Continuous Gesture Control of a Robot Arm: Performance Is Robust to a Variety of Hand-to-Robot Maps

Steafan E. Khan and Zachary C. Danziger

Abstract*—*Objective: **Despite advances in human-machine-interface design, we lack the ability to give people precise and fast control over high degree of freedom (DOF) systems, like robotic limbs. Attempts to improve control often focus on the static map that links user input to device commands; hypothesizing that the user's skill acquisition can be improved by finding an intuitive map. Here we investigate what map features affect skill acquisition.** ***Methods:*** **Each of our 36 participants used one of three maps that translated their 19-dimensional finger movement into the 5 robot joints and used the robot to pick up and move objects. The maps were each constructed to maximize a different control principle to reveal what features are most critical for user performance. 1) Principal Components Analysis to maximize the linear capture of finger variance, 2) our novel Egalitarian Principal Components Analysis to maximize the equality of variance captured by each component and 3) a Nonlinear Autoencoder to achieve both high variance capture and less biased variance allocation across latent dimensions** ***Results:*** **Despite large differences in the mapping structures there were no significant differences in group performance.** ***Conclusion:*** **Participants' natural aptitude had a far greater effect on performance than the map.** ***Significance:*** **Robot-user interfaces are becoming increasingly common and require new designs to make them easier to operate. Here we show that optimizing the map may not be the appropriate target to improve operator skill. Therefore, further efforts should focus on other aspects of the robot-user-interface such as feedback or learning environment.**

***Index Terms*—Assistive robotic manipulator, human-machine interface, machine learning, teleoperation.**

I. Introduction

RAPID advancements in robotics have produced systems with many degrees of freedom (DOF) [1], [2], but these sophisticated robots demand new high-dimensional interfaces that let users engage with all the available DOFs at once, moving beyond the common single-input-single-output strategies that are currently in use [1], [3]. One possibility is to use an abstract map in which the user's inputs have greater dimensionality than the number of DOFs of the robot, i.e., a surjective map [4], [5]. The surjective map transforms the user's high-dimensional input into a lower-dimensional representation (equal in dimension to the number of DOFs of the robot) that is used for control [4], [5]. Surjective maps are attractive because the user input can be almost any high-dimensional signal the user can produce reliably (e.g., electromyographic activity, neural electrical activity, or body kinematics [4], [6], [7]) and the control can be straightforwardly calibrated to the abilities of an individual user.

Building a surjective user-to-device map that optimizes control performance is a common topic in the motor control literature [6], [8], [9], [10], [11], [12], although development has focused mainly on low-DOF devices (e.g., 2-D control of computer cursors, or powered wheelchairs [10], [12]) to understand basic principles of learning or to prototype designs in simpler environments. These studies typically rely on linear dimensionality reduction techniques like Principal Components Analysis (PCA) to create the surjective maps [4], [10], [11], [13]. PCA is used to identify a subspace within the user's high dimensional input space which captures most of the variance of their natural voluntary movements. The motivating hypothesis is that capturing a large amount of the user's natural input variance is important for control, but it is unclear if this hypothesis extends to control of high-DOF robots where investigations are more limited [14], [15]. In fact, the limited success that PCA maps have had in high-DOF contexts [14], [15] has recently led investigators to hypothesize that PCA may not be suitable for high DOF control because the way it allocates variance in the user's calibration data set is extremely biased toward a small subset of all available device DOFs [16]. To achieve both a more uniform distribution of variance across device DOFs and a high total variance captured investigators have begun using nonlinear dimensionality reduction algorithms such as autoencoders [16], [17], [18], [19]. Although it was shown in a small sample that it is *possible* to control a high-DOF device [16], it remains unclear how much a nonlinear map can improve user control over linear methods, if at all.

Our goal here is to determine 1) to what degree can participants using a nonlinear autoencoder map outperform those using a simpler PCA map in high-DOF robot control, and 2) to understand if any differences in performance are attributable to PCA's biased variance allocation. This addresses both the

Manuscript received 27 March 2023; revised 31 July 2023; accepted 28 September 2023. Date of publication 13 October 2023; date of current version 26 February 2024. This work was work supported by the National Science Foundation under Grant CMMI-2128465. *(Corresponding author: Zachary C. Danziger.)*

Steafan E. Khan is with the Florida International University, USA.

Zachary C. Danziger is with the Florida International University, Miami, FL 33174 USA (e-mail: zdanzige@fiu.edu).

Digital Object Identifier 10.1109/TBME.2023.3323601

above hypotheses that 1) the increased representational power of nonlinear maps will improve high DOF device control over linear maps and 2) biased input variance allocation along the DOFs impairs controllability.

To test these hypotheses, we compare the performance of healthy human participants operating a 5 DOF robot arm and gripper using different surjective maps between their hand finger joints and the robot joints. We utilize the participant's finger joints because they are a readily available source of high dimensional signal over which the user has precise control. We create the surjective maps using a Nonlinear Autoencoder (NLAE), Principal Components Analysis (PCA), or our novel Egalitarian Principal Components Analysis (EPCA). EPCA is a linear dimensionality reduction algorithm that trades off variance capture for equal variance distribution. Unlike PCA, which distributes variance in a rank-ordered fashion, EPCA distributes variance uniformly (removing the bias) but at the expense of the total variance captured. By comparing the performance of participants in the NLAE group to the performance of participants in the two linear groups (EPCA and PCA) we identify what (if any) performance gains result from using the nonlinear map. By comparing the performance of participants in the PCA and EPCA groups we identify whether the biased variance allocation hypothesis is true. If biased variance allocation is detrimental to robot control, then we would expect participants using EPCA maps to outperform those using PCA maps, but if total variance captured is more important then we would expect the reverse. A preliminary version of this work has been reported [20].

II. Methods

A. Experiment Setup

Participants wore a CyberGlove on their dominant hand. The CyberGlove captured the motion of their fingers and thumb via 19 resistive sensors. The 19-D vector of sensor values was mapped to the angles of the 5 revolute joints of the robotic arm (Interbotix WidowX250 5DOF) using 1 of 3 surjective mapping algorithms: a nonlinear autoencoder neural network, principal components analysis, or our novel egalitarian principal components analysis. Commands were issued to the robot continuously at a rate of $\sim$ 10 Hz.

Each surjective mapping algorithm mapped the 19-D vector of sensor values to a 5-D latent representation in which each element of the 5-D vector was used to specify the angle of one of the robot's revolute joints. To ensure the joint angle commands generated were within the allowed ranges of motion of the robot joints, each element of the 5-D vector was normalized to the interval (0, 1) and thus specified the joint angle command as a fraction of the robot joint's range of motion, plus a bias of the minimum allowed joint position.

In the formal expression in Fig. 1, M_{group} represents the surjective mapping algorithm (described further below), $J_{i,i}$ represents the range of motion of the i_{th} robot joint, S_i represents the i_{th} sensor value, and ϕ_{i} represents the minimum joint position of the i_{th} robot joint. Participants were informed that their hands would be mapped to commands for the robot but were not told how the mapping would be performed or what robot variables they would be controlling.

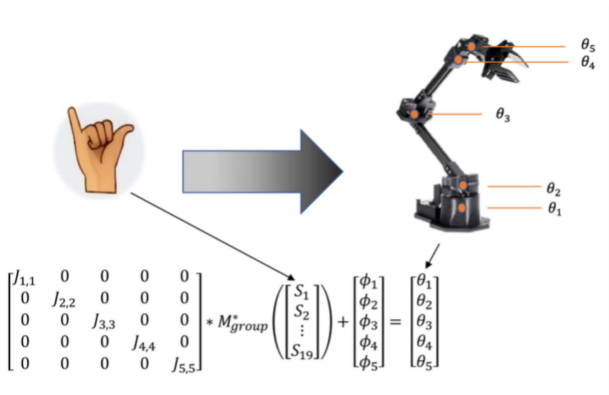

Fig. 1. Hand-to-robot map. Participants continuously controlled the robot configuration using the motion of their fingers and thumb. In each of our test groups the participant's hand was transformed into a robot command via a different surjective map M_{group}. Participants in all groups performed the same object pick-up-and-place task using the arm. $*$ Denotes the robot joint range of motion normalization, which is the same across all groups.

The robot gripper was also controlled by the user, but not via the surjective map M. Unlike the continuous arm joint angles, the gripper was commanded between only an open or closed state. Users toggled the gripper state by holding the robot still for 2 seconds.

Additionally, user commands were automatically checked to prevent the robot from colliding with the workspace table. Any command predicted to result in a collision was re-issued such that continuous operation was uninterrupted and stopped short of the surface. The robot commands were generated using the MATLAB (*Mathworks*) software and were sent to the ROS (*Robotics Operating System)* software which communicated with the physical robot [21].

B. Generation of the Surjective Maps

Calibration Data: For all groups, participants generated a calibration dataset to fit their map ($\mathrm{M}*_{\mathrm{group}}$) using the same protocol. The calibration set was obtained by having participants wiggle their fingers and thumb continuously for 30 seconds while the CyberGlove recorded their motions at approximately $\sim$10 Hz. This resulted in a calibration set of $\sim$300 hand postures that captured the user's natural posture tendencies and range of motion.

PCA Map: The surjective map for participants in the PCA group consisted of linearly projecting the participant's hand postures to a 5-dimensional latent space. The linear projection matrix, A, was created by solving the optimization problem posed in (1) by diagonalizing the data covariance matrix using PCA. X represents a 300 by 19 mean-centered matrix of sensor values from the calibration, the matrix norm is the Frobenius norm, and W is a 19 by 19 matrix of principal components. The first 5 PCs were used to form a 5 by 19 projection matrix A as depicted in (2).

$$W = \mathrm{argmin}_W \left(\left| \left| X^T - WW^T X^T \right| \right| \right) \tag{1}$$

$$A = \begin{bmatrix} W_{11} & W_{21} & \dots & W_{19,1} \\ \vdots & \vdots & \dots & \vdots \\ W_{15} & W_{25} & \dots & W_{19,5} \end{bmatrix} \quad (2)$$

The 5-D latent representation is therefore computed as the following:

$$M_{PCA}\ (s) = As \quad (3)$$

where A is the 5 by 19 projection matrix shown in (2), and s is the 19-D vector of sensor inputs.

EPCA Map: The surjective map for participants in the EPCA group was created by numerically solving the optimization problem posed in (4), which trades off reconstruction error in the calibration data (1st term) against the variability in the variance explained of the bases (2nd term). The optimization was performed under the constraint that the basis vectors of W have unit magnitude, which ensured that the EPCA algorithm did not rescale PCs to achieve an equal distribution of variance but rather solved for a completely different subspace. Here, X represents a 300 by 19 mean centered matrix of sensor values and W is a 5 by 19 projection matrix.

$$W = \text{argmin}_W \left(||X^T - W^+ W X^T|| + c\sqrt{\left(\sum_{i=1}^{r} \left(V_{ii} - \frac{1}{r} tr\,(V)\right)^2 \right)} \right) \quad (4)$$

where, $V = \frac{1}{n-1}(WX^T)(WX^T)^T$, $r = 5$ is the dimensionality of the latent space, $n = 300$, is the number of observations, and $c = 10$, is an empirically chosen constant. $+$ denotes the Moore-Penrose pseudoinverse.

$$M_{EPCA}\,(s) = Ws \quad (5)$$

Nonlinear Autoencoder Neural Network (NLAE): The surjective map for participants in the NLAE group was created by training an autoencoder neural network on the calibration dataset [22]. The initialization of the network weights was obtained by training the autoencoder network using example CyberGlove data consisting of ∼1700 hand postures collected across 6 individuals prior to the study (who did not participate in the study). The calibration dataset was partitioned into 70% training, 15% testing, and 15% validation datasets. The network was trained via Levenberg-Marquardt backpropagation until either the magnitude of the mean-squared-error gradient was less than 1E-7, or 1000 training iterations were performed.

The network architecture is the following:

Encoder:

$$layer_1 = \tanh\,(w_1 s + b_1) \quad (6)$$

$$layer_2 = w_2\ layer_1 + b_2 \quad (7)$$

Decoder:

$$Output\ = \tanh(w_3\ layer_2 + b_3) \quad (8)$$

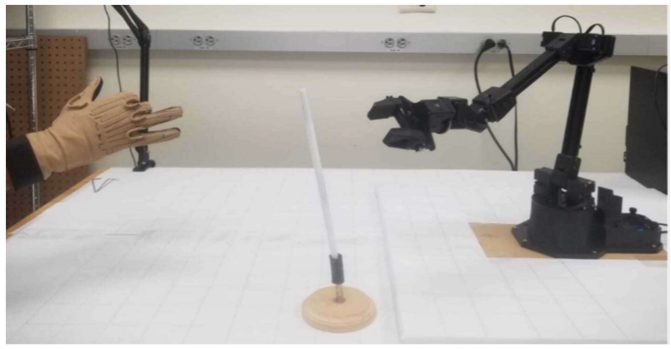

Fig. 2. Experiment Setup. The 5 DOF robotic arm and gripper (right) the object (middle) and the participant's hand in the CyberGlove (left).

where, w_j and b_j correspond to the network weights and biases for the jth network layer. The surjective mapping is shown below in (9).

$$M_{NLAE}\,(S) = Encoder\,(s) \quad (9)$$

C. Normalization

The output of $M_{group}(s)$ for all groups was normalized to the interval (0,1) using the following normalization function:

$$Z\ (s) = (M_{group}\,(s) - L_{\min}) \oslash (L_{\max} - L_{\min}) \quad (10)$$

where s is the vector of sensor values, M_{group} is the surjective mapping algorithm, $L_{\max}$ is the 5-D vector maximum of $M_{group}(X)$ (X being the calibration data), $L_{\min}$ is the 5-D vector minimum of $M_{group}(X)$ and $\oslash$ denotes Hadamard element by element vector division. Z(s) is denoted as $*$ in Fig. 1; where M^*_{group} represents $Z(M_{group}(s))$.

D. Protocol

Thirty-six healthy adults (aged 18–55) provided written informed consent to procedures approved by the Institutional Review Board at Florida International University. Participants were randomly assigned to one of 3 groups (n = 12/group), which differed in terms of the surjective mapping that they used for control. Prior to beginning the robot control task (below) participants completed the Purdue Pegboard test which is used to measure manual dexterity [23]. Groups were gender balanced and participant ages were on average 25 ± 6 years.

Participants were seated 3 feet away from a tabletop 5-DOF robotic arm and gripper which they controlled with their finger movements for approximately two hours. They were seated outside of the robot workspace and observed the motion of the robot from the left-hand side of the robot arm. Participants wore a CyberGlove 3 data glove on their dominant hand. We recorded 19 joint angles from the fingers, thumb, and palm arc and used these signals to control the configuration of the robot arm Fig. 2. The participant's task was to pick up cylindrical objects (0.5 cm diameter by 21 cm height) placed throughout the robot's workspace (45 unique object positions) and move them to a target bin (5-inch diameter) located on the near side of the table. The objects were supported by a spring which allowed the robot to bump into them without knocking them over. If

the object was pushed outside of the robot workspace or was dropped, the trial ended.

Subjects were placed into one of three groups, each controlled the robot by using a different surjective mapping between their finger joint angles and the robot configuration: PCA, EPCA, or NLAE.

Task: All groups performed the same two tasks: a 5-minute exploration task, and a 90-minute testing task. The exploration task consisted of unguided control of the robot during which participants explored how their hand postures related to the motion of the robot. The testing task consisted of 5 epochs of 9 trials each, during which participants would attempt to pick up the object and deposit it into the target bin. All trials started from the same nominal robot configuration and lasted a maximum of 2 minutes. Between epochs there was a rest period of approximately 2 minutes. Participants were told that they would be evaluated based on how many trial stages they completed per trial, where stages were defined as: object touch, object grasp, object lift, and object place (into the target bucket).

E. Data Analysis

To assess the total hand variance captured by each of the maps we computed the *Variance Accounted For (VAF)* depicted in (11).

$$VAF = \left(1 - \frac{var(X - \widehat{X})}{var\,(X)}\right) \tag{11}$$

where X is the 300 by 19 matrix of sensor values used to calibrate the mappings and $\hat{X}$ is the reconstruction of the calibration data from its 5-D latent representation. The reconstruction of the sensor data for the NLAE group was performed by passing the 5-D latent representation through layer 3 of the network. The reconstruction for the PCA and EPCA group was obtained by algebraically solving (3) and (5) for s. A VAF of 100% would indicate that the inverse mapping could perfectly reconstruct the calibration sensor data from its 5-D latent representation.

Our primary outcome measure was trial ***score***. During each trial participants were awarded one point for every trial stage they completed; where stages were defined as: touching the object with the robot end effector, grasping the object with the robot gripper, lifting the object, and placing the object into the target. The ***score*** measure was computed as the sum of points a participant earned throughout a trial. This measure was selected prior to the start of data collection.

$$Score = \sum_{i=1}^{4} points_i \tag{12}$$

The performance measure in (12) was formulated to capture our intuitions about what a successful trial consisted of (i.e., passing through each natural stage of the reach should count as increasingly successful), however, many variations on that expression may also meet those requirements. To ensure that the specific formulation for computing "score" we chose was not itself the main factor in determining our statistical outcomes, we tested numerous variants. Additional measures we examined to characterize the learning outcomes and robot control post hoc were:

- *Nonlinear Progressive Score Measures:* Variants of (12), in which more points are awarded for later trial stages.
- *Success Rate:* The total number of successful trials completed per epoch by participants.
- *Smoothness Index:* Number of peaks in the robot end effector speed profile during reaches to objects. A peak is a speed greater than 20% of the maximum robot end effector speed of each trajectory [16].
- *Normalized Path Length:* The path length of the robot end effector trajectory during reaches to objects normalized by the straight-line distance between the starting position of the end effector and the target object [16].
- *Contact time:* The time taken to successfully reach the object and touch it with the robot end effector, where failed trials are set to the maximum of 120 seconds.
- *Normalized Joint Variance:* Variance of the motion of each robot joint normalized by the total motion variance of all robot joints.
- *Planarity:* The percentage of hand posture variance captured by the first 5 Principal Components in each experiment trial. Where the PCs are obtained by performing PCA on the trial data (not the calibration data) [11].
- *Hand Variance:* The total variance of all 19 dimensions of the user's input during an experiment trial.

To investigate whether the type of surjective mapping had an effect on learning we performed a two-way repeated measures ANOVA on our primary outcome measure score, in which the within participants factor was epoch and the between participants factor was group. We used Shapiro-Wilk's test to verify the normality assumption and Mauchly's test to verify the sphericity assumption. Where sphericity was violated, we applied the Greenhouse-Geisser correction.

To characterize changes in participant's control of the robot across time we compared the Success Rate, Contact Time, Normalized Path Length, and Smoothness Index during the first and last experiment epochs using rANOVAs. To identify differences in control characteristics between groups we performed one-way ANOVAs on the Success Rate, Contact Time, Normalized Path Length, Smoothness Index, and Normalized Joint Variance during the last experiment epoch.

The significance level was set at $\alpha = 0.05$ and post hoc comparisons were performed using a Bonferroni corrected significance ($\alpha = 0.025$ for the rANOVAs). Where normality was violated we examined the Kurtosis and Skewness of the data and found the distributions were approximately normal [24], [25].

Additionally, to compare the effect of group with the effect of initial participant aptitude we performed a Hierarchical Regression in which we regressed baseline participant performance (cumulative sum of trial scores during the first epoch) and group against acquired proficiency (cumulative sum of trial scores during epochs 2-5); and subsequently regressed only baseline performance against acquired proficiency without the group predictor. Correlations were assessed using the Correlation Coefficient in MATLAB. All group comparisons were performed in the SPSS statistics software.

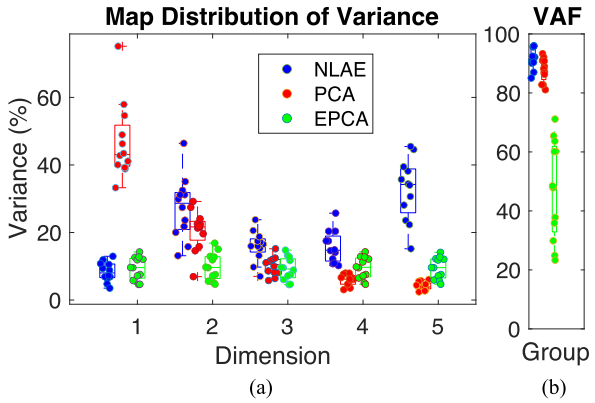

Fig. 3. Comparison of the calibration variance captured and distribution of calibration variance across latent dimensions by each map. (a) Variance Accounted For by each latent dimension of the maps for every experiment participant. The PCA map (middle) had the greatest variance accounted for by its first latent dimension with steeply declining variance accounted for by the subsequent dimensions. The EPCA map (right) had the same variance accounted for by each of its latent dimensions. The NLAE map (left) had a random amount of variance captured by each latent dimension. (b) The total variance accounted for by each of the maps. The NLAE (left) has the greatest VAF, followed closely by PCA (middle), and far exceeded the VAF by EPCA (right).

III. Results

We constructed 3 surjective maps between the hand and robot, each maximizing a different control principle to reveal what features of the map affected operator skill acquisition. We selected 1) Principal Components Analysis (PCA) to maximize the linear capture of finger variance, 2) our novel Egalitarian Principal Components Analysis (EPCA) to maximize the equality of variance captured by each latent dimension of the map, and 3) a Nonlinear Autoencoder (NLAE) to both maximize the ability to represent the user's full range of finger movements and achieve a more equitable distribution of variance across latent dimensions than PCA.

A. Representation of Calibration Data

The NLAE captured the greatest amount of variance in the calibration data. On average, it captured 92% of the user's calibration hand variance with 5 latent dimensions; compared to 88% for PCA and 47% for EPCA (Fig. 3(b)). By construction, the mappings also differed in how they distributed the user's calibration hand variance across their latent dimensions. The NLAE distributes user variance randomly across its code units while PCA distributes variance in a rank-ordered fashion across its basis vectors and EPCA distributes variance uniformly across its basis vectors (Fig. 3(a)).

B. Task Performance

Our primary goal was to identify how the map features affected user skill acquisition at a functional robot control task. By construction, each group's map differed significantly with respect to the calibration variance captured, and their distribution of the calibration variance across latent dimensions (Fig. 3). However, despite the stark differences between the maps, we identified no significant performance difference between groups. A two-way repeated measures ANOVA performed on scores (averaged over each epoch) yielded no significant interaction effect of group and epoch on score ($F(4, 8) = 22.121$, $p = 0.72$, Partial Eta Squared $= 0.04$). Our *a priori* power was estimated as 0.93 for a medium effect size of $F = 0.25$ for our two-way repeated measures ANOVA, meaning, if there was a substantial effect of the map on participant skill it is very likely that we would have detected it. We found no significant main effect of the group via a one-way ANOVA performed on the average points earned by each participant ($F(2, 33) = 2.81$, $p = 0.073$, Eta Squared $= 0.15$). There was a significant effect of epoch on participant scores ($F(4, 132) = 22.12$) $p < 0.01$; partial eta squared $= 0.40$), indicating that participants were able to improve at the task with practice regardless of the map. Repeating this analysis after substituting 'score' with other nonlinear progressive score measures did not change our findings.

To test whether participant's manual dexterity impacted their performance at the task we computed the Correlation Coefficient between participant's Purdue Pegboard Test Scores and their total points earned during the robot control. We found no significant correlation between dexterity and performance ($R = -0.08$, $p = 0.66$), indicating that manual dexterity did not influence participant's ability to learn the task (One NLAE participant was omitted from the correlation analysis due to an error in their Purdue Pegboard Test data collection).

Successfully completing all task stages proved to be difficult for the majority of participants. However, there were a few participants in each group that excelled at the task. Throughout the experiment participants in all groups were able to improve significantly in terms of their ability to successfully complete all task stages. Comparing the first and last epochs there was a main effect of practice on Success Rate for all groups ($F(1, 11) = 10.39$, $p = 0.008$); $F(1, 11) = 9.37$, $p = 0.011$; $F(1, 11) = 13.28$, $p = 0.004$ NLAE, PCA, and EPCA respectively). Additionally, we found no significant difference in Success Rate between groups during the last epoch ($F(2, 33) = 1.76$, $p = 0.19$, Partial Eta Squared $= 0.10$).

Throughout the experiment participants in all groups significantly reduced the time required to reach to and touch each of the objects ($F(1, 11) = 18.851$, $p = 0.001$; $F(1, 11) = 21.359$, $p < 0.001$; $F(1, 11) = 11.764$, $p = 0.006$); for NLAE, PCA, and EPCA respectively) (Fig. 9). Comparing group average contact times in the 5th (last) epoch we found a statistically significant difference between the groups ($F(2, 33) = 4.89$, $p = 0.01$, Partial Eta Squared $= 0.23$). Post hoc analysis using a Bonferroni corrected alpha revealed the PCA group was significantly different from the EPCA group ($p = 0.01$) with regards to contact time.

The participant task performance results (Figs. 4 and 5) indicated that participant's initial aptitude may be more predictive of acquired proficiency than the mapping they used. To compare the predictive power of initial aptitude to the predictive power of the group on participant skill acquisition we performed a Hierarchical Regression. We first regressed initial aptitude (participant's points earned in the first testing epoch) and group against

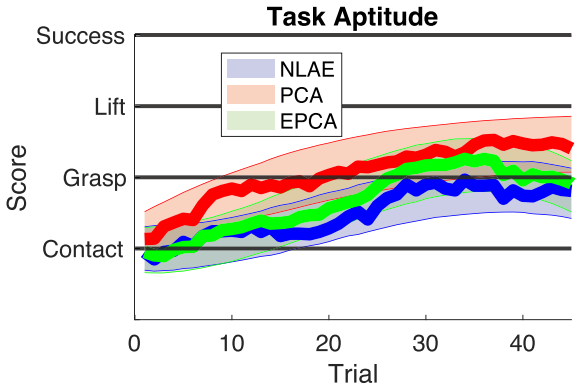

Fig. 4. Comparison of each group's performance at the robot control task. Within each group participant scores for each trial were averaged. The group averages were then smoothed (using a moving average filter over the 10 preceding trials) and plotted. Error bounds depict standard error. The large inter-participant variability dominates any small difference in performance between groups.

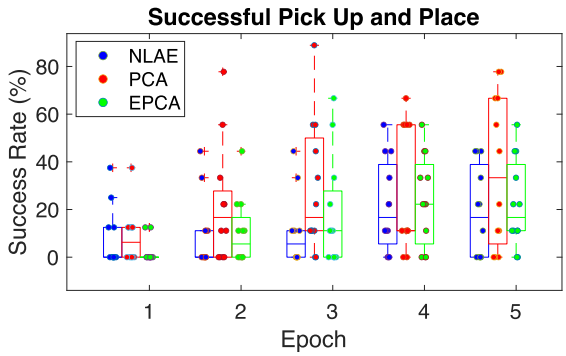

Fig. 5. There is no difference in participant's ability to complete all task stages between groups. Box plots depict individual participant's percentage of successful trials per experiment epoch.

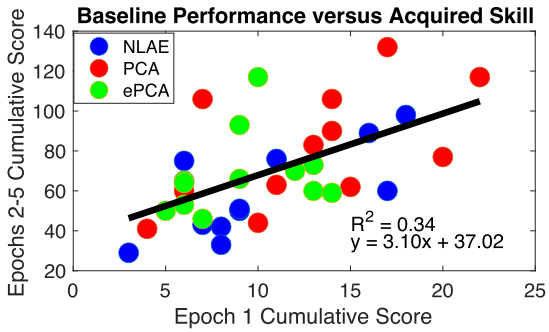

Fig. 6. Initial aptitude is highly predictive of acquired proficiency. Scatter plot depicts the participants' cumulative sum of trial scores in the first experiment epoch versus the cumulative sum of trial scores throughout the remainder of the experiment.

acquired proficiency (participant's total points earned in epochs 2–5) and found a strong correlation $R^2 = 0.40$ and a significant $R^2change (R^2change = 0.40, F(3, 32) = 6.99, p < 0.01)$.

Then we removed the group from our predictors and regressed only initial aptitude against acquired proficiency (Fig. 6). Without the group predictor we again identified a strong correlation $R^2 = 0.34$ and found a nonsignificant change in R^2 ($R^2change = -0.06$, $F(2, 32) = 1.53$, $p = 0.23$). These results (depicted below in Table I) show that participant's initial aptitude is highly predictive of acquired proficiency while the map participants were provided has no predictive power with regards to participants' acquired skill.

TABLE I
SPSS OUTPUT FROM THE HIERARCHICAL REGRESSION

Model Summary		Model 1	Model 2
R		.63[a]	.58[b]
R Square		.40	.34
Adjusted R Square		.34	.32
Std. Error of the Estimate		20.36	20.67
Change Statistics	R Square Change	.40	-.06
	F Change	6.99	1.53
	df1	3	2
	df2	32	32
	Sig. F Change	<.001	.23

a. Predictors: (Constant), EPCA, PCA, Baseline Performance
b. Predictors: (Constant), Baseline Performance

The nonsignificant F change for model 2 indicates that group has no predictive power with regard to participant's acquired skill.

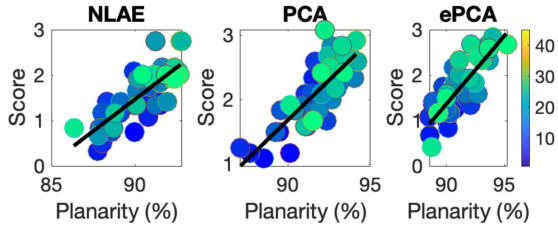

Fig. 7. Average group performance on each experiment trial plotted against the average planarity of participant hand motions on each trial. On average participants earned more points when they increased the planarity (variance captured by a 5-D linear latent space) of their inputs. Color depicts the trial number.

C. Analysis of Robot Control

To investigate the effect of the map's distribution of participant calibration hand variance on the participant's reliance on robot DOFs we computed the normalized robot joint variance during the last experiment epoch (Fig. 10). One-way ANOVAs performed on participants normalized robot joint variance identified no statistically significant difference between the groups ($F(2, 33) = 1.42$, $p = 0.26$, Eta Squared $= 0.08$; $F(2, 33) = 2.27$, $p = 0.12$, Eta Squared $= 0.12$; $F(2, 33) = 0.38$, $p = 0.69$, Eta Squared $= 0.02$; $F(2, 33) = 0.78$, $p = 0.47$, Eta Squared $= 0.05$; $F(2, 33) = 0.41$, $p = 0.66$, Eta Squared $= 0.02$; for joints 1–5 respectively). Comparing the distribution of robot joint variance shown in Fig. 10 with the distribution of user's calibration hand variance shown in Fig. 3, we did not identify the same pattern of variance distribution. The allocation of participant's hand variance across latent dimensions during calibration was not predictive of participant's reliance on robot DOFs during the control task.

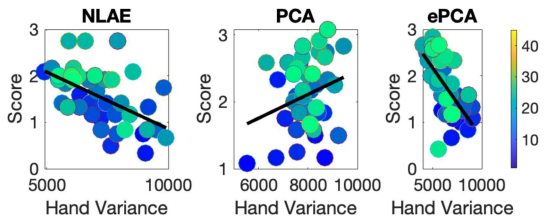

Fig. 8. Average group performance on each experiment trial plotted against average hand motion variance. Participants in the different groups learned different strategies to adapt to the structure of their maps but despite behavioral differences in user inputs, there are no differences in task performance.

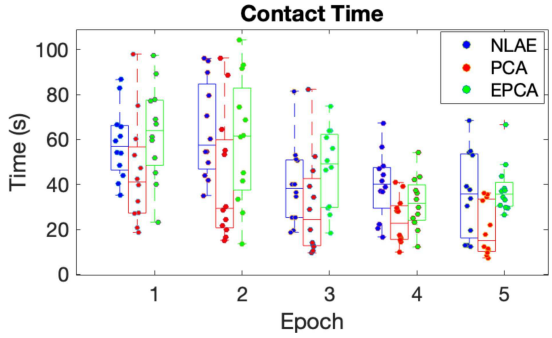

Fig. 9. Distributions of the time participants took to reach to and touch each object during the testing trials (averaged over epoch). Participants in all groups learned to reach objects faster as the experiment progressed. Comparison of group performance in epoch 5 yielded a significant difference between PCA and EPCA groups (removal of outliers did not change results).

Throughout the experiment participants in the PCA group increased the smoothness of their end effector trajectories ($F(1, 10) = 11.60$), $p < 0.01$, Partial Eta Squared = 0.54) (Fig. 11). Participants in the NLAE and EPCA groups did not appear to increase smoothness ($F(1, 11) = 0.28$, Partial Eta Squared = 0.025; $F(1, 11) = 0.79$, $p = 0.39$, Partial Eta Squared = 0.07 respectively). A comparison of the Smoothness Index of all groups in the final testing epoch revealed a significant difference between groups ($F(2, 32) = 5.23$, $p = 0.01$, Partial Eta Squared = 0.25). Post hoc analysis using a Bonferroni corrected alpha revealed a statistically significant difference between EPCA and PCA ($p = 0.01$). One PCA subject was excluded from the Smoothness Index analysis since their end effector velocity data was not recorded.

Participant's in the PCA group reduced the normalized path length of their end effector trajectories ($F(1, 10) = 13.90$, $p < 0.01$, Partial Eta Squared = 0.58) (Fig. 12). The NLAE and EPCA participants did not appear to reduce their normalized path length ($F(1, 11) = 0.06$, $p = 0.81$, Partial Eta Squared < 0.01; $F(1, 11) = 1.81$, $p = 0.21$, Partial Eta Squared = 0.14, respectively). A comparison of normalized path length in the last epoch found no statistically significant difference between any of the groups ($F(2, 32) = 1.26$, $p = 0.30$, Partial Eta Squared = 0.07). One PCA subject was excluded from the normalized path length analysis since their end effector trajectory data was not recorded.

D. Learning Correlates

Despite the difficulty of the task, participants were able to improve with practice, which raised the question of what participants had learned to do to improve. To answer this question, we searched for characteristics of the participants' inputs that correlated with improved task performance. We found a significant correlation between the average (5-dimensional) planarity of participants' hand postures and the average scores earned on a trial across all groups ($R = 0.78$, $p < 0.01$; $R = 0.80$, $p < 0.01$ $R = 0.82$ $p < 0.01$ for NLAE, PCA, and EPCA respectively) (Fig. 7). This result indicated that improving at the task required participants to learn to reduce the dimensionality of their inputs.

Next, to further characterize how participants in each group changed their inputs to improve at the task we examined the average variance of the participants' hand motions in each group over each experiment trial (Fig. 8). Participants in the PCA group did not significantly reduce the variance of their hand motions ($R = 0.26$, $p = 0.08$), while participants in the EPCA and NLAE groups did ($R = -0.53$, $p < 0.01$; $R = -0.66$, $p < 0.01$; for NLAE and EPCA respectively). Participants in the EPCA group may reduce the variance of their hand inputs as they realize that much of their hand motion has no impact on the robot control (due to the low VAF of the EPCA map), while participants in the NLAE group may reduce their hand variance because the nonlinearities of the map can relate small changes in the input to very large changes in robot commands. In both instances, the reduction in hand variance may be explained by a desire to increase feedback from hand motions.

The PCA group did not follow the same trend of hand variance reduction (Fig. 8) observed in the NLAE and EPCA groups, but did follow the trend of increasing planarity (Fig. 7). This result suggests that users in the PCA group learned to confine their finger articulation to a subspace of their full possible range of motion while simultaneously increasing the amount of finger motion within that subspace. PCA may differ from the other two groups in this respect because it both captures a large amount of calibration variance and is linear. Since the PCA bases capture a large amount of variance they represent hand motions in which many (if not all) of the user's hand joint angles are changing across a large range of motion. Therefore, to move throughout the PCA subspace (and change their commands to the robot joints) users have to make hand movements which have a large amount of variance associated with them. It is possible that, unlike in the NLAE and EPCA groups, users in the PCA group cannot reduce total hand motion variance to increase feedback.

IV. Discussions

In this human-machine interface (HMI) study, we implemented 3 abstract surjective maps that transformed 19 joint angles of the human hand into a command for the configuration of a 5 DOF, physically embodied robotic arm. We constructed our maps to investigate to what degree people using a nonlinear autoencoder could outperform those using a simpler linear PCA map for high-DOF robot control, and to identify if any performance differences were attributable to PCA's biased variance

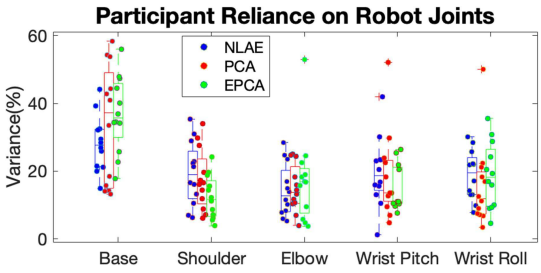

Fig. 10. Map distribution of calibration variance is not predictive of participant's reliance on robot joints. Box plots depict the percentage of robot motion variance allocated onto each robot joint during the last experiment epoch for each participant. One-way ANOVAs performed on the group data for each robot joint yielded no significant difference between groups (this result was unchanged by the removal of outliers).

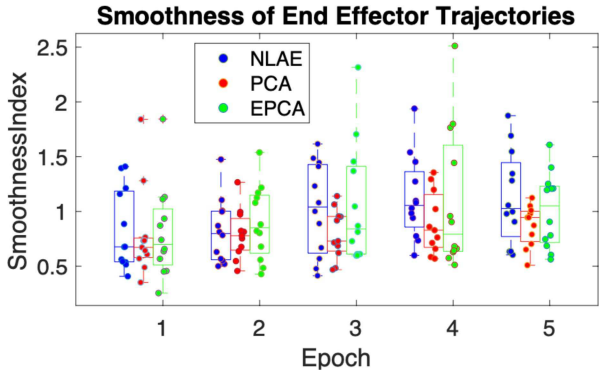

Fig. 11. Box plots depict the normalized smoothness index of the robot end effector velocity profile when reaching to objects (averaged over each epoch). Participants in the PCA group reach along smoother trajectories at the end of the experiment compared to the beginning. We found a significant difference in the smoothness of end effector trajectories between the PCA and EPCA groups in epoch 5 (removal of outliers did not change this result). However, this difference disappeared after normalizing the smoothness index by the reach duration, suggesting that the smoothness of reach trajectories was not different between the groups.

allocation. The study's main finding was that the map linking hand joints to robot joints did not affect performance (Fig. 4). This result runs contrary to the implicit (and often explicit) hypothesis underlying a large body of HMI work, namely that optimizing the user-to-device map is critical for users to become proficient with high dimensional HMIs [9], [16], [26], [27], [28], [29]. Instead, we find that subjects display a range of natural aptitude, from skilled to non-functional, and that this range is equivalent across the three very different maps (Figs. 5 and 6). Our findings suggest that efforts to improve high dimensional, robot HMIs should not focus on tuning and calibrating the static user-to-device map, and instead focus on other areas such as co-adaptation [12], [30] feedback [31], or learning environment [32].

A. Effect of Map Structure on User Performance

Since PCA distributes variance in a rank-ordered fashion along its basis vectors (Fig. 3(a)), the lowest variance principal components are often attributed to noise. In the case of the hand data collected in this study, on average PCs 4 and 5 accounted for only ∼6% and ∼4% of the user's calibration variance respectively. If these PCs represent calibration finger articulations that were less natural or accessible to participants, we might expect participants to be less able to control the corresponding robot DOFs, ultimately leading to a bias toward using the DOFs corresponding to the top PCs and an overall degradation in robot controllability. Other investigators have made similar speculations when constructing mappings from body kinematics to high DOF devices (e.g., simulated robots [16]).

To test this "PC control bias" hypothesis, we created the novel Egalitarian Principal Components Analysis (EPCA), which forces a similar allocation of variance among each of its basis vectors (Fig. 3(a)). This eliminated the supposed problem of low-variance-capture bases in PCA, at the expense of a reduction in overall variance explained. If biased variance capture is a large detriment to controllability we would expect the EPCA group to perform better than PCA, on the other hand, if total calibration variance explained is more important for controllability we would expect the reverse.

We found that there was no difference in performance between the PCA and EPCA groups, suggesting that steeply declining rank-ordered variance allocation among map bases does not impede controllability. In other words, the PC control bias hypothesis is false. Since, the task includes both positioning and orienting the end effector of the robot, PCA users needed to control all DOFs of the arm including the wrist DOFs (controlled by the low variance PCs). Not only could users in the PCA group learn to control the wrist DOFs, we further found that after practice there was no difference in the allocation of variance across the robot DOFs between any of the groups (Fig. 10.) This result indicates that, although a bias in the control may initially be present due to the imbalance in the distribution of variance across PCs, through feedback users can readily adapt to this aspect of the mapping.

The EPCA-PCA performance equivalence raises the possibility that both high total variance explained and unbiased variance among bases are required for boosting task proficiency. This tradeoff can be overcome by using the NLAE, which has higher total calibration variance explained and lower variability of variance explained among its latent dimensions than PCA. But even the NLAE group did not outperform PCA (and was possibly slightly worse). These results suggest that, for the class of maps calibrated on a user's unstructured exploration of their articulation space, the structure of the map is unimportant for robot controllability.

B. Robot End Effector Trajectories

Despite equivalent performance between groups, it is still possible that differences in the maps led participants towards different robot control strategies. To assess this possibility, we characterized the robot end effector trajectories in terms of the smoothness index and normalized path length (Figs 11 and 12). We found a significant difference in the Smoothness Index between the PCA and EPCA groups at the end of practice. Since,

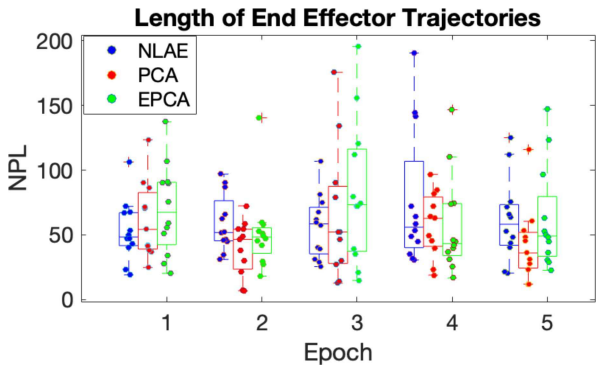

Fig. 12. Box plots depict the normalized path length of the robot end effector when reaching to each object (averaged over epoch). Participants in the PCA group reached along shorter trajectories at the end of the experiment compared to the beginning. We found no difference between groups in terms of normalized path length in epoch 5 (removal of outliers did not change this result).

we did not also identify a difference between the NLAE group and either the PCA or EPCA groups we concluded that the observed difference in Smoothness Index likely resulted from the slightly longer time that the EPCA participants took to reach the objects in the final epoch. Normalizing Smoothness Index by the contact time for each trial resulted in no significant difference between any of the groups.

We also did not identify any difference in normalized path length between the groups in the last experiment epoch. Although the normalized path length of our participants in the final experiment epoch is quite large (on average, approximately 50 times the straight line distance between the end effector start position and target), compared to results previously reported [16]. We suspect that our participants may not exhibit the same robot control characteristics as the literature because 1) we distribute our objects throughout more workspace territory than is typical in the literature, and 2) we compute end effector paths based on the joint-space commands participants send and not from measured end effector positions. Since participants can generate commands faster than the robot can actuate them, the actual trajectories of the robot end effector are shorter than the computed trajectories. In the literature, the effect of time delay may be less significant due to the use of simulated robots [16], or velocity controls [8], [33], [34], [35], [36].

C. Determinants of Skill Acquisition

Although there was no difference in performance between our groups there was enormous inter-participant variability which suggested that individual participant aptitude may be the real determinant of user skill acquisition. Our Hierarchical Regression (Table I) demonstrated that an individual's performance early in the experiment was highly predictive of the scores they would go on to accumulate throughout the experiment. Participants that learned to advance to higher task stages earlier consistently earned higher scores.

Since, identifying how participants learned to perform the higher task stages may offer valuable insight into how to guide learning, we performed an exploratory data analysis to determine whether there were common behaviors which participants employed when successfully completing the task. We found that regardless of the map all participants learned to reduce the dimensionality (increase planarity) of their inputs when controlling the robot (Fig. 7). This result agrees with previously reported findings [11].

Additionally, despite there being no difference in robot control strategy, utilizing our different maps did require users to learn to generate different hand inputs. Participants in the NLAE and EPCA groups learned to reduce the variance of their hand inputs while participants in the PCA group did not. We suspect that the reduction in hand variance was likely learned to increase feedback from robot commands. Although participants learned to generate different hand inputs to adapt to the different maps this did not have a significant effect on the robot control.

D. Implications for Human-Machine Interfaces

A long-standing goal of the field of human-machine interfaces has been to design biomimetic interfaces which provide humans control of devices in the same manner that they naturally control their own body parts [37]. Although the biomimetic approach can reduce the learning burden of the human operator, previous works have shown that after practice there is no difference in the proficiency of control [28], [38], [39]. Furthermore, it may not be possible to design biomimetic controls for devices which are dissimilar from biological body parts [37]. Constraining our interfaces to be biomimetic may place a fundamental limitation on the kinds of devices we can control [37]. Our results provide evidence that the focus of human-machine interface research should change from optimizing the map to optimizing other aspects of the interface, in particular those which have been shown to have a profound impact on learning and control.

For physically embodied robots one notable example is the time delay between the user's issued commands and the robot's physical actuation of the command [2]. In our experiment, participants issued commands to the robot at $\sim$10 Hz; however, the robot can take $\sim$1 second to move to a new configuration if it must traverse the maximum workspace distance. If a new command is sent before the last has been achieved, the motion of the robot is truncated, and it begins moving toward the new commanded configuration. This means that if users move their hands quickly and send very different robot commands, they may receive less feedback on how their hand postures relate to the robot configuration.

Another important finding of our work is that a person's aptitude for the task is highly predictive of their acquired proficiency. This result suggests that individual-specific interventions may provide a way to improve the skill acquisition of people who would otherwise struggle to operate the device. Adapting the map during operation is an individual-specific intervention that has been shown to facilitate learning in low-DOF HMIs [12], [30], [40]. Future works should investigate the effect of map adaptation on learning to control complex robots.

V. Conclusion

In the field of robotics, rapid advancements have created a need for new control paradigms which can provide people with continuous, precise, and fast control of high DOF systems. In

this work, we demonstrated that this can be achieved using a high-dimensional hand teleoperation scheme where users operate a physically embodied 5 DOF robot arm and gripper using the motion of their hand joints. We utilized our robot control scheme to investigate the effect of the structure of the hand-to-robot map on the skill acquisition of the user. Since the human-machine interface field has placed a large focus on the optimization of the user-to-device map, we hypothesized that optimizing our map structure, in terms of variance capture and distribution of variance, would lead to improvements in control. However, after comparing the performance of 3 groups of people controlling the robot via different maps (each constructed to maximize different variance characteristics) we found that this hypothesis was false. We identified that learning is largely robust to the map structure and our findings support a shift in focus of human-machine interface research. Future efforts to improve human-machine interfaces should focus on aspects of the interface other than the user-to-device map such as time delay, feedback, learning environment, or map adaptation.

Acknowledgment

The authors would like to thank Paulwin Arancherry, Pedro Ivan Alcolea, Tzu-Hsiang Lin, Hary Usaquen, and Maral Daneshyan for their support.

References

[1] V. Villani et al., "Survey on human–Robot collaboration in industrial settings: Safety, intuitive interfaces and applications," *Mechatronics*, vol. 55, pp. 248–266, 2018.

[2] J. Y. C. Chen, E. C. Haas, and M. J. Barnes, "Human performance issues and user interface design for teleoperated robots," *IEEE Trans. Syst., Man Cybern.. Part C, Appl. Rev.*, vol. 37, no. 6, pp. 1231–1245, Nov. 2007.

[3] B. E. Dicianno et al., "Joystick control for powered mobility: Current state of technology and future directions," *Phys. Med. Rehabil. Clin. North Amer.*, vol. 21, no. 1, pp. 79–86, 2010.

[4] M. Casadio et al., "Body machine interface: Remapping motor skills after spinal cord injury," in *Proc. IEEE Int. Conf. Rehabil. Robot.*, 2011, pp. 1–6.

[5] M. Macchini, F. Schiano, and D. Floreano, "Personalized telerobotics by fast machine learning of body-machine interfaces," *IEEE Robot. Automat. Lett.*, vol. 5, no. 1, pp. 179–186, Jan. 2020.

[6] M. Ison et al., "High-density electromyography and motor skill learning for robust long-term control of a 7-DoF robot arm," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 4, pp. 424–433, Apr. 2016.

[7] A. Branner et al., "Neuronal ensemble control of prosthetic devices by a human with tetraplegia," *Nature*, vol. 442, no. 7099, pp. 164–171, 2006.

[8] S. Chau et al., "A five degree-of-freedom body-machine interface for children with severe motor impairments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2017, pp. 3877–3882.

[9] C. Pierella et al., "Linear vs non-linear mapping in a body machine interface based on electromyographic signals," in *Proc. IEEE Int. Conf. Biomed. Robot. Biomechatronics*, 2018, pp. 162–166.

[10] F. A. Mussa-Ivaldi et al., "Sensory motor remapping of space in human–machine interfaces," *Prog. Brain Res.*, vol. 191, pp. 45–64, 2011.

[11] R. Ranganathan et al., "Learning redundant motor tasks with and without overlapping dimensions: Facilitation and interference effects," *J. Neurosci.*, vol. 34, no. 24, pp. 8289–8299, 2014.

[12] Z. Danziger, A. Fishbach, and F. A. Mussa-Ivaldi, "Learning algorithms for human-machine interfaces," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 5, pp. 1502–1511, May 2009.

[13] K. M. Mosier et al., "Remapping hand movements in a novel geometrical environment," *J. Neurophysiol.*, vol. 94, no. 6, pp. 4362–4372, 2005.

[14] J. M. Lee et al., "Diversity of learning to control complex rehabilitation robots using high-dimensional interfaces," 2022, *arXiv:2110.04663*.

[15] J. M. Lee et al., "An exploratory multi-session study of learning high-dimensional body-machine interfacing for assistive robot control," *BioRxiv*, 2023.

[16] F. Rizzoglio et al., "A non-linear body machine interface for controlling assistive robotic arms," *IEEE Trans. Biomed. Eng.*, vol. 70, no. 7, pp. 2149–2159, Jul. 2023.

[17] F. Rizzoglio et al., "Building an adaptive interface via unsupervised tracking of latent manifolds," *Neural Netw.*, vol. 137, pp. 174–187, 2021.

[18] A. A. Portnova-Fahreeva et al., "Learning to operate a high-dimensional hand via a low-dimensional controller," *Front. Bioeng. Biotechnol.*, vol. 11, 2023, Art. no. 1139405.

[19] A. A. Portnova-Fahreeva et al., "Autoencoder-based myoelectric controller for prosthetic hands," *Front. Bioeng. Biotechnol.*, vol. 11, 2023, Art. no. 1134135.

[20] S. E. Khan and Z. C. Danziger, "Learning high dimensional hand control of a robot arm is largely independent of mapping structure," *Soc. Neural Control Movement*, vol. 77, 2022, Art. no. 102576.

[21] M. Quigley et al., "ROS: An open-source Robot Operating System," *ICRA Workshop Open Source Softw.*, vol. 3, no. 2, 2009, Art. no. 5.

[22] A. A. Portnova-Fahreeva et al., "Linear and non-linear dimensionality-reduction techniques on full hand kinematics," *Front. Bioeng. Biotechnol.*, vol. 8, pp. 429–429, 2020.

[23] E. Proud and M. E. Morris, "Skilled hand dexterity in Parkinson's disease: Effects of adding a concurrent task," *Arch. Phys. Med. Rehabil.*, vol. 91, no. 5, pp. 794–799, 2010.

[24] B. Lantz, "The impact of sample non-normality on ANOVA and alternative methods," *Brit. J. Math. Stat. Psychol.*, vol. 66, no. 2, pp. 224–244, 2013.

[25] M. J. Blanca et al., "Non-normal data: Is ANOVA still a valid option?," *Psicothema*, vol. 29, no. 4, pp. 552–557, 2017.

[26] J. Miehlbradt et al., "Data-driven body–machine interface for the accurate control of drones," *Proc. Nat. Acad. Sci.*, vol. 115, no. 31, pp. 7913–7918, 2018.

[27] E. Peshkova et al., "Exploring intuitiveness of metaphor-based gestures for UAV navigation," in *Proc. IEEE Int. Symp. Robot Hum. Interactive Commun.*, 2017, pp. 175–182.

[28] H. R. Schone et al., "Should bionic limb control mimic the human body? Impact of control strategy on bionic hand skill learning," *bioRxiv*, 2023, Art. no. 423.

[29] M. Li et al., *Learning User-Preferred Mappings For Intuitive Robot Control*. Ithaca, NY, USA: Cornell Univ. Library, 2020.

[30] A. L. Orsborn et al., "Closed-loop decoder adaptation shapes neural plasticity for skillful neuroprosthetic control," *Neuron*, vol. 82, no. 6, pp. 1380–1393, 2014.

[31] F. Abdollahi et al., "Error augmentation enhancing arm recovery in individuals with chronic stroke: A randomized crossover design," *Neurorehabilitation Neural Repair*, vol. 28, no. 2, pp. 120–128, 2014.

[32] M. F. Levin et al., "Emergence of virtual reality as a tool for upper limb rehabilitation: Incorporation of motor control and motor learning principles," *Phys. Ther.*, vol. 95, no. 3, pp. 415–425, 2015.

[33] L. R. Hochberg et al., "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm," *Nature*, vol. 485, no. 7398, pp. 372–375, 2012.

[34] J. L. P. Collinger et al., "High-performance neuroprosthetic control by an individual with tetraplegia," *Lancet*, vol. 381, no. 9866, pp. 557–564, 2013.

[35] S. N. Flesher et al., "A brain-computer interface that evokes tactile sensations improves robotic arm control," *Science*, vol. 372, no. 6544, pp. 831–836, 2021.

[36] A. B. Schwartz et al., "Cortical control of a prosthetic arm for self-feeding," *Nature*, vol. 453, no. 7198, pp. 1098–1101, 2008.

[37] T. R. Makin et al., "Soft embodiment for engineering artificial limbs," *Trends Cogn. Sci.*, vol. 24, no. 12, pp. 965–968, 2020.

[38] S. M. Radhakrishnan et al., "Learning a novel myoelectric-controlled interface task," *J. Neurophysiol.*, vol. 100, no. 4, pp. 2397–2408, 2008.

[39] C. W. Antuvan et al., "Embedded human control of robots using myoelectric interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 4, pp. 820–827, Jul. 2014.

[40] S. Dangi et al., "Design and analysis of closed-loop decoder adaptation algorithms for brain-machine interfaces," *Neural Comput.*, vol. 25, no. 7, pp. 1693–1731, Jul. 2013.