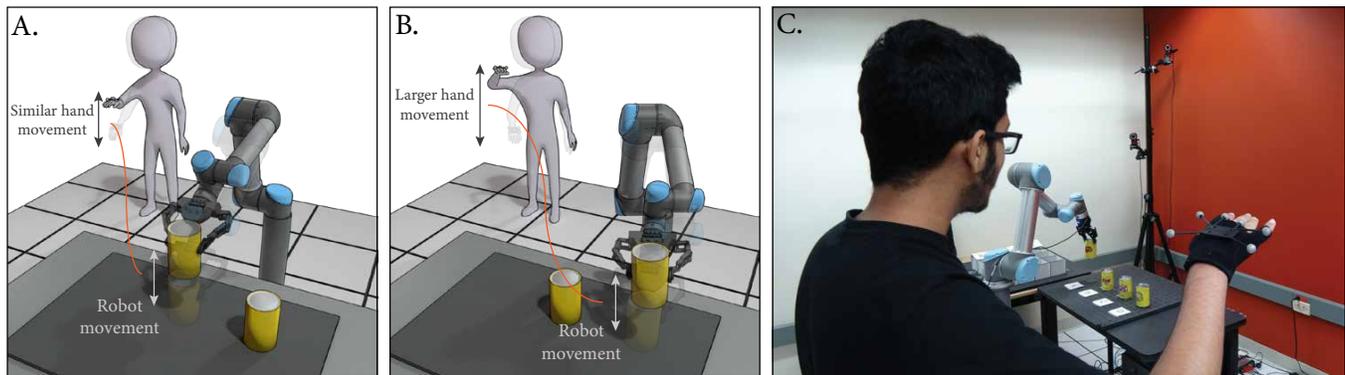


# Supporting Perception of Weight through Motion-induced Sensory Conflicts in Robot Teleoperation

Pragathi Praveena, Daniel Rakita, Bilge Mutlu, and Michael Gleicher  
University of Wisconsin–Madison, Madison, Wisconsin, USA  
{pragathi,rakita,bilge,gleicher}@cs.wisc.edu



**Figure 1:** In robot teleoperation, we propose that a conflict between information from an operator’s proprioceptive and visual senses (how much their hand moves versus how much the robot moves) is an effective cue for communicating weight of objects in a remote environment. *A.* The robot movement and operator hand movement are approximately the same when the object is light. *B.* The operator has to move their hand by a greater distance when the object is heavy, resulting in a *visuo-proprioceptive* weight cue. *C.* We demonstrate the feasibility of using such a cue in four tasks to enhance user performance and experience.

## ABSTRACT

In this paper, we design and evaluate a novel form of visually-simulated haptic feedback cue for communicating weight in robot teleoperation. We propose that a *visuo-proprioceptive cue* results from inconsistencies created between the user’s visual and proprioceptive senses when the robot’s movement differs from the movement of the user’s input. In a user study where participants teleoperate a six-DoF robot arm, we demonstrate the feasibility of using such a cue for communicating weight in four telemanipulation tasks to enhance user experience and task performance.

## CCS CONCEPTS

• **Human-centered computing** → **Interaction paradigms**; • **Computer systems organization** → **Robotics**.

## KEYWORDS

Communicative motion; visuo-proprioceptive cues; pseudo-haptics

## ACM Reference Format:

Pragathi Praveena, Daniel Rakita, Bilge Mutlu, and Michael Gleicher. 2020. Supporting Perception of Weight through Motion-induced Sensory Conflicts in Robot Teleoperation. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20)*, March 23–26, 2020, Cambridge, United Kingdom. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3319502.3374841>

## 1 INTRODUCTION

In *teleroobotics* systems, an operator controls a robot from a distance. This approach is useful in scenarios where complete automation is difficult and cognitive decisions made by an operator is beneficial, such as in unknown or unstructured environments [15]. Teleoperation systems provide operators with a range of feedback to help them experience the environment as if they are encountering it directly. Scenarios where operators need to be provided with feedback on manipulations in a remote workspace are particularly challenging for feedback design. Force feedback is one approach that is commonly used to give the operator a “sense” of the haptic properties of objects in the remote environment. These bilateral control systems have shown great promise in improving user experience and task performance, but they also pose challenges in system stability and can restrict the user’s natural range of motion [15]. An alternative approach, proposed by Lécuyer et al. [10, 12], is *pseudo-force* feedback where haptic feedback is simulated using vision. However, while this approach has shown promise in desktop and virtual reality interfaces, it has never been applied to the teleoperation of a robot.

Our work explores the use of visually simulated haptic feedback for robot teleoperation, specifically to communicate the weight of an object in a remote workspace to the human operator. We propose a method where we manipulate robot motion during direct control of the robot and create a conflict between the information from a user’s proprioceptive and visual senses (how much their hand moves versus how much the robot moves). We design our cue based on early experiments in human perception [3, 25, 28] that suggest that there are salient kinematic cues related to weight perception. A human user controlling a teleoperated robot develops a mental model of how the robot moves with respect to their inputs. Our key idea is that by deliberately manipulating the motion of the robot with respect to what the human operator would expect, we create a communication channel to convey information about the weight of an object, which scaffolds the forming of the user’s mental model about the robot’s motions and capabilities. Because weight can be conveyed with motion cues alone, we have a performance baseline with which to confirm the existence of the visual-proprioceptive effect. Prior work has shown that visually presented haptic cues can be effective in systems with low-latency, in virtual environments, or when controlling objects with fewer degrees of freedom than a robot [5, 12, 26]. However, based on prior literature, it is unclear if the same cues would work in a robot teleoperation setting, which might involve latency, complex movements, and the operator not sharing the robot’s perspective.

In this paper, we present the application and initial evaluation of a *visuo-proprioceptive cue* for communicating weight to a human user teleoperating a robot arm. We present results from a user study where participants completed four tasks in a remote workspace using a six-DoF robot arm. Our results highlight the promise of simulating weight using visual motion cues to enhance user experience and task performance in teleoperation. Additionally, we show that during teleoperation users do not only passively observe these cues but experience a visuo-proprioceptive sensory conflict. These results suggest that sensory conflicts discussed in previous work on optically-simulated haptic feedback can be effective ways of providing operator feedback in robot teleoperation settings.

The contribution of our work is threefold: (1) the conceptualization and application of *visuo-proprioceptive cues* as a design element for creating teleoperation feedback; (2) *data* from an initial evaluation of the visual motion cue in an HRI study; and (3) *insights and design guidelines* for the future use of and further research into visuo-proprioceptive cues as communicative signals in HRI.

## 2 RELATED WORK

Alternatives to force feedback have been studied in human-machine interaction, such as tactile [20], visual [2, 17], auditory [8, 9], and multimodal cues [13, 19]. In addition to providing a better understanding of the remote environment, haptic cues and cues designed using alternate modalities can influence perceptions of the robot. For example, stiffness of an input method can change perceptions of robot speed [24]; the robot’s sound can change perceptions of its competency [30]; and different visual signals from a drone can change perceptions of the robot as a collaborator [29]. For the design of our cue, we draw from prior work in human perception of

kinematic cues to estimate object weight, pseudohaptics, and other optically-simulated haptic feedback.

### 2.1 Human perception

Work in human perception [3, 25, 28] suggests that people can estimate the weight of an object lifted by another person when they visually observe the lifting action. Shim and Carlton [28] and Hamilton et al. [6] additionally suggest that the most salient cue that is used by observers to estimate the weight is variations in velocity when the object is being lifted off the ground. Importantly, kinematic cues have shown to have a strong predominance over intrinsic object properties such as size, material, or fullness for visual weight estimation [1]. Recent studies [18, 27, 32] show promising synthesis of kinematic patterns for robots to communicate weight while lifting objects. Sciutti et al. [27] specifically looked at whether people process human and humanoid-robot lifting actions in a similar manner. The results indicate that robot action reading is likely to rely on the same processes underpinning human action reading and that observers are able to infer weights of unknown lifted objects with similar accuracy in both cases. These studies suggest that humans can interpret robot motion cues to estimate weight. However, there is additional work that suggests that having sensory conflicts between the visual and proprioceptive senses can create stronger cues. We discuss this literature below.

### 2.2 Pseudohaptics

Lécuyer et al. [12] introduced the idea of using passive input devices (do not generate energy like active haptic devices), together with visual feedback to provide an operator with *pseudo-force* feedback. For example, when inserting a virtual cube into a narrow duct, reduction in the cube’s speed as it enters the duct causes the user to instinctively increase the pressure on the input device which results in the feeding back of an increased reaction force by the static device. The simultaneous deceleration of the object on the screen and the increased reaction force from the device gives the user the illusion of friction force between the cube and the duct [12]. In other work, visuo-proprioceptive sensory conflicts have been shown to create illusory perceptions of haptic properties such as weight and stiffness [5, 26]. Pusch and Lécuyer [21] describe *optically simulated haptic feedback* as visually presented haptic effects that have to be “understood” by the user in order to relate their observations to the simulated phenomena. An example is the manipulation of speed and size of a mouse cursor to simulate texture sensations of an image displayed on the computer screen [11]. While these effects may not have strong illusory effects as pseudo-force feedback does, studies have shown them to be effective and to increase user efficiency and satisfaction [31].

We expand this concept of using sensory conflict between the visual and proprioceptive senses to the context of robot teleoperation. We seek to understand the extent to which visuo-proprioceptive sensory conflicts might serve as an effective communicative cue to improve teleoperator performance and experience. Our implementation builds on recent work [22] that shows how arm scale teleoperation systems can be created by mapping the movement of the user’s arm to the robot and that this *mimicry* interface is easy for even novices to use.

### 3 TECHNICAL DETAILS

Our goal is to generate motion cues to communicate object weight and blend them with the human operator’s motion commands. We build on two key findings from prior literature [6, 28]:

- (1) When lifting an object with the hand, the *duration* of the preparation time prior to lifting the object is a salient cue.
- (2) In contrast, when *observing* another person lift an object, the visual information needed to identify the weight is available in the lift motion, especially a lower lift velocity.

Our motion cue manipulates the motion of the robot only in the vertical direction to simulate the effect of gravity. A gain factor,  $g$ , is used to create variations of the motion cue to convey information about a range of weights. A gain factor of 1.0 results in direct control of the robot (no cue), and lower gains signal larger weights.

Our input is a Cartesian workspace path that the operator commands the robot to follow, sampled as  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ . At any given time  $t$ , the path has a position and orientation,  $\mathbf{x}_t \in SE(3)$ . For a given gain,  $g$ , blending the motion cue into the commanded path results in new Cartesian goals  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ , where  $\mathbf{e}_t \in SE(3)$ .  $\mathbf{x}_{t_z}$  and  $\mathbf{e}_{t_z}$  are z-coordinates (corresponding to the vertical direction) of the end-effector pose goals. When the gripper closes around an object,  $\mathbf{e}_1$  is set to  $\mathbf{x}_1$  and the motion cue is applied.

If  $[\mathbf{e}_{t_z} - \mathbf{e}_{1_z}] \leq 1$  cm:

$$\mathbf{e}_{t_z} = \frac{g}{4} * [\mathbf{x}_{t_z} - \mathbf{x}_{t-1_z}] + \mathbf{e}_{t-1_z} \quad (1)$$

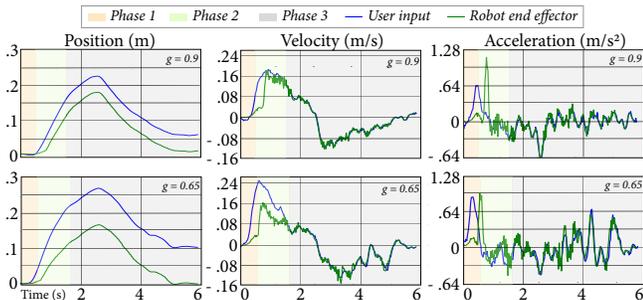
If  $1$  cm  $< [\mathbf{e}_{t_z} - \mathbf{e}_{1_z}] < 10$  cm:

$$\mathbf{e}_{t_z} = g * [\mathbf{x}_{t_z} - \mathbf{x}_{t-1_z}] + \mathbf{e}_{t-1_z} \quad (2)$$

If  $[\mathbf{e}_{t_z} - \mathbf{e}_{1_z}] \geq 10$  cm:

$$\mathbf{e}_{t_z} = [\mathbf{x}_{t_z} - \mathbf{x}_{t-1_z}] + \mathbf{e}_{t-1_z} \quad (3)$$

In our formulation, the lifting and lowering action is made up of three phases. Phase 1, defined by Equation 1 and inspired from the first key finding in prior literature, causes the robot to move a relatively small amount as compared to the input from the user. For example, if the gain is  $g = 0.5$ , for the first centimeter of robot motion, the end-effector moves 12.5% of the distance moved by the operator’s hand. This behavior simulates the effect of inertia when the operator begins to lift the stationary object. In Figure 2, we see a slight lag in the beginning before the robot follows the user input. In Phase 2, defined by Equation 2 and drawn from the second key finding in prior literature, once the object is in motion, the robot will move a fraction (defined by  $g$ ) of the user input. For example, if the gain is  $g = 0.5$ , for the next 9 cm of motion, the robot



**Figure 2: Position, velocity, acceleration of the robot (green) and user hand (blue) at gain = 0.9 (top) and = 0.65 (bottom).**

end-effector moves 50% of the distance moved by the operator’s hand. Similar cues have been implemented in prior literature where  $g$  is called the control-to-display ratio [5, 26]. The motion cue is switched off at a height of 10 cm, which marks the onset of Phase 3 (defined by Equation 3). Figure 2 shows that as  $g$  gets closer to 1.0, the motion profiles become more similar to user input. A lower  $g$ , such as  $g = 0.65$ , results in lower lift velocity.

The parameters for this cue were chosen based on our observations and participant feedback from a pilot study. There is potential to improve the design of the cue by tuning its parameters or manipulating other aspects of the motion (such as acceleration). However, for our initial evaluation, we designed a cue similar to prior work in pseudohaptics to serve as a baseline for evaluating it in our novel setting of robot teleoperation.

### 4 SYSTEM DETAILS

We integrated our cue within an implementation of the mimicry control teleoperation approach proposed by Rakita et al. [22, 23]. This implementation of teleoperation allows users to intuitively and safely control the robot using the free movement of their hands.

*Input Device* – A glove with six motion capture markers attached as a rigid frame on the back of the hand served as the input into our system. Eight Optitrack Flex cameras tracked these markers and provided precise pose information. A Logitech R800 clicker was used to operate the robot gripper and to serve as a clutch.

*Robot* – The user controlled a six-DOF Universal Robots UR5 robot arm equipped with a Robotiq 85 gripper. The robot was operated using velocity control via a network connection at 125 Hz. The Logitech R800 clicker could send open and close commands to the gripper. The gripper speed and force were set to default values.

*System Architecture* – A desktop computer running the Windows OS streamed the input data from the motion capture system. A second desktop computer running the Linux OS received the input motion capture data from the first desktop, computed target joint angle configurations for the robot, and sent them to the controller.

### 5 USER STUDY

To obtain preliminary evidence for the effectiveness of our visually presented haptic cue in providing operators with real-time feedback on object weights, we conducted a human-subjects study in which participants performed several telemanipulation tasks with a robot.

#### 5.1 Hypotheses

Our study aimed to test three hypotheses that we have developed based on findings in prior literature:

**H1:** If two objects are lifted by the participant, one object using direct control and the other with a blended visuo-proprioceptive cue, participants will report a difference between the objects and the perception of weight without being explicitly asked about it.

Prior literature [12, 26] has suggested that participants report “feeling” friction, gravity, or weight due to pseudo-force feedback without explicitly being asked about it.

**H2:** During teleoperation, participants do not only passively observe the visual motion cue, but experience a

stronger visuo-proprioceptive cue that leads to better task performance as compared to visual cues alone.

Previous work [18, 27, 32] has shown that people can interpret motion cues from autonomous robots. In teleoperation, however, the motion cue must be blended with the operator’s motion commands. Although operators may generate different movements while performing the same task, we expect them to be able to interpret the visuo-proprioceptive cue consistently.

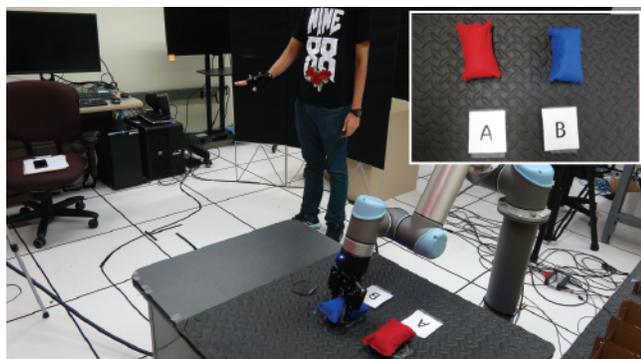
**H3:** The motion cue will improve user experience and task performance compared to receiving no cue because the user has a better understanding of physical properties of the workspace objects through feedback.

If the task requires decision-making based on weight of an object in the remote robot workspace, we expect communicating weight to help the user make better decisions and to reduce workload.

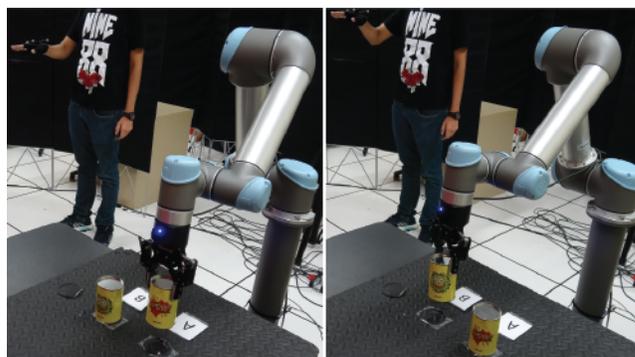
## 5.2 Experimental Design, Tasks, & Procedure

We designed a within-subjects experiment where each participant completed four tasks. The procedure was administered under a protocol reviewed and approved by the Institutional Review Board (IRB) of University of Wisconsin–Madison. Following informed consent, participants were provided with an explanation of the objectives of each task prior to starting it and were asked to fill a set of questionnaires following the task. At the end of the study, participants gave responses to a semi-structured interview about their overall experience. These responses were recorded using a digital voice recorder and transcribed by an experimenter. Two Logitech C930e cameras were used to record the participant and the robot motion. Additionally, a Sony HD camcorder was used to take videos of robot motion for Task 2 where other participants were asked to observe this motion on recorded video. The camera was placed on a tripod close to the participant, and care was taken to ensure that the viewpoint of the remote workspace for the video was similar to the participant’s viewpoint. In Task 2, the videos were presented to the participants on a 42-inch video monitor.

*Task 1* (Figure 3) — To test *H1*, we designed Task 1 such that participants were not explicitly asked about object weight. After each trial of the task, participants freely described their experience,



**Figure 3:** In task 1, participants telemanipulated two bean bags, one associated with a motion cue, and described the differences in their experiences between A and B.



**Figure 4:** In task 2, participants lifted two cans and made a decision as to which can was heavier under two conditions: *Controller* where the participant teleoperated the robot and *Observer* where the participant watched a video of robot from the previous participant’s *Controller* trial.

and we recorded their responses. In Task 1, participants were asked to lift and place down two bean bags with identical weights and material but of different colors (red and blue) placed at locations A and B. The values of  $g$  associated with the two bags were  $g = 0.35$  and  $g = 1.0$  such that one bag was associated with a motion cue while the other was not. Across participants, we counterbalanced the assignment of the motion cue and color of the bags at A and B.

Task 1 also served as a training task for subsequent tasks in this study. Before picking up any object, participants were asked to start with their hand in a waist-high position that was consistent each time to ensure that the starting position did not affect weight perception. Participants lifted the bean bag at A until they heard a beep and then placed the bag back down. The beep was sounded when the object was 10 cm from the starting position. The beep coincided with the motion cue switching off, of which participants were not made aware. Then the participants repeated the process with the bean bag at B. They performed three trials of this task. After the first trial, participants were asked to describe their experience. After the second and the third trials, participants were asked to describe any differences when picking up the two objects or between the objects themselves.

*Task 2* (Figure 4) — To test *H2*, we designed a  $2 \times 1$  within-participants, two-alternate forced choice task. The two conditions that the participants completed were the *Controller* condition where the participant teleoperated the robot and the *Observer* condition where the participant watched a video of robot movement from the previous participant’s *Controller* trial (the first participant used the results of the last pilot participant). The order of the conditions was counterbalanced across participants.

In this task, participants were presented with two cans on the table that were visually similar and had the same weight. Participants were explicitly told that the cans weighed differently and asked to determine the heavier can. We assigned a motion cue corresponding to one of five levels of gain—0.25, 0.35, 0.45, 0.65, and 0.9—that resulted in 15 comparisons, which participants completed in a randomized order. Of the 15 comparisons, 10 were between cans cued with different levels of gains and 5 were between cans cued with the same gain. Participants were asked to perform the

same lifting action as in Task 1, which involved picking the object at A until a beep was heard, placing it back down and repeating this action with the object at B. Participants wrote their response as to which can was heavier and rated their level of confidence with this decision. While the participant completed their responses, the experimenter reset the cans into starting positions before the next trial. After each condition, participants filled out the NASA Task Load Index (TLX) [7] questionnaire.

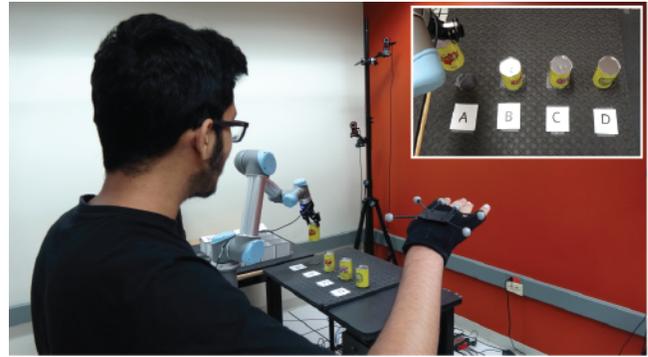
We chose to use the previous participant’s video of this task for the *Observer* condition instead of one or a small set of standard videos. Our rationale was that the utility of our cue depends on the subject’s interpretation of the robot motion. By having a constant stimulus across subjects, the results could be greatly skewed by the properties of that specific choice of stimulus. For example, data from a less fluent user is likely to be noisy and harder to interpret than data from a user with prior robot experience. While different stimuli may potentially introduce noise into the results, we believe that our approach improves the external validity of the experiment.

*Task 3* (Figure 5) – To test *H3*, we designed a  $3 \times 1$  within-participants categorization task. The order of the conditions was randomized across participants. All the cans used in this task were visually similar and weighed the same. Participants were asked to categorize 15 cans (5 cans per condition) into bins labeled “Recycle” or “Stock” based on whether the can was perceived to be light or heavy. In each condition, the first can to be picked by the participant was provided for reference as a can meant for recycling. On average, this task required participants to perform relatively complex robot control to lift and place the cans upright in the bin.

In one condition, participants received our motion cue. The gains associated with the cans were 1.0, 1.0, 0.45, and 0.65. The order of application was randomized across participants. In another condition, participants received a number cue, such that when a can was lifted, a number in the range of 1 to 6 would be revealed behind the can that was proportional to the intended weight of the object. The numbers associated with the cans were 1, 3, 5, and 6. The order of application was randomized across participants. Participants were told that the recycle can was associated with the number 1. In the last condition, participants received no cue regarding the weight of the object. After each condition, participants filled out the NASA TLX questionnaire.



**Figure 5: In task 3, participants were asked to categorize five cans into bins labeled “Recycle” (for light cans) or “Stock” (for heavy cans) based on the perceived weight.**



**Figure 6: In task 4, participants were asked to lift four cans (A, B, C, D) and order them based on the perceived weight.**

*Task 4* (Figure 6) – To test *H3*, we designed a  $3 \times 1$  within-participants rank-ordering task. The order of the conditions was randomized across participants. The participants were asked to lift four cans (A, B, C, D) and order them based on the perceived weight. Participants were not asked to rearrange the cans but to write down their rank-order in the questionnaire. They were asked to keep the number of lifts to a minimum and to write down an answer as soon as they were confident. No limit was imposed on the number of lifts because an upper limit was not known *a priori*. Participants were allowed to take notes in between trials, so that they did not have to memorize the order.

In one condition, participants received our motion cue with gains 0.25, 0.35, 0.45, and 0.65. The order of application was randomized across participants. In another condition, participants received no cue regarding the weight of the object. All cans used across the *Motion cue* and *No cue* conditions were visually similar and weighed the same. In the last condition, which we call *Haptic cue*, participants were asked to perform the ordering task with their hand with four cans that were visually similar and filled with bags of rice weighing 50 g, 100 g, 150 g, and 200 g. The order of application was randomized across participants. After each condition, participants filled out the NASA TLX questionnaire.

Tasks 3 and 4 were designed to test *H3* from different perspectives; Task 3 required complex robot movements, while Task 4 challenged participant sensitivity to small differences in perceived weight. The *Number cue* and *Haptic cue* conditions served as two different gold standards for the tasks by answering the question, what is the task performance with maximum information?

*Missing data* – One participant did not fill out the NASA TLX questionnaire for Task 2, which resulted in 10% of measures for workload missing for Task 2. One participant did not complete Tasks 3 and 4, resulting in 10% of the data from these tasks missing. Their data for Tasks 1 and 2 have been included in our analysis.

### 5.3 Measures & Analyses

The measures used in the tasks are described here, along with the task numbers that are relevant to each measure.

*Accuracy (T2, T3)* – We compared subject responses to virtual weight that was introduced due to the motion cue. Participant responses were converted to a binary scale, 1 corresponding to a right answer and 0 corresponding to a wrong answer. A response

was considered correct when a can associated with a lower gain (larger virtual weight) is chosen as the heavier object (*T2*) or the object to be stocked (*T3*). In the *No cue* condition in Task 3, the cans were assigned a gain, but no cue was applied, and subject answers were compared to this assignment. The intention was to establish a baseline condition close to chance to which the *Motion cue* condition could be compared.

*Rank-ordering score (T4)* – Spearman’s rank correlation [14] was used to assign a score to the participant’s rank-ordering. A score of 1 was assigned if the rank-order is correct and a score of -1 is assigned if the rank-order is reversed. Other scores fall within the range of (-1, 1). Similar to the *Accuracy* measure, subject responses were compared to virtual weight that was introduced based on the value of the gain associated with the motion cue. To ensure that all scores were positive, we added 1 to the Spearman’s rank correlation measure, which resulted in a numeric score between 0 and 2 that can be analyzed using parametric tests.

*Number of lifts (T4)* – The number of lifts for each condition was counted from the video that was recorded from each participant.

*Perceived confidence (T2, T3, T4)* – Participants answered a single question measured on a seven-point rating scale with two opposing anchors (1 = strongly disagree; 7 = strongly agree) and levels of numbers in between. Although controversial, Carifio and Perla [4] and Norman [16] support the use of parametric tests for such scales.

*Workload (T2, T3, T4)* – We used the NASA Task Load Index (TLX) [7] to assess user’s perceived workload. The total workload is divided into six subscales—mental demand, physical demand, temporal demand, performance, effort, and frustration—that are each measured on a 100-point rating scale, where higher ratings indicate more workload. The scores were averaged with equal weighting to calculate the TLX (shown in Figure 9).

For Task 2, we analyzed data from accuracy of participant responses and perceived confidence using repeated-measures chi-squared test, including the condition, *Controller* or *Observer*, as a within-participant variable. For other measures, we analyzed data using one-way repeated measures analyses of variance (ANOVA), including the type of cue, *No cue*, *Motion cue* or *Number cue/Haptic cue*, as a within-participants variable. If the ANOVA test showed significant differences, we used Tukey’s HSD test to determine where the differences lied while accounting for multiple comparisons.

## 5.4 Participants

We recruited 10 participants (5 male, 5 female) from the University of Wisconsin–Madison campus between the ages of 18 and 49 ( $M = 24.8, SD = 9.58$ ). Participants reported low familiarity with robots ( $M = 2.9, SD = 1.79$ , measured on a seven-point scale). Two participants were left-handed. Three participants reported an interaction with a robot in prior robotics research studies. The study took 90 minutes, and all participants received \$ 15 as compensation.

## 6 RESULTS

In this section, we present the qualitative and quantitative results from the four tasks in our user study.

*Task 1* – At the end of three trials of teleoperating a robot to lift one cued and another non-cued bag, two out of ten participants

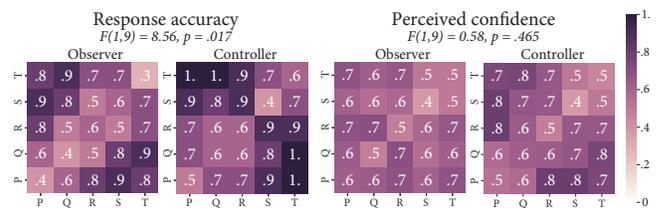
reported a perception of weight for the object with the motion cue. They mentioned that (participant number in boldface),

**P4**: “B [cued bag] is much heavier.”

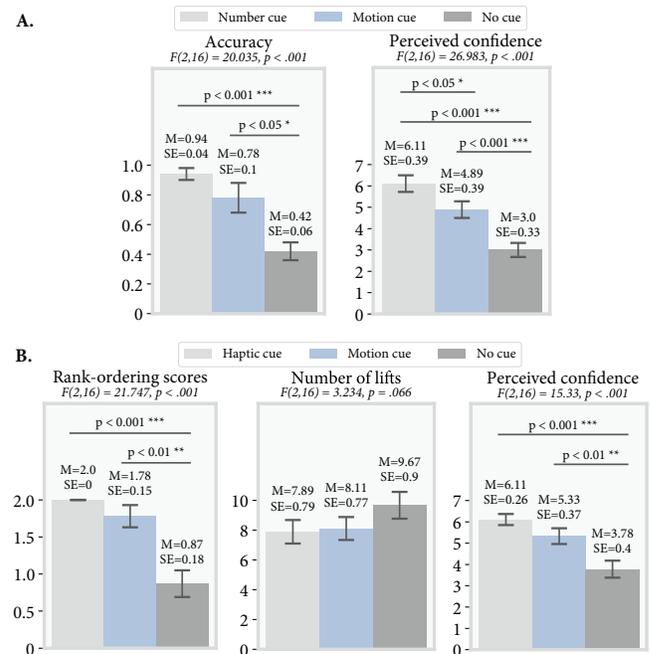
**P5**: “feel like it’s a lot lighter... the blue bag [non-cued bag]. Because it [robot] would bring it up a lot faster.”

After the first trial, only three out of ten participants reported any difference between their experiences lifting the two bean bags. However, their observations were not related to the motion cue, but rather the difficulty of movement due to lack of practice or bad positioning. For example,

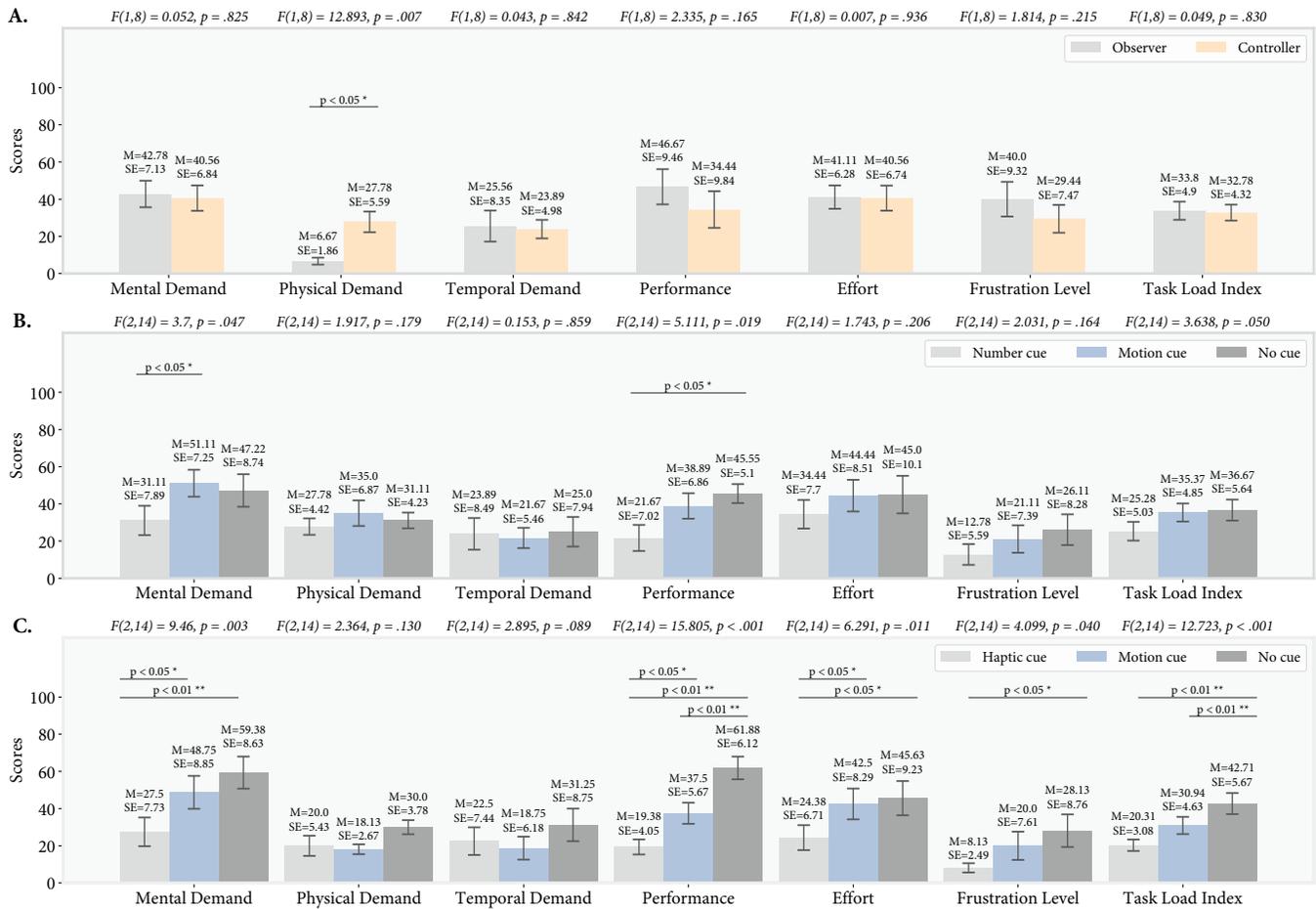
**P8**: “I like that the joint there goes with my own arm. I thought that was very intuitive. And I got it easier with the second bean bag for some reason in this practice.”



**Figure 7: Heatmaps for performance measures for Task 2 (alternative forced choice). The Response accuracy heatmaps show the fraction of correct answers for each comparison present in the task. The Perceived confidence heatmaps show the normalized perceived confidence of the participants in their responses. The levels of gains used in the task are 0.25 (P), 0.35 (Q), 0.45 (R), 0.65 (S), and 0.9 (T).**



**Figure 8: Bar charts with standard errors for performance measures for (A.) Task 3 (binning) and (B.) Task 4 (sorting). Horizontal lines indicate significant Tukey HSD test results.**



**Figure 9: Results from the NASA TLX questionnaire are shown as bar charts with standard errors for (A.) Task 2 (2-alternative forced choice), (B.) Task 3 (binning) and (C.) Task 4 (sorting). Horizontal lines indicate significant Tukey HSD test results.**

After the second and third trial, eight out of ten participants reported the effects of our *visuo-proprioceptive* cue. For example,

**P1:** “The robot arm was a bit slow to go up compared to my hand, it was responding very slowly for the red bag [cued bag], but otherwise the rest of the movements were, I think, smoother than before.”

**P2:** “I feel that I need to make more effort to grab B [cued bag], because of this movement that it does.”

**P3:** “With A [cued bag] it felt like slower or something. It felt like my arm was a lot higher than the robot’s arm was”

Only two participants (P4 and P5 quoted earlier) attributed this difference in robot movement to difference in object weights.

**Task 2** – In Figure 7, our results show that between the two conditions, *Controller* and *Observer*, there exist significant differences in accuracy but not in the perceived confidence. The *Response* heat maps reveal that choosing the heavier object is harder when the two objects are cued using gains that are closer together, as illustrated by the lighter colors nearer to the diagonal running from the bottom left corner to the top right corner. This effect is particularly visible in the *Observer* condition. For comparisons between cans that are cued with the same gain value, the can at B is considered heavier

and as the correct answer. We see on the diagonal of the heatmaps that the accuracy is close to chance, which also verifies that the position of the can on the table does not influence weight perception. With regard to significant differences in user experience between the two conditions, as seen in Figure 9-A, participants report a higher physical demand in perceived workload in the *Controller* condition.

**Task 3** – In Figure 8-A, we see that there is a significant difference in accuracy of response and perceived confidence between the conditions where participants received a cue (either the *Number cue* or *Motion cue*) and no cue. Since we expected perfect accuracy in the *Number cue* condition, the three out of nine participants (P1, P2, P7) who made errors were probed about their performance. The participants reported that the number cue was inconsistent with how heavy the can “felt.” Additionally, in the *No cue* condition, three out of nine participants (P2, P5, P7) deviated from the instructions that required them to restrict their movement to lifting the can and placing it in the appropriate bin. In the absence of any cue, they swayed the can back and forth mid-air to discern the weight of the object. In Figure 9-B, we see no significant differences in perceived workload between the *Motion cue* and *No cue* conditions.

Presence of the *Number cue* leads to lower perceived mental workload as compared to providing the *Motion cue* and better perceived performance as compared to providing *No cue*.

*Task 4* — Figure 8-B shows that participants have higher ranking scores and higher perceived confidence in both the *Haptic cue* and *Motion cue* conditions as compared to the *No cue* condition. In Figure 9-C, we see that participants also experience a higher overall perceived workload in the absence of a cue. Between the conditions where participants received a cue, participants reported higher performance and lower effort in the *Haptic cue* condition where they received direct haptic feedback through their hand.

## 7 DISCUSSION

Our study assessed participant performance and experience on a range of tasks, progressing from simple to more challenging tasks, and showed support for hypotheses H2 and H3.

**H1:** While prior literature suggests strong illusory effects of pseudohaptics in VR and screens, the results of Task 1 suggest that, in robot teleoperation settings, the visuo-proprioceptive cue is subtle. Specifically in prior work, the strength of the cue may have been greater because subjects used a grounded input device (such as a mouse) that could provide a passive reactive force [12], or other settings may provide better shared perspective between the visual and proprioceptive channels, such as in VR [26]. While a different input device or a better viewpoint of the robot workspace could improve the strength of the cue, we posit that the cue is unlikely to lead to a strong illusory effect in a teleoperation setting because of the complexity of movement associated with robot control. The effect of the visuo-proprioceptive cue may also be more obvious to expert users who have to focus less on robot control.

**H2:** The results from Task 2 support our hypothesis that participants do not only passively observe the visual motion cue (in the *Observer* condition) but experience a stronger visuo-proprioceptive cue (in the *Controller* condition). This stronger cue leads to better task performance as compared to visual cues alone, especially when it comes to being able to differentiate between different levels of the cue. While these visuo-proprioceptive sensory conflicts do not necessarily result in a haptic illusion, operators rely on a more cognitive interpretation of the sensory conflict. Participants in the *Controller* condition were able to quickly relate their observations to perceived variations in object weight without any explicit learning. Importantly, we observe that different operators generate different kinds of movements while doing the same task but are systematically able to interpret the motion cue provided by our approach. The finding that in the *Controller* condition the cue relies on both the visual and proprioceptive senses is also consistent with the higher physical demand reported by participants.

**H3:** The results of Tasks 3 and 4 support the hypothesis that our cue improves user experience and task performance when compared to receiving no cue. Participants are more accurate and confident in their performance when provided with our motion cue as compared to no cue. In the more challenging task (Task 4), they also reported lower perceived workload when provided with the cue. In the task with complicated motion paths (Task 3), the cue was often missed when complex robot control was required. When participants lifted the objects with their hand, four out of

nine participants stalled at the apex of the lift to discern the weight. Our current motion cue switches off near the apex. Adding a sense of downward pull when the object is stationary at the apex may be an additional design consideration. Other improvements may come from additional feedback when an object is being shaken or extending the motion cue for a longer time period if the operator is unsure of their decision.

*Implications*—In this paper, we explored how visually simulated haptic feedback can be used within robot teleoperation interfaces. Our findings highlight visuo-proprioceptive cues as a rich space for HRI research and design and suggest several insights for future work. We summarize these insights below:

- (1) Novice users notice the cue, but do not automatically associate it with weight, which may change with expertise. The illusory effect of pseudohaptics may be difficult to obtain in teleoperation due to complex robot control and movement.
- (2) Once participants are aware of the haptic property they are looking for, they are able to recognize and use the visuo-proprioceptive cue and make consistent decisions.
- (3) Participants are not only passively observing robot motion during teleoperation, but they rather rely on the coupling between the visual and proprioceptive feedback.
- (4) The current design of the cue may not be readily noticed in the presence of complex robot movement. While more complex cues have not been explored in prior pseudohaptics literature to communicate weight, the existence of complex movements in a teleoperation setting present both the necessity and the opportunity to develop better cues.
- (5) As tasks get more challenging, presence of the motion cue is more valuable. This outcome is demonstrated by more significant differences in user experience and task performance measures for the conditions with and without the cue.

*Limitations*—Our work has many limitations. First, our study used data from a small number of novice participants, and further research with larger samples of people with a range of teleoperation skills would widen our understanding and extend the generalizability of our findings. Additionally, while our study shows that visuo-proprioceptive cue conflict can be used to convey information in a teleoperation system, more work is needed to demonstrate its utility in real-world teleoperation scenarios. Specifically, future work must establish the utility of using the cue conflict for conveying actual weight. Finally, our work explored a single example of a visuo-proprioceptive cue. In the future, we hope to explore different ways to map between user movements and resulting motion to understand the design space of sensory conflicts that can be created. We also hope to use visuo-proprioceptive cues to provide other optically simulated haptic feedback, for example, to show stiffness, friction, or texture.

## ACKNOWLEDGMENTS

We thank our colleagues Bengisu Cagiltay, Curt Henrichs, Olivia Zhao and Luis Molina for assistance with running the experiments. We thank the anonymous reviewers for their excellent feedback. This research was supported in part by NSF award 1830242 and NASA University Led Initiative (ULI) grant 80NSSC19M0124.

## REFERENCES

- [1] Kaat Alaerts, Stephan P. Swinnen, and Nicole Wenderoth. 2010. Observing how others lift light or heavy objects: which visual cues mediate the encoding of muscular force in the primary motor cortex? *Neuropsychologia* 48, 7 (2010), 2082–2090.
- [2] Angelica I. Aviles-Rivero, Samar M. Alsaleh, John Philbeck, Stella P. Raventos, Naji Younes, James K. Hahn, and Alicia Casals. 2018. Sensory Substitution for Force Feedback Recovery: A Perception Experimental Study. *ACM Trans. Appl. Percept.* 15, 3, Article Article 16 (April 2018), 19 pages. <https://doi.org/10.1145/3176642>
- [3] Geoffrey P. Bingham. 1993. Scaling Judgments of Lifted Weight: Lifter Size and the Role of the Standard. *Ecological Psychology* 5, 1 (1993), 31–64.
- [4] James Carifio and Rocco Perla. 2008. Resolving the 50-year debate around using and misusing Likert scales. *Medical education* 42, 12 (2008), 1150–1152.
- [5] Lionel Dominjon, Anatole Lécuyer, J-M Burkhardt, Paul Richard, and Simon Richir. 2005. Influence of control/display ratio on the perception of mass of manipulated objects in virtual environments. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005*. IEEE, 19–25.
- [6] A.F. de C. Hamilton, D.W. Joyce, J.R. Flanagan, C.D. Frith, and D.M. Wolpert. 2007. Kinematic cues in perceptual weight judgement and their origins in box lifting. *Psychological research* 71, 1 (2007), 13–21.
- [7] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*, Peter A. Hancock and Najmedin Meshkati (Eds.). Advances in Psychology, Vol. 52. North-Holland, 139–183.
- [8] Masaya Kitagawa, Daniell Dokko, Allison M Okamura, and David D Yuh. 2005. Effect of sensory substitution on suture-manipulation forces for robotic surgical systems. *The Journal of thoracic and cardiovascular surgery* 129, 1 (2005), 151–158.
- [9] Katherine J Kuchenbecker, Jamie Gewirtz, William McMahan, Dorsey Standish, Paul Martin, Jonathan Bohren, Pierre J Mendoza, and David I Lee. 2010. VerroTouch: High-frequency acceleration feedback for telerobotic surgery. In *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*. Springer, 189–196.
- [10] Anatole Lécuyer. 2009. Simulating haptic feedback using vision: A survey of research and applications of pseudo-haptic feedback. *Presence: Teleoperators and Virtual Environments* 18, 1 (2009), 39–53.
- [11] Anatole Lécuyer, Jean-Marie Burkhardt, and Chee-Hian Tan. 2008. A study of the modification of the speed and size of the cursor for simulating pseudo-haptic bumps and holes. *ACM Transactions on Applied Perception (TAP)* 5, 3 (2008), 14.
- [12] Anatole Lécuyer, Sabine Coquillart, Abderrahmane Kheddar, Paul Richard, and Philippe Coiffet. 2000. Pseudo-haptic feedback: can isometric input devices simulate force feedback?. In *Proceedings IEEE Virtual Reality 2000 (Cat. No. 00CB37048)*. IEEE, 83–90.
- [13] Anduaem Tadesse Maereg, Atulya Nagar, David Reid, and Emanuele L Secco. 2017. Wearable vibrotactile haptic device for stiffness discrimination during virtual interactions. *Frontiers in Robotics and AI* 4 (2017), 42.
- [14] John H McDonald. 2009. *Handbook of biological statistics*. Vol. 2.
- [15] Günter Niemeyer, Carsten Preusche, Stefano Stramigioli, and Dongjun Lee. 2016. Telerobotics. In *Springer handbook of robotics*. Springer, 1085–1108.
- [16] Geoff Norman. 2010. Likert scales, levels of measurement and the “laws” of statistics. *Advances in health sciences education* 15, 5 (2010), 625–632.
- [17] Allison M Okamura, Lawton N Verner, Tomonori Yamamoto, James C Gwilliam, and Paul G Griffiths. 2011. Force feedback and sensory substitution for robot-assisted surgery. In *Surgical Robotics*. Springer, 419–448.
- [18] Oskar Palinko, Alessandra Sciutti, Laura Patané, Francesco Rea, Francesco Nori, and Giulio Sandini. 2014. Communicative lifting actions in human-humanoid interaction. In *2014 IEEE-RAS International Conference on Humanoid Robots*. IEEE, 1116–1121.
- [19] Bernd Petzold, Michael F Zaeh, Berthold Faerber, Barbara Deml, Hans Egermeier, Johannes Schilp, and Stella Clarke. 2004. A study on visual, auditory, and haptic feedback for assembly tasks. *Presence: Teleoperators & Virtual Environments* 13, 1 (2004), 16–21.
- [20] Domenico Prattichizzo, Claudio Pacchierotti, and Giulio Rosati. 2012. Cutaneous force feedback as a sensory subtraction technique in haptics. *IEEE Transactions on Haptics* 5, 4 (2012), 289–300.
- [21] Andreas Pusch and Anatole Lécuyer. 2011. Pseudo-haptics: From the Theoretical Foundations to Practical System Design Guidelines. In *Proceedings of the 13th International Conference on Multimodal Interfaces (ICMI '11)*. ACM, New York, NY, USA, 57–64.
- [22] Daniel Rakita, Bilge Mutlu, and Michael Gleicher. 2017. A motion retargeting method for effective mimicry-based teleoperation of robot arms. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 361–370.
- [23] Daniel Rakita, Bilge Mutlu, and Michael Gleicher. 2018. RelaxedIK: Real-time Synthesis of Accurate and Feasible Robot Arm Motion. In *Proceedings of Robotics: Science and Systems*. Pittsburgh, Pennsylvania. <https://doi.org/10.15607/RSS.2018.XIV.043>
- [24] Daniel J Rea and James E Young. 2018. It’s All in Your Head: Using Priming to Shape an Operator’s Perceptions and Behavior during Teleoperation. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 32–40.
- [25] Sverker Runeson and Gunilla Frykholm. 1981. Visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance* 7, 4 (1981), 733.
- [26] Majed Samad, Elia Gatti, Anne Hermes, Hrvoje Benko, and Cesare Parise. 2019. Pseudo-Haptic Weight: Changing the Perceived Weight of Virtual Objects By Manipulating Control-Display Ratio. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 320.
- [27] Alessandra Sciutti, Laura Patane, Francesco Nori, and Giulio Sandini. 2014. Understanding object weight from human and humanoid lifting actions. *IEEE Transactions on Autonomous Mental Development* 6, 2 (2014), 80–92.
- [28] Jaeho Shim and Les G. Carlton. 1997. Perception of Kinematic Characteristics in the Motion of Lifted Weight. *Journal of Motor Behavior* 29, 2 (1997), 131–146.
- [29] Daniel Szafir, Bilge Mutlu, and Terrence Fong. 2015. Communicating directionality in flying robots. In *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 19–26.
- [30] H. Tennent, D. Moore, M. Jung, and W. Ju. 2017. Good vibrations: How consequential sounds affect perception of robotic arms. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 928–935. <https://doi.org/10.1109/ROMAN.2017.8172414>
- [31] Koert van Mensvoort, Dik J Hermes, and Maurice van Montfort. 2008. Usability of optically simulated haptic feedback. *International Journal of Human-Computer Studies* 66, 6 (2008), 438–451.
- [32] Allan Zhou, Dylan Hadfield-Menell, Anusha Nagabandi, and Anca D Dragan. 2017. Expressive robot motion timing. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 22–31.