x_t, h_t) := \underset{x \in \mathcal{C}^n}{\arg\min} \left\| \phi(x) - h_t \right\|^2 \tag{5}$$

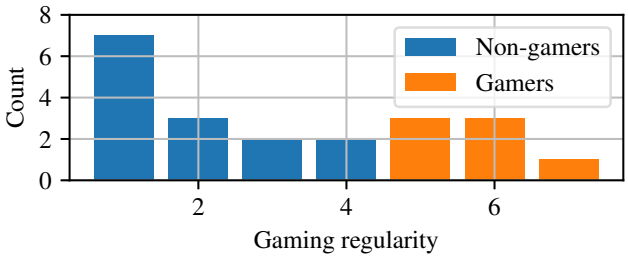


Fig. 2: Participant field habitual traits.

where $\phi(\cdot)$ is the mapping from the configuration space to the translational and angular task space.

4) Reduced task mode (RT): Humans typically model manipulation tasks in the task space. In this mode, we relieve control of a number of task space dimensions from the human and allow these to be moderated by the autonomy. In order for the autonomy to perform this moderation, however, we must, again, assume some pre-collected model of the environment e. For example, a spraying task requires five task space dimensions, however, to a human this can be considered a two-dimensional task, i.e., the position on the wall to spray. In this example, we can thus allow the human to control a target position defined on a two-dimensional surface (i.e. the spraying surface that we assume is modeled by e) and pass control of the three remaining dimensions to the robot. We model this mode by

$$f(x_t, h_t) := \underset{x \in \mathcal{C}^n}{\arg\min} \left\| \phi(x) - y^*(h_t; e) \right\|^2 \tag{6}$$

where $y^* = y^*(h_t; e) \in \mathbb{R}^k$ is some k-dimensional task space goal computed using the m-dimensional input from the human such that $m < k$. As in the spraying example, the computed value of y^* can be used to encode ideals such as perpendicularity to the spraying plane and a given stand-off distance from the surface.

IV. Experiments

A. Participant selection

We have obtained results in this study by evaluating the performance of 21 participants (16 male, 5 female). The age distribution of the participants were 7 (21-25), 11 (26-30), 2 (31-35), and 1 (36+).

During preliminary investigations a difference in performance was noticed for participants who regularly played video games. During the final experiments, we asked participants to provide a rating on how often they play video games[†]. They were asked "How often do you play video games (e.g., Xbox, PS4, PC)? [1 Never], [2 Bi-monthly], [3 Monthly], [4 Bi-weekly], [5 Weekly], [6 Regularly (but not daily)], [7 Daily]". The responses are shown in Fig. 2. A participant was considered a *gamer* if they gave a rating greater than or equal to 4, and a *non-gamer* otherwise.

[†]Note, those who participated in the preliminary investigation were barred from the final experiments to avoid skill transfer.

B. Experimental design

Fitts' law is widely considered to be a robust measure of performance for target acquisition (pointing) tasks making the analogy between movement time and transmission of information [25]. A common usage of Fitts' law from the human-computer interaction literature is to compare the usability of computer input devices (e.g. a mouse, trackball, and a stylus with a tablet) [26]. The law makes two underlying assumptions: (1) task difficulty is linearly correlated with performance, and (2) a complete move is performed through a number of iterations of feedback-guided corrective sub-movements, i.e., the *deterministic iterative-corrections model* [27]. In this work, we use Fitts' law as a method to specify task difficulty.

Fitts' established the information capacity of the human motor system by deriving a model for the *index of performance* I_p (in $\mathrm{bit\,s^{-1}}$) expressed by

$$I_p = I_d/T \tag{7}$$

where I_d is the *index of difficulty* (in bit) and T (in seconds) is the average movement time. The index of performance is a metric that quantifies task performance; higher values for I_p indicate better performance. The index of difficulty is a metric that defines the task difficulty; higher values for I_d imply the task is more difficult. Under the deterministic iterative-corrections model and by analogy with Shannon's Theorem 17 [28] a formula for I_d is derived, see [27] for details, given by

$$I_d = \log_2(2D/W) \tag{8}$$

where D is the distance to a target and W is the width of the target, see Fig. 3b. Define a *condition* by the tuple (W, D). A difficult task (i.e. high I_d) can be seen when, for example, W is small and D is large, and an easy task (i.e. small I_d) when W is large and D is small.

Operators in industry must maintain a number of motion constraints to achieve high task performance. Inspired from the concrete spraying task, we define two additional performance metrics that generalize Fitts' law.

Define the *angular length* L_a as the total change in θ_t for a complete move and expressed by

$$L_a = \frac{1}{T}\int_0^T \theta_t dt \tag{9}$$

where θ_t as in Fig. 3d. This quantity describes how well a user is able to maintain perpendicularity to the wall during transitioning between one target and the next.

Define the *delta length* denoted L_δ as the least absolute deviation in the stand-off distance with respect to a given ideal stand-off distance δ^* over the duration of a complete move and expressed by

$$L_\delta := \frac{1}{T}\int_0^T |\delta_t - \delta^*| dt \tag{10}$$

where δ_t is the stand-off distance as in Fig. 3d.

For each mode of teleoperation, participants completed 17 conditions: 16 have been generated by scaling the condition values used in Fitts' original work [25] while one has been hand-tuned to add spread to the I_d values. The conditions used in this study are shown in Tab. I representing the range $I_d \in [1.3219, 5.9069]$.

TABLE I: Condition values used in experiments.

W (m)	D (m)
0.0125	0.0469, 0.0938, 0.1875, 0.3750
0.0500	0.0938, 0.3750, 0.7500
0.0250	0.0469, 0.0938, 0.1875, 0.7500
0.1000	0.1875, 0.3750, 0.7500
0.1500	0.1875, 0.3750, 0.7500

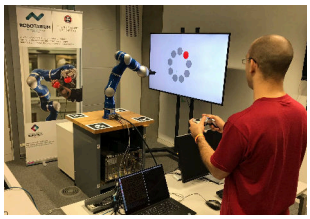
C. System description

We have implemented each mode of teleoperation described in Sec. III-C using an open-loop position control framework on a 7-DoF KUKA LWR robot arm. The operator interfaced with the system using an F710 Logitech gamepad. In Fig. 4, we show the mappings between the interface and robot for each mode. Each experiment was initiated by the user clicking a button on the gamepad. As a safety feature, in order to stream control commands to the system the user held down the top-left shoulder button akin to, but less straining, than a dead-man switch on a teleoperated machine. Interprocess communication is handled by the Robot Operating System (ROS) [29] and modes requiring numerical optimization were computed using the Extensible Optimization Toolkit (EXOTica) [30]. Goal joint positions were streamed at $100\,\mathrm{Hz}$. Targets and the focus point were displayed on a $1.65\,\mathrm{m}$ cross-diagonal display placed specifically such that the transform between the robot base and the television screen was known. Targets were scaled to the real world and the position of the focus point was found using the robot forward kinematics.

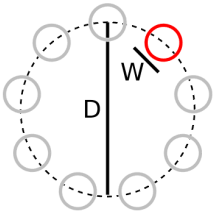
An ideal stand-off distance of $\delta^* = 0.55\,\mathrm{m}$ was chosen, justified as follows. In industry, the MEYCO Oruga (Fig. 1) has an approximate maximum reach of $r_O = 6\,\mathrm{m}$ and the ideal stand-off distance is $\delta^*_O = 2\,\mathrm{m}$. The KUKA LWR arm has an approximate maximum reach of $r_K = 1.5\,\mathrm{m}$. We thus find δ^*_K by balancing the ratios and adding a ten percent safety distance.

D. Experimental protocol

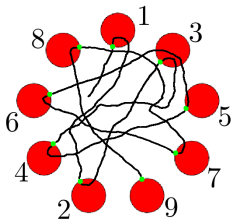
Participants were tasked with teleoperating a KUKA LWR robot arm to reach-and-point to a number of circular targets, indicated in red, arranged in a circle as in Fig. 3b. Targets were displayed on a screen in a known position and orientation with respect to the robot base, see Fig. 3a. For each condition, targets were presented in the order as in Fig. 3c. Conditions were randomized for every participant to minimize skill transfer. The participants were instructed to use as many controllable dimensions available to them in order to manipulate the robot in such a way that completes the task as fast as possible while maintaining the angular and delta length constraints; i.e., simultaneously minimize T, L_α, and L_δ to the best of their ability.



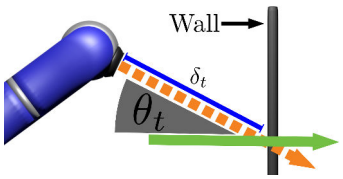
(a) An operator controlling the KUKA LWR arm.



(b) Multi-directional task.

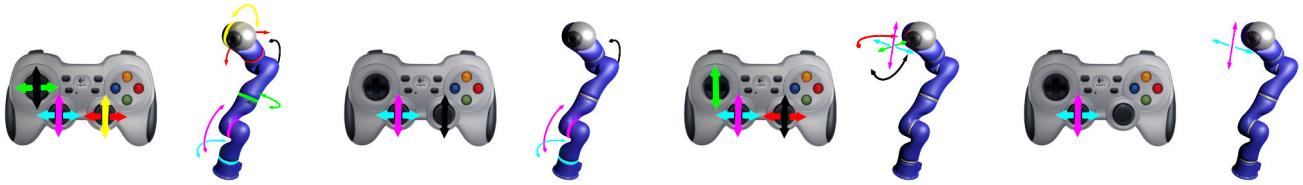


(c) Target order of appearance for a single condition.



(d) Angle to the plane θ_t and stand-off distance δ_t used to compute performance metrics.

Fig. 3: Experimental setup. Whilst condition order of appearance is randomized, target order was always kept the same as in (c). The black line indicates an example path of the focus point during manipulation; the path was not shown to participants, only the focus point indicated by the green dot.



(a) Full joint space control. (b) Reduced joint space control. (c) Full task space control. (d) Reduced task space control.

Fig. 4: Gamepad mappings for the modes of teleoperation implemented in experiments. Markings indicate where the human interacts with on the gamepad and where that interaction is perceived to be on the robot model.

Participants were allowed to move around the laboratory during the experiment, akin to concrete sprayer operators. For safety, they were not allowed within 1.5m of the robot.

Targets and conditions were presented to the participant in succession. As one target was deemed acquired the next immediately followed and as one condition was completed (i.e., one full cycle of targets in the order shown in Fig. 3c) the next immediately followed. Target acquisition is when the focus point (i.e., the point on the display screen the robot is pointing) comes into contact with the target. At the start of every mode the robot was reset to the same starting configuration. Following each mode participants were asked to fill out a questionnaire (Sec. IV-E).

E. Questionnaire

Participants were asked to provide a rating for each mode on their speed perception, accuracy perception, fatigue, and the mental capacity required. The questionnaire, shown in Tab. II, was devised from the ISO 9241 standard and work by Douglas et al. [31]. Questions 1-5 use a 7-point scale and 6-7 are open-ended.

TABLE II: Questionnaire used in investigation.

	Questions:	Rating:	1	7
1)	The mental effort required for operation was ..		high	low
2)	Accurate pointing was		difficult	easy
3)	Operation speed was		fast	slow
4)	Finger fatigue		high	none
5)	Overall, the mode of teleoperation was		difficult	easy
6)	Did you have any trouble with this mode of teleoperation?			
7)	Do you have any comments in general about using this mode of teleoperation?			

V. Results

A. Low-dimensional task space elicits high performance

Data for 21 participants was collected for 17 conditions with 9 targets for each condition. In order to compute our performance metrics we collected target acquisition timestamps and joint states.

We have filtered the data based on one criteria. Data for the first target of every condition was ignored since when transitioning between one condition and the next the data is not representative of skill in either condition. We have collected an extensive database of results constituting, in total, 8882 data points. Using each metric (time, angular length, and delta length) and the index of difficulty we compute an index of performance for every metric and for each point in our data set. The distributions of $\log(I_p)$ is shown in Fig. 5. Using this data we have been able to perform a one-sided paired-sampled t-test to test our hypotheses.

Precisely the same conclusions for time and delta length metrics have been observed. We see **H1** is accepted as expected for reducing control space-dimensionality, except the FJ-RJ comparison. For the angular length **H1** is accepted; this is unsurprising since RJ supports end-effector alignment. For cross-modes, e.g. FJ-RT, we see that **H1** is accepted.

When comparing task space modes and joint space modes we see that **H2** has been accepted. An interesting point to note for the time metric and delta length metric is that **H2** is accepted when comparing the RJ and FT modes. This suggests that low-dimensional joint spaces in some cases elicit worse performance than higher-dimensional task spaces. We postulate that the RJ mode may fight intuition

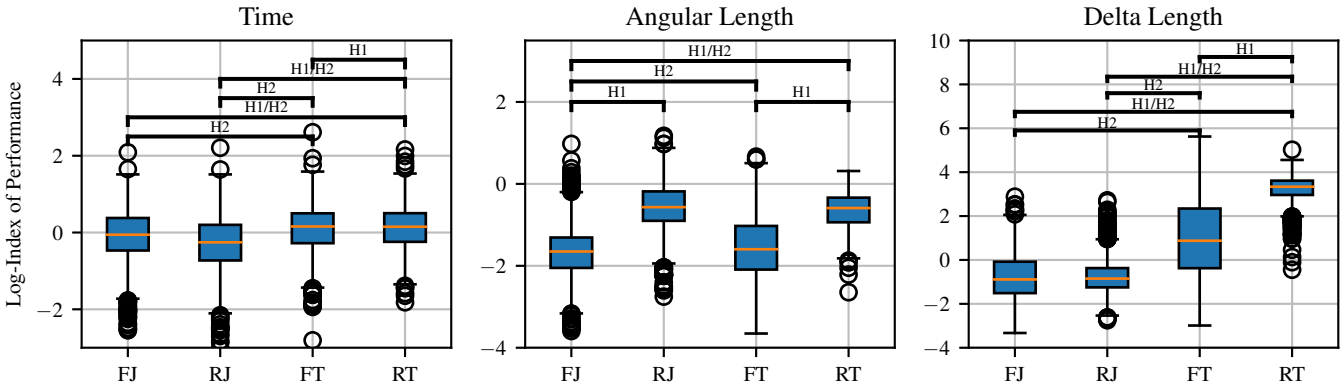


Fig. 5: Log-Index of performance distribution results. By means of a one-sided paired-sample t-test, pairs accepted under **H1** and **H2** ($\alpha = 0.01$) are indicated above.

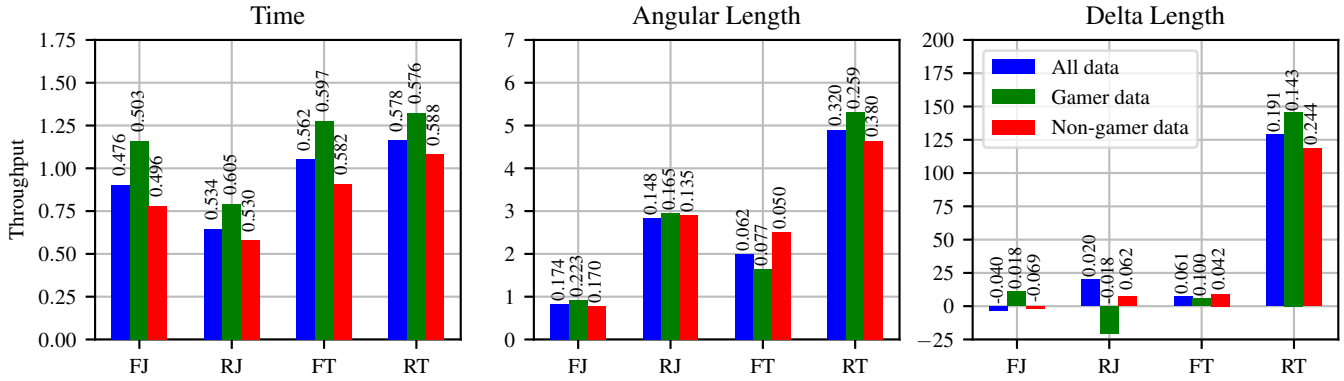


Fig. 6: Throughput results that estimate the index of performance. We report the results for the full data set and the two subsets: gamer and non-gamer. We also provide the Pearson correlation coefficient above each bar.

resulting in the machine performing motions that do not match the operators innate model of the system dynamics.

B. Habitual traits effect performance

During preliminary investigations, it was noted that participants who were known to play video games regularly seemed to give higher performance than those that did not. In our final experimental design we decided to include this comparison as part of our multivariate analysis. As described in Sec. IV-A, we asked each participant to indicate their gaming regularity and made the distinction between gamers and non-gamers. Using this we split our data set into two subsets.

By the assumption that difficulty and performance are linearly correlated we are able to compare across the various modes of teleoperation using linear regression. In order to compare the general performance for each trait we use an estimate of *throughput* [25]. The throughput is another indicator of overall performance. For the time metric, the throughput Th is estimated by the reciprocal of a scalar b where a, b are regression parameters such that $a + bI_d = T$; so $Th = 1/b$. For the angular length and delta length, throughput for each has been estimated in the same fashion. Throughput estimates for our full data set and the gamer/non-gamer subsets are shown in Fig. 6.

Regarding time performance, in general, excluding the result for the RJ mode, a similar pattern is seen when compared to the results in Fig. 5; the RT mode has highest throughput in general and task space modes have higher throughput than joint space. For each mode of teleoperation we observe a higher throughput for gamers than the non-gamers. We noted that gamers generally seemed more familiar using two analog joysticks at once as opposed to the non-gamers. We posit, due to this ability, gamers were able to achieve faster times.

Regarding angular length performance, we see the general trend in performance as discussed in the previous section. Gamers have higher throughput than non-gamers for each mode except the FT mode. It should be noted that the correlation coefficients for these are generally low which renders these results as potentially spurious. Drawing conclusions from these results may be unreliable. The throughput results for the delta length suffer from similar issues and so reliable conclusions unfortunately cannot be made for these either.

C. Participants approve low-dimensional task spaces

The results of the questionnaire are shown in Tab. II. The mean and standard deviation of the responses on questions 1-6 are shown above a paired-sampled t-test to determine

TABLE III: Results of the questionnaire and paired-sampled t-test ($\alpha = 0.05$). Bold indicates the one-sided hypothesis **H3/H4** is accepted. *Note, since Q3 does not evaluate the users preference (highlighted in bold) indicates the result of a two-sided significance test.

Question	FJ	RJ	FT	RT
1	1.81 ± 1.33	3.29 ± 1.59	3.57 ± 1.57	5.05 ± 1.36
2	2.57 ± 1.66	2.90 ± 1.73	3.14 ± 1.39	2.95 ± 1.69
3*	4.62 ± 0.92	4.71 ± 1.06	3.81 ± 1.08	4.29 ± 1.15
4	3.29 ± 2.00	3.62 ± 1.88	4.67 ± 1.43	4.90 ± 1.37
5	3.10 ± 1.55	4.24 ± 1.26	4.86 ± 1.28	5.95 ± 1.07

Question		RJ	FT	RT
1	FJ	**0.0048 (3)**	**0.0009 (3)**	**5.2e-7 (3/4)**
	RJ	-	0.5867	**0.0003 (3/4)**
	FT	-	-	**0.0011 (3)**
2	FJ	0.5538	0.1435	**0.0281 (3/4)**
	RJ	-	0.6339	**0.0104 (3/4)**
	FT	-	-	0.0727
3*	FJ	0.7245	**0.0202**	0.2596
	RJ	-	**0.0068**	0.1428
	FT	-	-	0.1158
4	FJ	0.4907	**0.0120 (3)**	**0.0044 (3/4)**
	RJ	-	**0.0100 (4)**	**0.0005 (3/4)**
	FT	-	-	0.3086
5	FJ	**0.0244 (3)**	**0.0009 (3)**	**3.3e-6 (3/4)**
	RJ	-	0.1198	**1.0e-5 (3/4)**
	FT	-	-	**0.0047 (3)**

the responses' statistical significance.

The results for question 1 show that participants felt the FJ mode required the highest amount of mental effort. There seems to be no statistical difference between the RJ and FT modes. Participants indicated the RT required the least amount of mental effort.

Overall, participants indicated the task space modes were the easiest to point accurately. There is no statistical difference between FJ and RJ modes both having ratings indicating that participants felt these were the most difficult to accurately point. The results of the t-test indicate we reject **H3/H4** when comparing FJ, RJ, and FT modes. However, whilst **H3/H4** are accepted when comparing between RT and the joint space modes, **H3** is rejected when comparing the task space modes. The p-value for the RJ-FT comparison is reasonably high and the FJ-FT comparison is higher than the p-value for the FT-RT comparison. Comparing the results of this question to the responses given for questions 6 and 7 (shown below), we suggest that, despite the computed p-value, accepting **H3** for the FT-RT comparison has potential grounding as a conclusion.

During our experiments, maximum joint velocities were reduced to a conservative range for safety. Question 3 attempted to ascertain whether the participants felt the robot motion was too slow or indeed too fast. Mean values indicate participants felt the operation speed was neither too fast nor too slow. There are not significant differences between the results apart from the FT mode compared to the joint space modes.

The results for question 4 indicate the participants experienced the least finger fatigue for task space control modes with no statistical difference between the two. The joint space modes caused the most fatigue.

For question 5, participants rated the RT mode as the easiest to use and the FJ the most difficult. There is no significant difference between the FT and RJ modes.

We summarize common responses to questions 6 and 7.

Full joint mode:

6) "*Wrist joints felt slower than base joint.*"
"*Joint mappings felt inverted and was easy to get into strange configurations.*"
7) "*Hard to maintain constraints.*"
"*Only used two joints at a time.*"

Reduced joint mode:

6) "*Attention was directed to focus point, not robot, making operation very difficult.*"
"*Often requiring to re-adjust.*"
7) "*Slightly better than full joint mode.*"
"*Easy for small D values.*"

Full task mode:

6) "*Unexpected motions at times.*"
"*Maintaining delta length was effectively impossible. Perhaps with more practice this mode would be more efficient.*"
7) "*Felt it was possible to go faster.*"
"*Favored over both joint space control modes; intuitive, mostly easy to use.*"

Reduced task control mode:

6) "*Some unexpected motions.*"
"*Sometimes felt too slow.*"
7) "*Very intuitive and easy to use.*"

Upon observation, the responses generally correlate with the answers to questions (1)-(5). The mappings for both joint modes were identified as inverted by many of the participants. The direction the joint moved under these modes was matched with the joystick direction. For example, for the FJ mode, pushing forward on the left joystick moved joint 2 in the direction such that the end-effector and focus point moved downwards. Participants felt the robot end-effector and focus point should instead move upwards. An issue for some participants for the task space modes is that the robot would make unexpected motions in certain configurations. The source of this issue is that at times a target configuration was computed requiring a joint velocity surpassing the maximum; an unfortunate consequence of unconstrained optimization. Whilst these issues were, at times, observed they did not render the task impossible to complete.

VI. Conclusions

In this paper, we developed, implemented, and carried out a study to determine which modes of teleoperation elicit high task performance for unskilled human operators on a task inspired by concrete spraying in industry. A Fitts' law paradigm was used to quantify difficulty and performance of each mode. We have generalized Fitts' law for two additional performance metrics and an extensive data set was collected from our experiments. The results and analysis performed support several conclusions regarding control, sub-task allocation between human and autonomy, and how habitual traits can effect performance.

A reduced task space control mode has been shown to outperform all three other modes of teleoperation with regards to three performance metrics; time, angular length,

and delta length. The results of the questionnaire support this conclusion as the RT mode was generally favored the highest. The RT mode is the closest model that directly regulates the performance parameters. Given that our hypotheses were accepted this opens new avenues for shared control design. Based on these results, we posit that there exists a positive correlation between cognitive load and number of human controllable dimensions for task space control modes. Both joint space modes were not favored by participants in this study with some noting that the reduced joint space control felt like having the "*worst of both worlds*" with regards to joint space and task space control.

This work assumed unskilled participants and so the long-term learning effects on performance are not considered here. In future work, we intend to study the effect on the participants learning curve in order to predict average learning time. The ability of the participants classed as gamers to achieve higher performance than non-gamers may be a consequence of their familiarity with the game pad controller used as the interface in our experiments. Future investigations will take into account the effect the interface used by comparing other interfaces such as a joystick, 6-DoF space mouse, and a combination of a computer mouse and keyboard. Metrics quantifying cognitive load will be explored and compared against the number of controllable dimensions in-order to investigate our proposition of a positive correlation between the two quantities.

We designed our experiment using a flat plane to represent a spraying surface. This design choice was necessary in our experimental setup which used a television screen. Otherwise, a realistic setup would be very challenging (e.g. displaying visuals on a curved surface). An alternative was to conduct the experiments in a simulated environment, however this would then provide the limitation that the user could not choose their viewpoint as in the concrete spraying example. Additionally, not allowing the user to adjust their viewpoint would bring into question the applicability of the angular and delta length metrics.

This study indicates the mode specification highly impacts the design and performance of shared control systems. We intend to use the knowledge acquired here to inform the development of new formulations for shared autonomous and collaborative methods that adapt dynamic motion constraints on-the-fly using multi-modal sensory data.

REFERENCES

[1] Health and Safety Executive (HSE), "Workplace fatal injuries in Great Britain 2018," *www.hse.gov.uk/statistics/pdf/fatalinjuries.pdf*, 2018.

[2] M. Ballou, "Shotcrete rebound – how much is enough?" *Shotcrete magazine, American Shotcrete Association*, 2003.

[3] F. Abi-Farraj, N. Pedemonte, and P. R. Giordano, "A visual-based shared control architecture for remote telemanipulation," in *IEEE/RSJ IROS*, Oct 2016, pp. 4266–4273.

[4] N. Y. Lii, Z. Chen *et al.*, "Toward understanding the effects of visual- and force-feedback on robotic hand grasping performance for space teleoperation," in *IEEE/RSJ IROS*, Oct 2010, pp. 3745–3752.

[5] R. R. Murphy, K. L. Dreger *et al.*, "Use of remotely operated marine vehicles at minamisanriku and rikuzentakata japan for disaster recovery," in *2011 IEEE Int. Symposium on Safety, Security, and Rescue Robotics*, Nov 2011, pp. 19–25.

[6] O. Khatib, "A unified approach for motion and force control of robot manipulators: The operational space formulation," *IEEE Journal on Robotics and Automation*, vol. 3, no. 1, pp. 43–53, February 1987.

[7] M. Zucker, N. Ratliff *et al.*, "CHOMP: Covariant Hamiltonian optimization for motion planning," *IJRR*, vol. 32, pp. 1164–1193, 2013.

[8] L. B. Rosenberg, "Virtual fixtures: Perceptual tools for telerobotic manipulation," in *IEEE Virtual Reality Annual Int. Symposium*, ser. VRAIS '93. IEEE Computer Society, 1993, pp. 76–82.

[9] W. Merkt, Y. Yang *et al.*, "Robust shared autonomy for mobile manipulation with continuous scene monitoring," in *IEEE CASE*, Aug 2017, pp. 130–137.

[10] T. Inagaki, "Adaptive automation: Sharing and trading of control," in *Handbook of cognitive task design*. CRC Press, 2003, pp. 171–194.

[11] G. Girmscheid and S. Moser, "Fully automated shotcrete robot for rock support," *Computer-Aided Civil and Infrastructure Engineering*, vol. 16, no. 3, pp. 200–215, 2001.

[12] H. Chen, W. Sheng *et al.*, "Automated robot trajectory planning for spray painting of free-form surfaces in automotive manufacturing," in *IEEE ICRA*, vol. 1. IEEE, 2002, pp. 450–455.

[13] J. Vogel, K. Hertkorn *et al.*, "Flexible, semi-autonomous grasping for assistive robotics," in *IEEE ICRA*, May 2016, pp. 4872–4879.

[14] L. Armesto, J. Moura *et al.*, "Constraint-aware learning of policies by demonstration," *IJRR*, vol. 37, no. 13-14, pp. 1673–1689, 2018.

[15] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," in *IEEE ICRA*, vol. 2, March 1985, pp. 500–505.

[16] J. W. Crandall and M. A. Goodrich, "Characterizing efficiency of human robot interaction: a case study of shared-control teleoperation," in *IEEE/RSJ IROS*, Sept 2002, pp. 1290–1295.

[17] S. Nia Kosari, F. Rydén *et al.*, "Forbidden region virtual fixtures from streaming point clouds," *Advanced Robotics*, vol. 28, no. 22, pp. 1507–1518, 2014.

[18] A. Dragan and S. Srinivasa, "Formalizing assistive teleoperation," in *Robotics: Science and Systems*, July 2012.

[19] A. D. Dragan and S. S. Srinivasa, "A policy-blending formalism for shared control," *IJRR*, vol. 32, no. 7, pp. 790–805, 2013.

[20] F. Abi-Farraj, T. Osa *et al.*, "A learning-based shared control architecture for interactive task execution," in *IEEE ICRA*, May 2017, pp. 329–335.

[21] M. Honegger and A. Codourey, "Redundancy resolution of a cartesian space operated heavy industrial manipulator," in *IEEE ICRA*, vol. 3, May 1998, pp. 2094–2098.

[22] E. You and K. Hauser, "Assisted teleoperation strategies for aggressively controlling a robot arm with 2d input," in *Proc. of Robotics: Science and Systems*, Los Angeles, CA, USA, June 2011.

[23] D. Kim, R. Hazlett-Knudsen *et al.*, "How autonomy impacts performance and satisfaction: Results from a study with spinal cord injured subjects using an assistive robot," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 42, no. 1, pp. 2–14, Jan 2012.

[24] A. Leeper, K. Hsiao *et al.*, "Strategies for human-in-the-loop robotic grasping," in *2012 7th ACM/IEEE Int. Conf. on Human-Robot Interaction (HRI)*, March 2012, pp. 1–8.

[25] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement." *Journal of experimental psychology*, vol. 47, no. 6, p. 381, 1954.

[26] I. S. MacKenzie, A. Sellen, and W. A. S. Buxton, "A comparison of input devices in element pointing and dragging tasks," in *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, ser. CHI '91. New York, NY, USA: ACM, 1991, pp. 161–166.

[27] E. Crossman and P. Goodeve, "Feedback control of hand-movement and fitts' law," *The Quarterly Journal of Experimental Psychology Section A*, vol. 35, no. 2, pp. 251–278, 1983.

[28] C. E. Shannon and W. Weaver, *A Mathematical Theory of Communication*. Champaign, IL, USA: University of Illinois Press, 1963.

[29] M. Quigley, K. Conley *et al.*, "ROS: an open-source Robot Operating System," in *ICRA Workshop on Open Source Software*. Kobe, Japan, 2009.

[30] V. Ivan, Y. Yang *et al.*, *EXOTica: An Extensible Optimization Toolset for Prototyping and Benchmarking Motion Planning and Control*. Cham: Springer Int. Publishing, 2019, pp. 211–240.

[31] S. A. Douglas, A. E. Kirkpatrick, and I. S. MacKenzie, "Testing pointing device performance and user assessment with the iso 9241, part 9 standard," in *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, ser. CHI '99. New York, NY, USA: ACM, 1999, pp. 215–222.