Formalized Task Characterization for Human-Robot Autonomy Allocation

Michael Young[1], Christopher Miller[1], Youyi Bi[2], Wei Chen[2], and Brenna D. Argall[3]

Abstract— **Humans and robots team together to perform tasks in various domains. Some tasks are easier to perform than others, but little work focuses on discovering the underlying mechanisms that affect perceived difficulty and task performance. To fill this gap, we propose a formalized approach to task characterization for human-robot teams using Taguchi design of experiments and conjoint analysis. With this, we conduct a 20 person study where participants operate a 6 degree of freedom robotic arm to perform manipulations defined by 6 kinematic features. We find that rotational features of a task contribute significantly more to decreased performance and increased difficulty than translational features. The participants also perform the activities with autonomous assistance. The data shows a reduction in the effect of these features on performance and difficulty when assistance is active. Furthermore, we examine when to trigger assistance based on thresholds set from outlier detection. The analysis indicates that rotational features and features leading to kinematic singularities are useful for triggering assistance. Future work will use these results to inform a dynamic autonomy allocation framework when the autonomous assistance should step in.**

I. INTRODUCTION

Significant research in psychology has focused on understanding which aspects of a task influence a human's performance. Little work, however, has concentrated on shared-control tasks performed jointly by humans and robots, especially in the domain of assistive robotics. The goal of autonomous assistance is to increase overall performance and extend the user's ability. Many people with motor impairments cannot independently perform various activities of daily living (ADLs). Engineers design assistive robots with the goal of increasing the independence of this population. Moreover, engineers endow assistive robots with autonomy to make them easier, safer, and more efficient. However, evidence suggests motor-impaired populations do not want autonomous assistance unless absolutely necessary [1]. The questions are 1) which tasks require assistance, and 2) what are the features defining these tasks?

*We would like to gratefully acknowledge support by grant from U.S. Office of Naval Research under the Award Number N00014-16-1-2247

** We would like to gratefully acknowledge the financial support from National Science Foundation (NSF) CMMI-1436658

[1]Michael Young and Chris Miller are with Mechanical Engineering, Northwestern University, and the Shirley Ryan AbilityLab, Chicago, IL - USA mikesyoung@u.northwestern.edu chris.x.miller@u.northwestern.edu

[2] Youyi Bi and Wei Chen are with Mechanical Engineering, Northwestern University, Evanston, IL - USA youyi.bi@northwestern.edu weichen@northwestern.edu

[3]Brenna D. Argall is with Electrical and Computer Engineering, Mechanical Engineering, and Physical Medicine and Rehabilitation Northwestern University, and the Shirley Ryan AbilityLab, Chicago, IL - USA brenna.argall@northwestern.edu

In this work, we aim to shed light on the question of which aspects of shared-control tasks influence human-robot team performance. We perform a study that characterizes the human-robot team's performance and the user's perceived difficulty based on specific features of robotic arm manipulation tasks. More specifically, we are interested in the kinematic features of the task under both unassisted (full-teleoperation) and autonomy-assisted operation. To do this, we utilize techniques typically used in industrial engineering and automotive manufacturing for design of experiments and feature extraction. Our overarching research goal is to detect which kinematic features contribute most to perceived difficulty and decreased performance so that autonomy can adjust the level of assistance accordingly.

In Section II, we cover related work in task complexity and shared-control. Our technical contribution begins in Section III where we outline a formal methodology for feature extraction and design a human-robot team experiment. The task complexity findings are in Section IV with discussion including insights for autonomy allocation in Section V. We conclude the paper with ideas for future work in Section VI.

II. BACKGROUND

This section covers related work in task characterization, task complexity, and shared-control for human-robot teams.

Researchers in the industrial and manufacturing setting generally use the framework of task complexity to determine whether a human can perform the task well. Many consider task complexity to have three viewpoints, covered in the review paper [2]. The *interaction* viewpoint considers the interaction between the user and the task, which means the user's experience and proficiency of operation are taken into account. The *resource requirement* viewpoint evaluates the requirements of the user by the task such as cognitive load. The *structuralist* viewpoint takes into account the elements of a task, where elements can be defined in a variety of ways. Some examples include possible paths or number of task goals [3].

In robotics, and specifically in human-robot teams, when researchers aim to understand what drives performance and safety, this implicitly addresses task complexity. One indicator of success is user acceptance of the robotic system [4]. Many researchers investigate human factors—such as age, skill, and gender—in the acceptance and performance of their system [5]. Others develop mechanisms to share control between the human and robot and investigate which features of the mechanism increases the likelihood of users accepting autonomous assistance [6]. In many ways, these works implicitly address the interaction and resource requirement viewpoints of task complexity, but their findings are usually not generalizable across tasks. The structural viewpoint is largely ignored in previous robotics work. One work, however, finds that the performance and safety of humans operating a robotic arm decreases based on structural features such as proximity to obstacles [7].

Though little work has been aimed at characterizing task complexity for robotics, many techniques are well suited for this type of work. For instance, conjoint analysis is commonly used to determine which features of a product indicate a consumer's likelihood to buy the product [8], [9]. This method predicts the value, or coefficient, of a feature at a specific level to maximize the overall utility of a particular design. For example, automobile manufacturers will evaluate how a car's headlight shape, the *feature*, influences the buyer's decision for different *levels* (i.e. round or square). A higher coefficient value for the level round would indicate a higher effect on consumer's purchasing patterns.

We are also interested in mechanisms that use information about the task to adjust how the human-robot team interacts and operates. One recent work develops a framework for the design of human-machine systems that accepts task variables such as task complexity for all three viewpoints, as well as human and robot attributes [10]. The framework is high level and highlights the need for accurate measurement of task complexity. In a similar manner, task features are taken into account to design shared-control paradigms for human-robot teams [11]. Other approaches use performance metrics of the task, rather than features, to adjust the parameters of the autonomous assistance in real-time [12]. Some systems identify specific tasks to trigger autonomous assistance [13]. These mechanisms require identification of the exact task, rather than a combination of features, to make an assessment as to whether the level of autonomy should be adjusted.

We aim to expand these fields of research by developing a formal process for characterizing task complexity for human-robot teams based on a set of features. In future work, knowledge of how features affect performance will inform a dynamic autonomy allocation framework when to shift levels of shared-control.

III. METHODS AND DESIGN

Little work has formalized task characterization in robotics. The methods presented in this section seek to provide both a formalized framework for this domain and a grounding for a specific platform and category of tasks, manipulation of objects using a robotic arm.

A. Feature Extraction

Conjoint analysis [9] is a technique well-suited for understanding which features play a role in a person's decision. We apply this technique to a slightly different purpose: to quantify features that increase the user's perceived difficulty and decrease overall performance. We start by building models for each individual participant using the features:

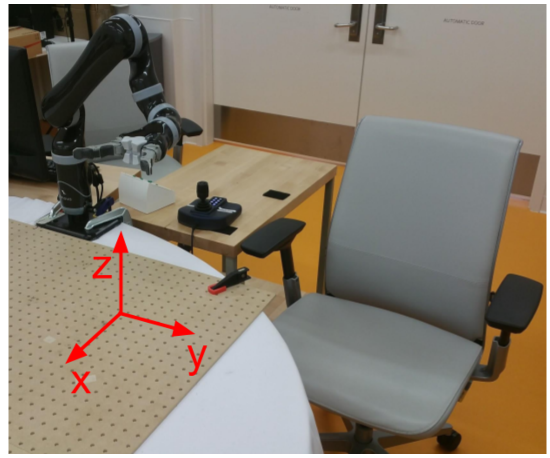

Fig. 1. Mico arm, 3-axis joystick, and world frame (x-, y-, z-axes).

$$U(q) = c + \sum_{i=1}^{N} a_i q_i \tag{1}$$

where $U(q)$ is the utility value output from the model, q_i are features treated as dummy variables, N is the number of features, c is the constant term, and a_i is the coefficient of feature i. A higher value for a given coefficient indicates that the feature or feature interactions impacts the utility significantly more.

In our implementation, $U(q)$ is the TLX score or task time for the particular participant, the dummy variables are 1 for present and 0 otherwise, and we estimated the values of a_i using linear regression for each participant.

B. Pilot Study

Conjoint analysis requires careful selection of features to acquire meaningful results. To identify these features, we perform an in-house pilot study where 6 participants use the Mico robotic arm (Kinova Robotics, Canada) and IPD Ultima 3-axis joystick (CH-Products Industrial, CA, USA) to perform manipulation tasks (Figure 1). To operate the robotic arm, users move the joystick forward/backward, left/right, and twist which affects translational motion in the x-, y-, and z-axes or rotational motion in roll, pitch, and yaw, respectively (Fig. 1). The user presses a button to switch between translational and rotational control.

The pilot study aims to investigate the impact of kinematic features, object thickness, accuracy of object placement and line of sight (LOS) on perceived task difficulty. Kinematic features define the movement of an object within a robot's workspace using Euler angle *rotations* and *translations*

	ϕ	θ	ψ	Distance	Thick	Acc.	LOS
a_i	2.25	5.57	2.06	-1.48	-9.23	-1.12	-4.06
σ^2	10.4	5.1	4.04	5.44	6.52	3.88	6.14

TABLE I
COEFFICIENT VALUES (MEAN AND STD.) FOR EACH FEATURE OF THE PILOT STUDY

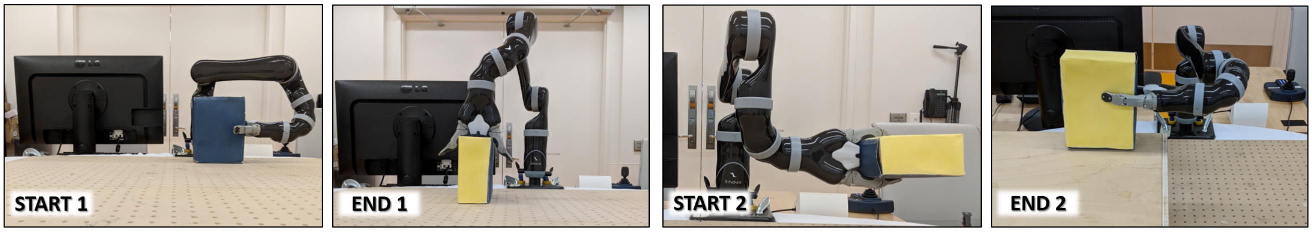

Fig. 2. Example tasks. Start and End 1 depict the starting and ending pose of feature set $[0, 0, 0, 1, 0, 1]$ which models pure rotations: the user picks up the object, rotates it, and places it in the same location. Start and End 2 depict the starting and ending pose of feature set $[1, 1, 1, 1, 1, 1]$ which models rotation about and translation through all three axes.

through the x-, y-, and z-axes (forward-back, left-right, and vertical axes from the user's point of view, respectively) in the world frame. Angles in the rotation space are parameterized using an 1-2-3 Euler angle rotation sequence through Euler angles ϕ, θ, ψ. An analysis of the pilot data using conjoint analysis (Eq. 1 without the interaction terms) indicate that rotational features are most significant among all features investigated (Table I). Furthermore, the large variance ($\sigma^2 = 5.44$) of the distance feature indicates that some users find distance (translation) adds to difficulty. The results of the pilot study thus indicate the selection of kinematic features for the conjoint analysis study.

C. Task Design

With the goal of investigating kinematic task features, we design a set of manipulation tasks that each test a combination of features. The object to be manipulated is a rectangular prism with varying face colors such that every orientation of the object is unique. The environment includes a shelf (for z-axis translations) but no other clutter so as to isolate the features of interest. To capture the effect of a feature, we choose to limit the features to binary levels (present or not present). We set the translational features as a constant fraction of the radius r of the robot arm's workspace to generalize the results to other robotic manipulators with different workspaces. The fraction values are empirically chosen as $\frac{1}{2}r$ for forward-back and left-right translation and $\frac{1}{3}r$ for vertical translation. We set the rotational features as 90 degrees about the rotational axis of interest.

D. Task and Feature Selection

For the 6 kinematic features with 2 (binary) levels per feature, a full factorial design includes 64 (2^6) tasks. However, using careful design of experiments (DOE), we reduce the number of tasks tested to help avoid human fatigue and ensure that the task feature space remains full rank. We use a Taguchi design [14], a version of fractional factorial DOE. In our implementation, the design does not include interactions and optimizes for a minimum number of tasks while ensuring significance of individual feature effects. Taguchi design lowers the number of tasks to 16 and we remove an additional 3 tasks due to identical kinematic representations and the null case, leaving 13 task feature sets.

We define a description f_i of the ith task with respect to feature levels as a binary vector of the 6 features $f_i = [x, y, z, \phi, \theta, \psi]$. The total feature space F is a matrix representing all tasks, where the ith row is f_i (Equation 2).

$$F = \begin{array}{c} \begin{matrix} x & y & z & \phi & \theta & \psi \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{array} \tag{2}$$

E. Shared Control Implementation

Our work furthermore aims to evaluate whether the challenges identified by conjoint analysis are mitigated by the addition of assistance from robotic autonomy. In the trials with assistance, the human's control signal u_{human} and the autonomy's control signal u_{auto} are combined using the blending paradigm in Equation 3 to produce the executed command $u_{control}$. The blending parameter ($\alpha = 0.5$) is selected empirically, as the goal of this study is not to investigate different levels of assistance but rather to determine when autonomy would be useful.

$$u_{control} = (1 - \alpha) \cdot u_{auto} + \alpha \cdot u_{human} \tag{3}$$

An in-house potential field library provides the autonomy control signals. An attractor is placed at the goal location to guide the participant in completing the task. The attractors assisted the participant in both translational and rotational motions. For tasks where the participant is required to place the object atop a shelf, the sides of the shelf behave as repellers to help the participant avoid object-shelf collisions and complete the task successfully.

F. Procedure

To complete the experimental tasks, participants used the same hardware as the pilot study. The full experimental protocol and consent form were approved by the Northwestern Institutional Review Board.

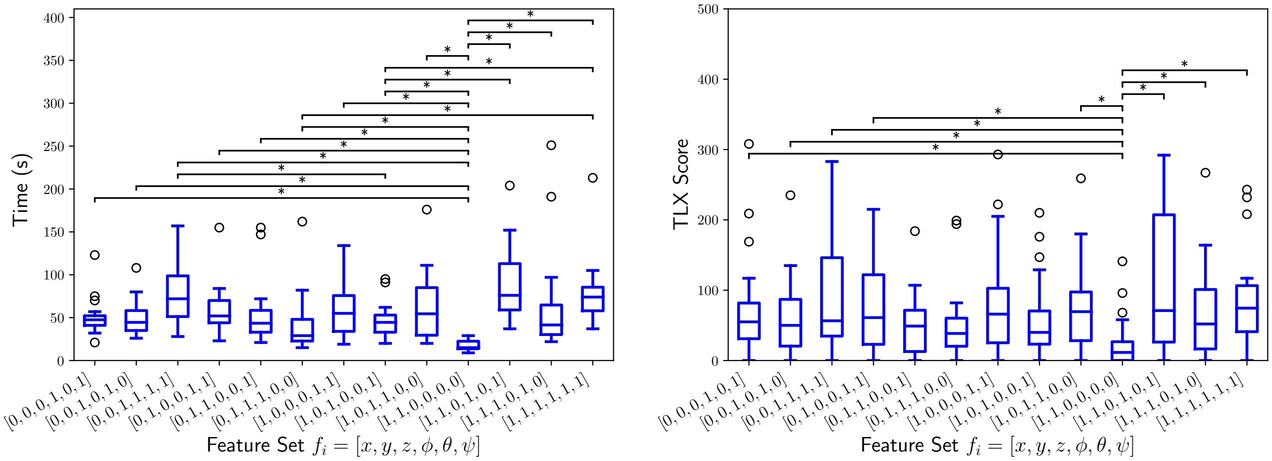

Fig. 3. Aggregated performance (left) and difficulty (right) metrics for the 13 tasks defined by feature sets under full-teleoperation (unassisted). Feature set $[1, 1, 0, 0, 0]$ is significantly lower than all other task times and 11 out of 12 TLX scores.

1) Participants: The group consisted of 10 female and 10 male participants (22 to 36 years old) with varying levels of familiarity with robotic devices.

2) Tasks: To study the effect of assistance on task completion for a given feature set, each task was completed both with and without assistance. Each participant completed 26 tasks (13 with assistance and 13 without). To reduce learning effects, the order of tasks was counterbalanced across participants, without repetition.

3) Protocol: At the start of each session, the participant was given approximately five minutes to become familiar with the joystick interface and controlling the arm.

Then at the start of a task, the participant was shown the object's start and end position and orientation. At any time, the participant could ask for a verbal description of the object's end position and orientation. Figure 2 depicts examples of the start and end positions for two tasks.

A task would begin when the participant closed the arm's gripper to grasp the object. The task would end when experimenters verified the object was placed in the correct ending position. After completing each task, the participant would complete a NASA TLX survey [15] and the task completion time—defined as the duration between when the gripper was closed and reopened—was recorded.

4) Metrics: We chose one subjective and one objective metric, which respectively measured the user's perceived difficulty and performance.

1) **Perceived Difficulty:** The raw TLX score. Numerous studies show the reliability of using this metric [16]. A higher value indicates increased difficulty.
2) **Performance:** The completion time of each task.

IV. RESULTS

The results indicate that rotational features and forward translation contribute more to perceived difficulty and task time of the tested tasks. Furthermore, we see the assistance from robotic autonomy reduces the effect of these features. We present the details of our findings in this section.

For all analyses, we use the non-parametric Kruskal-Wallis test to check for significance within groups and the Wilcoxin test with Bonferonni correction for pairwise comparisons (Kruskal-Wallis is chosen because an analysis of the data using a Shapiro-Wilk test finds that only 65% of the tasks likely stem from a normal distribution). Throughout, we denote statistical significance of $p < 0.05/m$ as *, $p < 0.01/m$ as **, and $p < 0.001/m$ as ***, where m is the Bonferonni correction.

A. Task Comparison

The first goal of this work is to characterize task complexity based on kinematic features for a robotic manipulator. To this end, we compute average trial times and TLX scores under full-teleoperation (no autonomy assistance). For full-teleoperation, the average task time is 57.7 ± 37.4s with a median of 48.0s; the average TLX score is 73.9 ± 61.89 with a median of 55.0. When comparing tasks to each other, we find significant variation (Fig. 3). Specifically the task with feature set $[1, 1, 0, 0, 0, 0]$ yields significantly lower task times and TLX scores than other tasks. This is likely due to the fact that it is the only feature set with only translational components and no rotational components. This analysis alone, however, does not fully explain the effects of individual features on the metrics.

To gain insight into individual effects of kinematic features, we utilize conjoint analysis. First, we build models (Eq. 1) for each participant and metric. The average (over subjects) coefficient values are reported in Table II. For time, ϕ is significantly ($p < 0.05$) higher than x and y, and for TLX score, ϕ is significantly higher than y and z. This sug-

TABLE II
COEFFICIENT VALUES FOR EACH FEATURE IN FULL TELEOPERATION

Metric	x	y	z	ϕ	θ	ψ
Time (s)	8.33	3.43	7.18	18.96	15.86	17.00
TLX Score	8.69	-2.20	-0.83	29.85	17.09	18.86

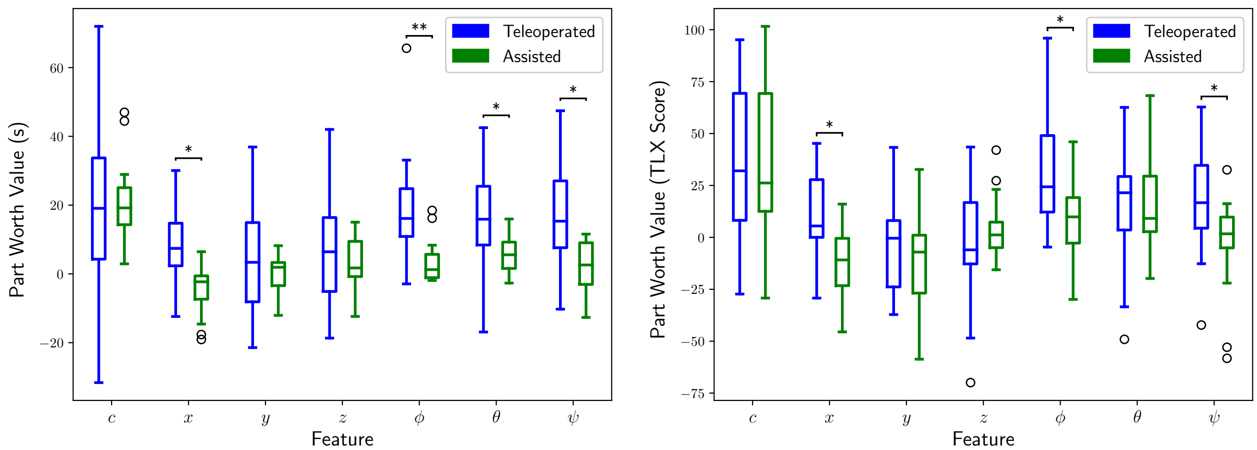

Fig. 4. Pairwise comparison of coefficient values under full-teleoperation and autonomy-assisted operation, performance (left) and difficulty (right). Full-teleoperation has significantly higher coefficients for all Euler angle features (ϕ and ψ only for TLX Score) and forward-back translation x.

gests that a task with a rotation about the forward-back axis is more difficult and takes longer to accomplish than those with only translational features. Translation in x also tends to be higher than in y and z. One possible explanation is moving an object forward and then changing the orientation can reach the limitations of the arm's workspace, causing kinematic singularities that often increases the time and difficulty of a task. Lastly, on average the coefficients for rotations are higher than translations. This indicates that the reason for increased times and scores is due to rotations. In other words, rotational task features contribute more to task complexity than translational features for robotic arm operation.

B. Assistance Comparison

The results in Section IV-A suggest that specific kinematic features influence perceived difficulty and task completion time. We therefore examine whether the addition of autonomous assistance in the control loop might improve those metrics. Participants perform the same tasks with assistance to achieve an average task time of 26.0 ± 14.6s (median 22.0s) and average TLX Score of 40.0 ± 44.3 (median 29.0). These values are significantly ($p < 0.001$) lower than those produced under full teleoperation.

Looking only at the averaged information, one might falsely assume that autonomy should always assist the user; however, the results in Table III show task dependence. For example, the task with feature set $[1, 1, 0, 0, 0, 0]$ (zero rotations) already has low TLX scores and task times under teleoperation, and feature set $[0, 0, 1, 0, 1, 0]$ does not see significant improvement with the addition of assistance.

We further analyze results from the conjoint analysis, this time comparing the coefficients between the full-teleoperation model and assistance model (Fig. 4). The conjoint analysis models have the following mean coefficients of determination (R^2) across subjects: Teleoperation TLX (0.52) and Time (0.55), Assistance TLX (0.47) and Time (0.46). The R^2 indicate there is room for improvement in exact prediction of TLX score and task time; however, it

TABLE III

TASK TIME AND TLX SCORE OF ALL FEATURE SETS FOR BOTH TELEOPERATION (TEL) AND ASSISTED (AST) OPERATION (OP) WITH PAIRWISE SIGNIFICANT DIFFERENCES (SIG)

Feature Set $[x, y, z, \phi, \theta, \psi]$	Op	Time (s) *Mean(SD)*	Sig	TLX *Mean(SD)*	Sig
[0,0,0,1,0,1]	Tel Ast	50.4(20.5) 22.9(8.91)	**	76.5(67.8) 37.2(40.4)	*
[0,0,1,0,1,0]	Tel Ast	49.8(20.4) 43.4(20.0)		61.3(41.4) 83.3(86.3)	
[0,0,1,1,1,1]	Tel Ast	78.5(32.9) 37.2 (21.4)	**	102(70.7) 76.6(63.4)	
[0,1,0,0,1,1]	Tel Ast	58.8(27.1) 23.4 (7.51)	***	75.2(45.8) 32.3(23.6)	**
[0,1,1,0,0,1]	Tel Ast	54.6(35.1) 27.3 (12.5)	**	55.6(33.3) 32.9(28.2)	*
[0,1,1,1,0,0]	Tel Ast	42.2(32.6) 20.5 (9.58)	**	62.0(51.6) 23.9(21.6)	*
[1,0,0,0,1,1]	Tel Ast	60.3(32.0) 18.1 (5.62)	**	81.9(70.9) 28.6(27.4)	*
[1,0,1,0,0,1]	Tel Ast	46.5(19.6) 20.7 (11.1)	**	62.3(46.5) 26.3(31.3)	*
[1,0,1,1,0,0]	Tel Ast	63.5(39.1) 19.3 (7.3)	***	81.2(50.9) 33.2(25.9)	*
[1,1,0,0,0,0]	Tel Ast	17.6(6.15) 17.5 (4.36)		21.8(25.3) 25.7(29.2)	
[1,1,0,1,0,1]	Tel Ast	88.1(41.8) 34.4 (17.1)	**	122(89.4) 47.2(34.1)	*
[1,1,1,0,1,0]	Tel Ast	66.9(56.8) 22.1 (8.41)	**	72.1(64.2) 28.8(26.8)	*
[1,1,1,1,1,1]	Tel Ast	77.6(36.5) 31.1 (11.5)	***	87.5(56.7) 44.8(32.9)	*

also suggests sufficiency in explaining the reason behind the metrics. The addition of interaction terms can increase the R^2 value but does not improve heteroskedascity, thus we report only the linear model (Eq. 1). For task time, all rotations (ϕ, θ, ψ) and x are significantly ($p < 0.05$) less important in the assistance case. For TLX score, x, ϕ, and ψ contribute significantly ($p < 0.05$) more in full teleoperation than the assistance case. We notice that θ contributes to task time but not perceived difficulty. Additionally, the two translational features y and z do not contribute significantly to either metric as a whole.

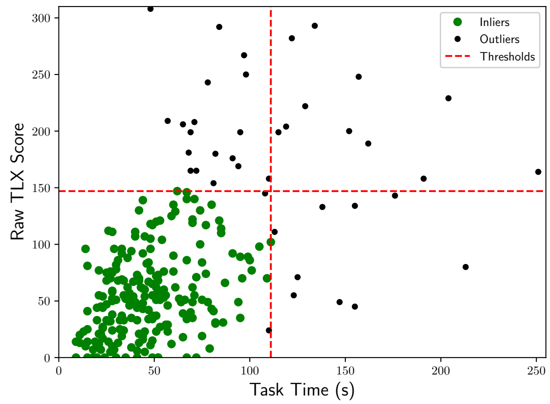

Fig. 5. DBSCAN results used to set task time and TLX score thresholds.

C. Validation Using Statistical Outliers

To cross-check the results of Sections IV-A and IV-B, we perform a parallel conjoint analysis using outlier detection to identify trials with poor performance or difficulty to compute an alternative utility value (Eq. 1), build models using this alternative utility function, and examine the feature sets of these models. For this work, we set thresholds to eliminate outliers in task time and perceived difficulty. To detect outliers, we use DBSCAN, an algorithm for detecting clusters with noise based on density [17]. DBSCAN requires two parameters for implementation: (1) the neighborhood radius ϵ and (2) minimum number of neighboring points. We set $\epsilon = 15$ based on the 4-distance graph and used the default value of neighboring points. The results provide maximum values for each metric based on the noise segmentation (Fig. 5): 111 seconds and 147 TLX score. If the task time or TLX score exceeds these limits, we appropriate this trial as requiring autonomous assistance.

For each feature set, we count the number of trials which exceed either or both thresholds (for task time or TLX score). We then use this count as the utility value $U(q)$ in Equation 1 (instead of TLX score or task time), and build a model. Of the 260 trials performed without assistance, 14.6% (38 trials) exceed either or both thresholds. The coefficients for the feature set $[x, y, z, \phi, \theta, \psi]$ are $[1.17, 0.17, -0.27, 3.37, 0.62, 1.48]$, respectively. Like in our prior analysis, we again observe high coefficient values for x, ϕ, and ψ. The value of θ (rotation about the y-axis) is less strong, which is the only difference with the prior model.

It is also worth noting that only 80% of participants contributed to the 38 cases that require autonomous assistance. This means that not all users will require assistance, regardless of feature combinations.

V. DISCUSSION

In this section, we briefly cover the insights from the formalized feature extraction for use in autonomy allocation. Furthermore, we acknowledge the limitations of the methods and grounding experiment as areas for future work.

A. Insights for Dynamic Autonomy Allocation

The characterization of a task's complexity based on kinematic features affords a dynamic autonomy allocation several advantages. Specifically, our results show that autonomy assistance can alleviate the effect of certain features which attribute to task difficulty and task time. If the autonomy can correctly identify these features, it can step in to provide assistance when these features are present. Limiting assistance to stepping in only when these task features are present also helps to keep the human maximally engaged in domains when this is what they desire. Our task characterization performed using conjoint analysis offers three main insights for adding autonomy assistance:

1) **Rotational assistance:** To add assistance when any rotational feature is detected.
2) **Singularity assistance:** To add assistance in situations where kinematic singularities may arise.
3) **User dependence:** To consider the individuality of the user before adding assistance. Our results do show that not all humans require assistance and whether feature set flags should be personalized remains an area for future investigation.

B. Limitations

We have demonstrated that a systematic design of experiments and analysis can extract features of interest for robotic tasks; still, limitations exist. This work did not explore the interactions of kinematic task features due to the number of experiments and trials required. With this smaller study, we now have the foundational data to justify further experimentation to include interactions, as well as a myriad of other task features (e.g., accuracy of placement, object size). Note that the inclusion of more features and interactions will require exponentially more data to generate enough statistical power for meaningful results. Additionally, this work was performed with only one platform, a 6-DOF robotic arm and a 3-axis joystick interface, and it remains to be seen to what extent these results might be platform or interface dependent. Our results do suggest that with careful feature selection it is possible to employ the DOE and analysis methods successfully.

VI. CONCLUSION

We have presented a formal task characterization methodology for human-robot shared-control and grounded these methods with a robot arm manipulation experiment. The results indicate that a formal design of experiments can extract the importance of certain features that influence the complexity of tasks in human-robot teams. Specifically in the domain of robotic manipulation, the analysis shows that rotational kinematic features are likely the cause of decreased performance and increased perceived difficulty compared to other features. The data further shows that the addition of robotic assistance not only improves performance but alleviates the contribution from rotations. These insights will be used in future work to build an autonomy allocation framework for an assistive robotic arm.

References

[1] D. Gopinath, S. Jain, and B. D. Argall, "Human-in-the-loop optimization of shared autonomy in assistive robotics," *IEEE Robotics and Automation Letters*, vol. 2, no. 1, pp. 247–254, 2017.

[2] P. Liu and Z. Li, "Task complexity: A review and conceptualization framework," *International Journal of Industrial Ergonomics*, vol. 42, no. 6, pp. 553–568, 2012.

[3] D.-H. Ham, J. Park, and W. Jung, "Model-based identification and use of task complexity factors of human integrated systems," *Reliability Engineering & System Safety*, vol. 100, pp. 33–47, 2012.

[4] J. Y. Chen and M. J. Barnes, "Human-agent teaming for multirobot control: A review of human factors issues," *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 1, pp. 13–29, 2014.

[5] I. H. Kuo, J. M. Rabindran, E. Broadbent, Y. I. Lee, N. Kerse, R. Stafford, and B. A. MacDonald, "Age and gender factors in user acceptance of healthcare robots," *Proc. IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN)*, pp. 214–219, 2009.

[6] A. Erdogan and B. D. Argall, "The effect of robotic wheelchair control paradigm and interface on user performance, effort and preference: an experimental assessment," *Robotics and Autonomous Systems*, vol. 94, pp. 282–297, 2017.

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

[8] P. E. Green, A. M. Krieger, and Y. Wind, "Thirty years of conjoint analysis: Reflections and prospects," *Interfaces*, vol. 31, no. 3_supplement, pp. S56–S73, 2001.

[9] R. D. Luce and J. W. Tukey, "Simultaneous conjoint measurement: A new type of fundamental measurement," *Journal of Mathematical Psychology*, vol. 1, no. 1, pp. 1–27, 1964.

[10] K. Stowers, J. Oglesby, S. Sonesh, K. Leyva, C. Iwig, and E. Salas, "A framework to guide the assessment of human–machine systems," *Human Factors*, vol. 59, no. 2, pp. 172–188, 2017.

[11] J. Beer, A. D. Fisk, and W. A. Rogers, "Toward a framework for levels of robot autonomy in human-robot interaction," *Journal of Human-Robot Interaction*, vol. 3, no. 2, p. 74, 2014.

[12] Q. Li, W. Chen, and J. Wang, "Dynamic shared control for human-wheelchair cooperation," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, 2011, pp. 4278–4283.

[13] M. Barnes and J. Grossman, "The intelligent assistant for electronic warfare systems," *(Rep. No. NWC TP 5885) China Lake: US Naval Weapons Center*, 1985.

[14] G. Taguchi, "Introduction to quality engineering: designing quality into products and processes," Tech. Rep., 1986.

[15] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," *Advances in Psychology*, vol. 52, pp. 139–183, 1988.

[16] S. G. Hart, "NASA-task load index (NASA-TLX); 20 years later," *Proc. of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 9, pp. 904–908, 2006.

[17] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise." in *Proc. of Knowledge Discovery and Data Mining (KDD)*, vol. 96, no. 34, 1996, pp. 226–231.