

Combating Spatial Disorientation in a Dynamic Self-Stabilization Task Using AI Assistants

Sheikh Mannan
sheikh.mannan@colostate.edu
Colorado State University
Fort Collins, CO, USA

Paige Hansen
p.hansen@colostate.edu
Colorado State University
Fort Collins, CO, USA

Vivekanand Pandey Vimal
somde@brandeis.edu
Brandeis University
Waltham, MA, USA

Hannah N. Davies*
hannahndavies@gmail.com
Emory University School of Medicine
Atlanta, GA, USA

Paul DiZio
dizio@brandeis.edu
Brandeis University
Waltham, MA, USA

Nikhil Krishnaswamy
nkrishna@colostate.edu
Colorado State University
Fort Collins, CO, USA

ABSTRACT

Spatial disorientation is a leading cause of fatal aircraft accidents. This paper explores the potential of AI agents to aid pilots in maintaining balance and preventing unrecoverable losses of control by offering cues and corrective measures that ameliorate spatial disorientation. A multi-axis rotation system (MARS) was used to gather data from human subjects self-balancing in a spaceflight analog condition. We trained models over this data to create “digital twins” that exemplified performance characteristics of humans with different proficiency levels. We then trained various reinforcement learning and deep learning models to offer corrective cues if loss of control is predicted. Digital twins and assistant models then co-performed a virtual inverted pendulum (VIP) programmed with identical physics. From these simulations, we picked the 5 best-performing assistants based on task metrics such as crash frequency and mean distance from the direction of balance. These were used in a co-performance study with 20 new human subjects performing a version of the VIP task with degraded spatial information. We show that certain AI assistants were able to improve human performance and that reinforcement-learning based assistants were objectively more effective but rated as less trusted and preferable by humans.

CCS CONCEPTS

• **Human-centered computing** → **Interaction paradigms; Empirical studies in HCI**; • **Computing methodologies** → *Control methods; Learning paradigms.*

KEYWORDS

Spatial disorientation, Balancing, AI assistance, Real-time human-agent interaction

ACM Reference Format:

Sheikh Mannan, Paige Hansen, Vivekanand Pandey Vimal, Hannah N. Davies*, Paul DiZio, and Nikhil Krishnaswamy. 2024. Combating Spatial

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

HAI '24, November 24–27, 2024, Swansea, United Kingdom

© 2024 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-1178-7/24/11.

<https://doi.org/10.1145/3687272.3688329>

Disorientation in a Dynamic Self-Stabilization Task Using AI Assistants. In *International Conference on Human-Agent Interaction (HAI '24)*, November 24–27, 2024, Swansea, United Kingdom. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3687272.3688329>

1 INTRODUCTION

Maintaining spatial awareness and orientation is critical in domains like piloting, spaceflight, and even driving. Spatial disorientation has been and continues to be a leading cause of fatal aircraft incidents [7, 15] and occurs when sensory information (e.g. from the visual, somatosensory, and vestibular systems) is erroneous, which can lead to unrecoverable crashes, injury, or loss of life [15, 30].

An AI agent in this situation could potentially use numerical signals to track the pilot and vehicle’s positioning in the relevant orientational plane(s), detect if there is a risk of losing control [10, 58, 59], and even alert the pilot to make corrective maneuvers. However, there is a record, particularly in high-risk domains like aviation, of either over-trust or under-trust in highly automated systems. For instance, Sadler et al. [43] demonstrated that pilots’ trust in the recommendations of an automated system correlated with the level of transparency (such as justification) in the recommendation. The *shared autonomy* literature indicates that even when an agent knows an optimal strategy, failing to comply with a suboptimal strategy its human partner insists on may have a negative effect on trust and lead to disuse of the system [17, 22, 31, 44].

In this paper, we hypothesize that when attempting to balance themselves under disorienting conditions, humans will be more receptive to assistance from an AI whose strategy to regain balance is more human-like, even if that strategy is objectively less optimal than a less human-like strategy. We establish a novel task of AI assistance in regaining balance in disorienting situations using a documented, realistic simulation of vehicle control in a spaceflight analog condition. In this scenario, subjects are deprived of gravitational cues and use a joystick to self-balance while seated in a *multi-axis rotation system* (MARS) programmed to behave like an inverted pendulum [34, 51, 52]. We took data of human subjects attempting to keep the MARS balanced and used it to train digital twins of “pilots” that exemplified performance characteristics of humans of various proficiency levels. We then trained multiple “assistant” models that, due to different data and techniques, demonstrate performance strategies that may align with or differ

*This work performed at Brandeis University.

drastically from those of humans. The pilot and assistant models were placed in a co-performance simulation with a *virtual inverted pendulum* (VIP) programmed with identical physics to the MARS. Assistants attempted to help the pilot models keep the VIP balanced and avert crashes by offering corrective cues when the pilot risked a destabilizing loss of balance. To test transfer to a different environment, we ran assistants with digital twins trained over data from humans performing an instance of the VIP task directly. Finally, the best-performing assistant models were deployed in a human-subject study in the VIP environment to demonstrate the feasibility of AI assistance as a countermeasure in disorienting situations.

Following a discussion of related literature (Sec. 2), in Sec. 3 we describe the MARS and VIP tasks, the human performance data collected from each, and how performance is evaluated in these tasks. Sec. 4 describes the training of different reinforcement learning models using an analogous environment to the MARS/VIP tasks, and deep learning models using human performance data, as well as the selection of “digital twin” models of different human proficiencies. Sec. 5 presents the evaluation methods and results for two studies: the high-throughput digital twins study wherein 21 candidate assistant models made corrective suggestions to the different digital twins of human pilots, and a human-subject study wherein the five best performing assistants from the digital twins study engaged in co-performance and co-training in the VIP task with real humans. In Sec. 6 and Sec. 7, we discuss our findings and their implications for human-agent collaboration and trust when performing real-time situated tasks, and directions for future work.¹

Our novel contributions are: 1) a novel task of AI assistance in disoriented self-balancing, a challenging action-learning task with well-controlled parameters; 2) an assessment of different reinforcement learning and deep learning models’ abilities to prevent destabilizing loss of control in this task using a high-throughput digital twins setting; 3) a human-agent co-performance study with the best-performing AI assistants involving human-in-the-loop (HITL) AI training, to demonstrate transfer to real humans.

2 RELATED WORK

Spatial disorientation and balance. An inverted pendulum (with center of mass above the pivot point) is a common model of human upright balance in the study of postural dynamics [40]. Panic et al. [34] and Vimal et al. [53] explored the relevance of the MARS balancing task to the perception of gravitational cues during unstable vehicle control. Subjects were strapped into a MARS device programmed with inverted pendulum (IP) dynamics and instructed to stabilize themselves about the direction of balance (DOB) using a joystick. Because the risk of spatial disorientation-related accidents is heightened when visual information is limited [25, 46], subjects were blindfolded. Typically, humans rely on gravitational cues when balancing, which are detected by the vestibular and somatosensory systems as participants tilt away from the gravitational vertical, however in spaceflight conditions gravitational cues are not reliable. To create a disorienting spaceflight analog condition, Panic et al. [33] and Vimal et al. [51, 52] placed participants in the Horizontal Roll Plane, where they were always perpendicular to the

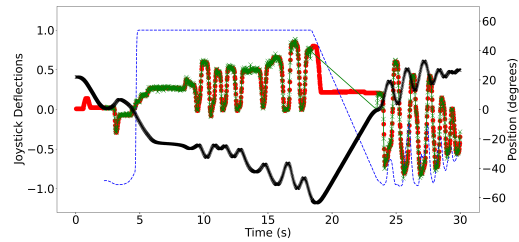


Figure 1: Joystick deflections predicted by a DDPG (blue) and an LSTM trained over human data (green) compared to an actual 30-sec. MARS participant trial sample (participant deflections in red and angular position in black). This instance of the LSTM displays a test RMSE of .013 while the DDPG gets .803.

gravitational vertical and no longer tilting relative to it. 90% of participants reported spatial disorientation and in data 100% exhibited characteristic positional drifting [52]. Participants showed minimal learning and frequent “crashes” (reaching pre-programmed $\pm 60^\circ$ boundaries, after which the MARS automatically reset to the DOB).

AI algorithms. Tasks like IP balancing are well-known use cases in reinforcement learning [2, 5, 13] that serve as demonstration benchmarks for newer continuous control algorithms like SAC [16] or DDPG [23]. These have demonstrated proficiency at solving non-linear control tasks like IP balancing using reward signals extracted from observation of applied environmental physics. However, due to the nature of environmental physics input vs. sensorimotor input, we hypothesize that they learn to perform the task very differently from humans. Here we explore the application of different RL and deep learning algorithms to AI assistance in disoriented balancing.²

Embodiment. Seminal literature operationalizes embodiment as a two-way process between brain and body/environment [47, 50]. Most modern approaches treat embodiment primarily at a surface level, such as the richness of visual [3, 4, 18, 19] or interoceptive [1] representation of an agent’s form, or in terms of the physical form an agent takes, wherein the actions it is capable of are conditioned upon its articulators [11, 36, 37]. We adopt a definition of embodiment akin to Ziemke’s *structural coupling* [60], where a system embodied in an environment takes environmental states as input, which changes the condition of the system [38]. However, the types of information a system is exposed to, such as through different types of sensors, also condition the relations the system develops between itself and its environment [8], such that exposure to different types of data (or different ways of measuring the same underlying environmental state) may mean that different inputs and modeling strategies cause a model to learn different policies within the same action space, and thus may learn to perform the same task equally well through potentially radically different strategies. Therefore we will speak of different “embodiments” of the task problem space as reflected through these different strategies.

²Although the RL approaches we use are also parameterized by multilayer neural networks, we use “deep learning” to refer to non-reinforcement learning algorithms, such as sliding window or sequence modeling approaches.

¹Our code and data is available at <https://github.com/cs-u-signal/HITL-VIP>.

Fig. 1 shows how this manifests in the MARS task. “Actual” shows the deflections of a human subject—the subject balanced the MARS using small, intermittent deflections that are characteristic of proficient human performance [55]. Joystick deflections predicted by a DDPG model are shown in blue and deflections predicted by an LSTM trained over actual human motions from other MARS trials are shown in green. The model trained on human data *embodies* the problem space similarly to a proficient human, making it a superior predictor of novel humans’ small, intermittent actions, while the DDPG predicts long, continuous deflections that are more indicative of poor human performance, even though a DDPG successfully performs the task independently (see also [27]). In this paper, we assess if and how an AI model’s ability to embody a task space similarly to a human translates to increased ability to assist a human in the task or greater human preference for that assistant compared to others.

Trust. Life-critical situations involving dynamic systems, such as aviation use cases, may require rapid decisions that may have irreversible consequences. According to Schneiderman [45]’s human-centered AI (HCAI) framework, systems deploying high levels of autonomy would need to be *reliable, safe, and trustworthy* (RST) within known parameters. Human users must also not become over-reliant on automation such that if the system makes a mistake, the human can still override the AI using situational knowledge and value judgments. Furthermore, automated systems may prompt over-reliance in emergency scenarios even after the systems have demonstrated a lack of reliability [41, 48, 57], or an under-reliance on a typically robust system after a single failure point [35].

3 MARS AND VIP TASKS

Fig. 2 compares the MARS and VIP paradigms. Both can be configured in challenging but non-disorienting modes, with standard sensory information; or difficult and disorienting modes, with degraded information. In both disorienting conditions, subjects show the same characteristic drifting and lack of learning when compared to the coherent conditions.

The **MARS task** is described in Sec. 2. In the data we use here (described below), MARS dynamics were governed by $\ddot{\theta} = k_P \sin\theta$, where θ is degrees deviation from the DOB and pendulum constant $k_P = 600^\circ/s^2$.

In the **virtual inverted pendulum (VIP)** paradigm, an analog to the MARS programmed with the same physics, subjects balance a visually simulated circular array of dots (random dot kinematogram, RDK) which rolls in the plane of the display screen. This is visually rendered for humans and can be directly actuated by an algorithm. In the disorienting VIP condition (similar to the Horizontal Roll Plane in the MARS), the RDK is 50% coherent: alternating subsets of dots displace coherently across consecutive frames while the other half jump randomly. This eliminates configural displacement cues relative to the upright DOB while providing low-level retinal motion cues. Similar performance degradations occur between MARS upright vs. supine and VIP 100% vs. 50% coherence conditions, even with practice, exhibited by an increased number of crashes accompanied by more frequent destabilizing actions (Fig. 2).

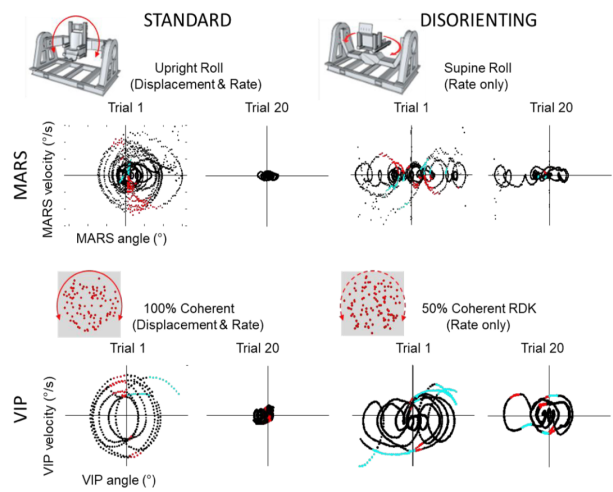


Figure 2: Typical performance in the MARS and VIP tasks, before practice (Trial 1) and after practice (Trial 20). Phase plots show angular velocity vs. angular displacement from the DOB. The “standard” conditions provide angular displacement and velocity cues, and subjects improve significantly between first and last trial, seen as clustering around the origin (balance point) by Trial 20. The “disorienting” conditions eliminate sensory signals about displacement from the DOB, increasing positional drift (shown as phase loop oscillations around the X-axis) and destabilizing joystick commands that accelerate away from the DOB in the current direction of motion, with minimized learning and continued positional drift in Trial 20. Cyan dots indicate destabilizing deflections, where position, velocity and joystick deflection all have the same sign. Red dots denote *anticipatory* deflections, where position and joystick deflection have the same sign but velocity has the opposite sign—usually done to slow the IP down when velocity is perceived as being too high.

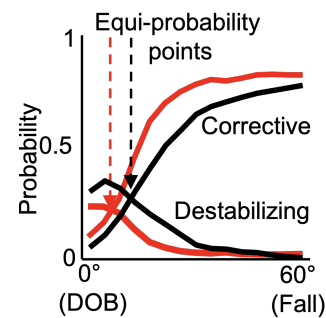


Figure 3: Complementary evolution of discrete destabilizing and corrective commands as a function of angular deviation away from the DOB and toward a fall boundary, seen in MARS (red) and VIP (black) tasks.

The same performance characteristics are exemplified in both MARS & VIP, as defined by subject matter experts in the neuroscience of balance dynamics [55]. In spaceflight, the pilot has cues about their motion but no orientation to gravity. The VIP 50% coherent reflects this by providing motion cues without cues

about configurational orientation. The VIP & MARS tasks possess similar underlying physical models of instability and in both tasks, the probability of destabilizing commands decreases and corrective commands increase as the pendulum position crosses the DOB and moves closer to a fall boundary (Fig. 3). This shows how the VIP task is demonstrably analogous to the MARS task, as the MARS task is demonstrably analogous to spaceflight conditions. Importantly, these distributions of actions come from human task performance and demonstrate a particular kind of human action policy based on sensorimotor perception, an example of structural coupling.

3.1 Data Collection

MARS. Wang et al. [58] released MARS human performance data from 34 healthy adults (18 female, 16 male). Each subject experienced two experimental sessions on consecutive days, each consisting of 20 100-sec. trials where they attempted to balance themselves with minimal oscillations while blindfolded. The data contains angular positions and velocities, and joystick deflections sampled at 50 Hz.

VIP. The VIP data consists of 31 healthy adults (22 female, 9 male). Subjects took part in 12 30-sec. trials in one session in the disorienting condition, with the same goal as the MARS task. Angular pos./vel. and joystick deflections were sampled at 200 Hz.

3.2 Performance Evaluation

An ideal performer in both of these tasks would be one that immediately rotates to the balance point and stays there with little to no motion. Calculable metrics from the collected data include number of crashes (excursions beyond $\pm 60^\circ$), proportion of destabilizing deflections (% *destab.*—see Fig. 2), mean/standard deviation of angular position θ , average magnitude of velocity ($\mu|Mag|_{vel}$), and root-mean-square (RMS) velocity. Lower metric values usually mean improvement, e.g., fewer crashes, more time spent near the DOB, less oscillation, slower motion, smaller deflections, etc.

MARS. Vimal et al. [55] used a Bayesian Gaussian Mixture method and the aforementioned features to cluster subjects into 3 statistically distinct groups that represent Proficient, Somewhat-Proficient, and Not-Proficient performance (hereafter *Good*, *Medium* and *Bad*). We use these clusters to split the MARS data into training subsets and to characterize digital twin performance in the task.

VIP. Plotting VIP subjects' RMS velocity vs. crash frequency revealed a positive linear relationship ($r = .73$). Based on the number of crashes in the 12th trial and crash reduction between Trials 1 and 12 (i.e., final performance and overall improvement), we assigned participants relative rankings and divided them in tertiles to mirror the Good/Medium/Bad MARS classification. VIP participants typically exhibited more crashes, destabilizing actions, and RMS velocity compared to MARS participants of the equivalent proficiency. These factors and differences in sample rate and environment allow us to test the transfer of AI assistants to novel digital twins.

4 MODEL TRAINING

Our goals in model training were both to train AI models capable of independently performing an IP balancing task parameterized with MARS physics and to create digital twins of humans that replicate different kinds of participant performance in the task. In some cases

these two categories overlapped, leading to a testable hypothesis: that a model that performs well at the task may also be able to assist a “pilot” (real or simulated) in performing the task better.

Fig. 4 shows an I/O schematic of all models. At time t_x , models take in a window containing the past *winSize* seconds of angular positions and velocities and predict the joystick deflection made at time $t_{x+future}$. DL models over human data additionally take in the past *winSize* joystick deflections made by the subject. If *winSize* = 0.0, the input consists of the values at t_x only. If *future* = 0.0, the next joystick deflection is predicted.

4.1 Reinforcement Learning Models

We trained reinforcement learning-based models that learn directly from exposure to environmental physics using a custom variation of Gymnasium's classic-control Pendulum environment [49]. This included 1) a problem space bounded at $\pm 60^\circ$ from the DOB, like the MARS/VIP task; 2) a random starting point for the inverted pendulum within the newly defined problem space; 3) a custom reward function³ given by Eq. 1, to encourage small continuous adjustments like those of Good MARS participants [52, 55].

$$r = \begin{cases} 0, & \text{if } -30^\circ \leq \theta \leq 30^\circ \\ -(\theta^2 + .1\omega^2 + .01d^2), & \text{if } \theta < -30^\circ \cup \theta > 30^\circ \end{cases} \quad (1)$$

The RL algorithms were directly exposed to environmental physics, unlike the DL models which received only implicit physics through the human performance data. The default SAC and DDPG implementations routinely converged to an optimal strategy that manifests as rotating immediately to the DOB and holding position there. We also trained and evaluated behavior cloning (BC) [42] and adversarial inverse RL (AIRL) [14] using Good MARS participant data, to teach the models strategies closer to what humans would execute, in terms of replicating behavior or uncovering implicit reward functions in the data.

RL models take the current angular position and velocity to predict the next joystick deflection (*winSize* = 0.0, *future* = 0.0, cf. Fig. 4). We used Stable-Baselines3's SAC and DDPG implementations [39], and trained them with the default MLP policy, BC, or AIRL. Gaussian distributed noise was added to the action space to encourage exploration as the IP is considered an under-actuated task [21]. We trained 5 RL models: 1) SAC & DDPG each with the standard policy; 2) SAC & DDPG each trained using BC; 3) AIRL implemented with a SAC-based generator model.

4.2 Deep Learning Models

To replicate human-like real-time performance of the MARS task, we trained Multilayer Perceptron (MLP), Vanilla Recurrent (RNN), Long-Short Term Memory (LSTM) [20], and Gated Recurrent Unit (GRU) [9] network architectures over actions made by humans in the actual MARS data. Architectures were trained using different window sizes (0.0s, just the current timestep—MLP models only; and 0.2s, 0.3s, and 0.5s). Training data was also split into Good, Medium, and Bad proficiencies and individual models were trained on data of a specific proficiency. An additional set of models was trained using a combination of 1) Good & Medium and 2) Good, Medium & Bad proficiency data, to see if models could learn strategies employed by

³Where θ is the angular position, ω is the angular velocity, and d is the joystick deflection.

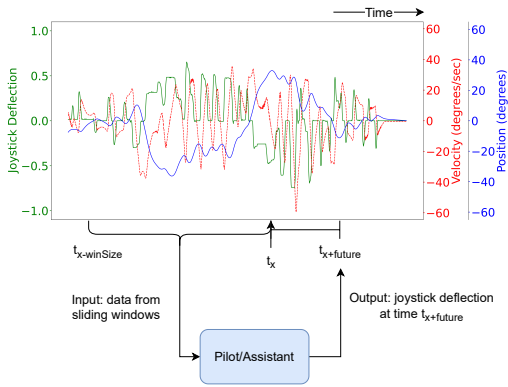


Figure 4: Model input and output structure. “Pilot/Assistant” stands in for any one of the trained prediction models.

certain proficiency groups in scenarios that were not experienced by the others. In total, we trained 40 individual DL models, all of which, even when they successfully avoid crashes in solo task performance, demonstrate suboptimal strategies with human-like oscillation and intermittent deflections. These behavioral differences show the differences in how the RL and DL models learn to situate themselves in (or “embody”) the problem space.

4.2.1 Selecting Representative Pilots. As the same performance characteristics are exemplified in both MARS & VIP, we identified models that most closely approximate performance categories from Vimal et al. [55], reflected in Table 1.

All DL models were made to perform the VIP task (3×30 s trials), with angular position, velocity, and joystick deflection recorded at each timestep. We extracted performance features shown in Table 1. Since the data distribution could not be assumed to be spherical, these features were used in k -means clustering ($k = 3$) to approximate the split into Good, Medium, or Bad groups. Following Vimal et al. [52], the cluster of models that displayed higher oscillations and greater average magnitude of deflections was considered Bad while the cluster that displayed smaller, more intermittent actions was considered Good (with the remainder considered Medium). We then took the models in each cluster that were trained over the equivalent data subset (e.g., models in the Good cluster trained over Good data, m.m.), and used the VIP performance characterization technique from Sec. 3.2 to identify which model best exemplified the characteristics of each proficiency group: **Good**—LSTM trained over Good data with a window size of 0.2s; **Medium**—GRU trained over Medium data with a window size of 0.3s predicting 0.1s into the future; **Bad**—MLP trained over Bad data with a window size of 0.5s. That each exemplar used a different architecture also suggests that the human subjects exhibited different strategies in performing the MARS task, to different effects.

Models trained over MARS data were found to also exemplify characteristics of different proficiencies when compared to participants in the VIP task, suggesting a level of generalizability between the two different environments. The selected architectures were then retrained using data from 3 VIP participants of each proficiency group to produce digital twins of VIP pilots. Table 1 shows the performance characteristics of each pilot exemplar model. All

other models that were trained over Good MARS data were reserved to act as candidate assistants, for a total of 21.⁴

5 EVALUATION

We perform two evaluations: 1) A high-throughput evaluation of pilot digital twins in co-performance of the VIP task with candidate assistants; 2) A human subject study of human co-performance of the VIP task with the best-performing assistants from the digital twin study.

The 4 major components of the evaluation pipeline include: 1) VIP 2) Crash predictor; 3) Assistant; 4) Pilot. The **VIP** component is *PyVIP*, a Python implementation for easy integration of ML models. The **Crash predictor** is a trained instance of the best crash prediction architecture reported in Wang et al. [58]: a stacked GRU trained over inputs like those in Sec. 4 that predicts the likelihood of a crash occurring. Due to the crash predictor’s high false positive rate, we added a *crash probability threshold* of 0.8 where only highly imminent danger would permit assistant suggestions.⁵ The assistant would provide suggestions when either 1) crash probability is greater than the threshold *and* angular distance from the DOB exceeds 12° , or 2) angular distance from the DOB exceeds 15° . The **Assistant** observes task performance and makes suggestions when certain conditions are met. The **Pilot** may be a digital twin or an actual human that controls the VIP. We used 6 digital twins—each of the architectures mentioned in Sec. 4.2.1, trained over both MARS and VIP data. Humans control the VIP with a joystick.

5.1 Digital Twins Study

In these experiments, the pilot has an 80% probability of accepting and executing an assistant suggestion instead of its own next action. If accepted, the pilot makes suggested deflections with a noise of $\mathcal{U}(-.05, .05)$ added to simulate human imprecision, after a $.4 + \mathcal{U}(-.05, .05)$ s delay to simulate reaction time.⁶

We ran each evaluation for 3 30-sec. trials. Data was sampled at 200 Hz with each sample comprising of the angular position and velocity of the VIP, joystick deflection, crash probability, pilot and assistant’s joystick deflections, which entity’s deflection was performed, and whether the deflection made was destabilizing. 468 individual digital twin trials were collected, or 3.9 hours of data.

5.1.1 Results. Table 2 shows performance differences between the digital twins when unaided and when aided by different assistants. Following the performance evaluation from Sec. 3.2 (where lower metric values signal improvement), SAC-AIRL is the overall strongest assistant for digital twins, decreasing crashes, % destabilizing deflections, and RMS velocity to a statistically significant level (all $p < 0.0001$ according to a paired two-tailed t -test).

RL models are generally better assistants than DL models over human data. Interestingly, *MLP-GMB-0* (MLP trained over all proficiencies with no window) decreased crashes as much as *SAC-AIRL*,

⁴The selected Good pilot architecture was retrained with different weight initialization to create a distinct instance of the model to act as an assistant.

⁵We follow Wang et al. [58]’s hypothesis that too many false positives could cause a human pilot to lose trust in the assistant, but also need to not admit too many false negatives. 80% represents a balance in these constraints.

⁶There is no prior work establishing the probability of human subjects following AI advice in *this* task, but work on other tasks report $\sim 80\%$ correctness of/willingness to rely on AI advice [12, 24, 56]. A 0.4s reaction time is fast for an average human [29], but well slower than a trained pilot [6].

Pilot	Crashes↓	% destabil.↓	$\mu \theta $ (°)↓	$\sigma(\theta)$ (°)↓	$\mu Mag _{vel}$ (°/s)↓	vel RMS↓
Good	7 / 17	15.9 / 54.0	16.6 / 20.3	21.5 / 23.7	53.3 / 53.0	70.6 / 77.7
Med.	9 / 40	21.4 / 63.3	20.1 / 28.8	18.9 / 29.7	68.3 / 122.7	93.9 / 152.3
Bad	27 / 23	36.7 / 52.8	21.4 / 19.4	25.9 / 14.3	114.1 / 57.4	135.9 / 92.2

Table 1: Performance statistics of pilot exemplar models (values are averaged over 3×30 sec. trials except # crashes, which is summed). Slashes separate models trained over MARS and VIP data. Columns from L–R: # crashes, % destabilizing actions, mean and SD distance from DOB, mean and SD angular velocity magnitude, and RMS velocity. Lower values are better (Sec. 3.2).

Assistant	Crashes↓	% destabil.↓	$\mu \theta $ (°)↓	$\sigma(\theta)$ (°)↓	$\mu Mag _{vel}$ (°/s)↓	vel RMS↓
SAC	0 / -5 / -25	2.2 / 4.9 / -18.3	0.4 / 1.2 / -3.6	-0.9 / -6.6 / -7.9	18.0 / -25.3 / -56.9	18.7 / -38.7 / -67.5
	-8 / -25 / -12	-31.8 / -38.3 / -28.5	-2.0 / -7.6 / -1.0	-1.9 / -6.7 / 7.9	27.0 / -37.7 / 7.5	24.4 / -42.2 / -6.9
SAC-AIRL	-2 / -7 / -17	1.9 / -3.8 / -14.4	1.9 / -0.1 / 5.4	0.9 / -7.7 / -0.3	-9.8 / -38.7 / -62.8	-11.2 / -53.5 / -71.2
	-15 / -33 / -12	-36.8 / -41.5 / -28.5	-0.5 / -6.1 / 0.9	-6.8 / -10.9 / 9.1	-22.5 / -66.6 / -8.7	-29.9 / -72.9 / -20.3
DDPG	1 / -2 / -21	2.3 / 3.7 / -17.3	2.4 / -2.4 / -2.2	2.1 / -3.1 / -3.4	25.3 / -11.0 / -43.1	24.6 / -21.2 / -51.0
	-11 / -23 / -12	-36.4 / -36.4 / -25.1	2.6 / -7.5 / -1.0	3.5 / -3.6 / 9.2	47.4 / -34.1 / 5.5	43.5 / -38.6 / -12.5
MLP-GMB-0	-2 / 4 / -11	-0.5 / 7.0 / -8.8	2.4 / 3.1 / 2.2	2.2 / 2.8 / -1.7	18.8 / 7.0 / -43.2	24.5 / 1.3 / -48.0
	-1 / -18 / -4	-21.6 / -29.7 / -21.5	2.1 / -6.8 / 4.0	0.8 / -2.5 / 12.6	31.4 / -15.5 / 12.2	38.7 / -18.6 / 3.8
LSTM-G-0.2	3 / 14 / -7	8.0 / 23.5 / -1.6	3.6 / -0.2 / 0.9	2.8 / 0.6 / 1.5	25.4 / 15.6 / -33.7	34.6 / 15.2 / -33.2
	4 / -13 / 1	-11.0 / -20.5 / -8.8	1.7 / -7.3 / 3.2	2.7 / -4.0 / 12.6	32.9 / -13.2 / 27.4	41.5 / -14.2 / 24.8

Table 2: Differences in performance with and without assistance (e.g., 0 means no change in that metric, lower values are better—Sec. 3.2) In each cell, top line refers to MARS pilot models and bottom to VIP pilot models. Slashes separate Good/Medium/Bad pilot models. Under Assistant, G/M/B denotes the proficiency of the assistant training data, decimals denote window size. Assistants shown achieved a significant reduction in at least one metric value. See appendix for results for all 26 assistants.

but for the Good MARS digital twin only. The Medium exemplar architecture performs significantly worse when trained over VIP data than over MARS data, and the Bad VIP pilot performs much more like the Medium MARS pilot, while Good pilot models are roughly consistent with each other across tasks, suggesting that many strategies lead to poor task performance and relatively few do well. This also speaks to SAC-AIRL’s ability to reduce crashes in both tasks for all proficiency levels. Other high-performing models also reduce crashes (DDPG: $p = 0.0002$, MLP-GMB-0: $p = 0.0033$) and destabilizing deflections (DDPG: $p = 0.0006$, MLP-GMB-0: $p = 0.0080$) for digital twins trained over both MARS and VIP data, demonstrating transfer between digital twins trained over the different task data.

5.2 Human Subject Study

Results from the high-throughput digital twins setting indicated that the 5 assistant models shown in Table 2 had a statistically significant effect on one or more metrics in co-performance with digital twins trained over both MARS and VIP data. These models—3 RL-based models and 2 models trained over human data—were included as candidate assistants in the human subject study. This gave us a robust but tractable sample of assistants to assess in co-performance of the VIP task with real human subjects.

We recruited 20 healthy adult subjects (6 female, 13 male, 1 non-binary). Each subject participated in 2 experimental sessions separated by approximately one week. In **Session 1**, subjects 1) attempted to balance a 50% coherent PyVIP RDK (3×30 s trials); 2) controlled the RDK with assistance from an AI model, rendered as left/right arrows indicating the direction of suggested deflection (3×30 s); 3) watched the same AI control the RDK while providing

directional suggestions via the joystick (3×30 s). Participants were randomly assigned one of the candidate assistant models during Session 1—subjects were grouped into fours and each group received assistance from a single type of architecture. Subjects were not told which type of model they were receiving assistance from.

Between sessions, each assistant model was fine-tuned using data from Task 3 in Session 1. Episodes (consisting of input window and predicted action) where the direction of agent-predicted deflection conflicted with the direction of human deflection were stored. These human-in-the-loop (HITL) disagreement samples were used to fine-tune the model: the actor networks of the SAC and DDPG were fine-tuned using behavior cloning over the new data, the SAC-AIRL model was updated using AIRL over the new data, and the deep learning models underwent standard fine-tuning.

In **Session 2**, subjects 1) undertook Task 1 as in Session 1 (solo RDK balancing— 3×30 s); 2) undertook AI-assisted balancing as in Session 1 Task 2 but with a different assistant model (3×30 s); 3) undertook AI-assisted balancing with the version of their Session 2 Task 2 assistant fine-tuned with data from Session 1 subjects who interacted with that model type (3×30 s). In Session 2 Task 2, subjects were assigned a non-fine-tuned model of a different architecture and in Session 2 Task 3 they were given an instance of that same model fine-tuned with HITL data from Session 1. Participants were not informed that the Session 2 Task 3 model was fine-tuned on Session 1 human data.

Finally, subjects took a survey, based on Muir [28], about their solo performance, how AI assistance changed their performance, and the level of trust they had in the assistant from each task.

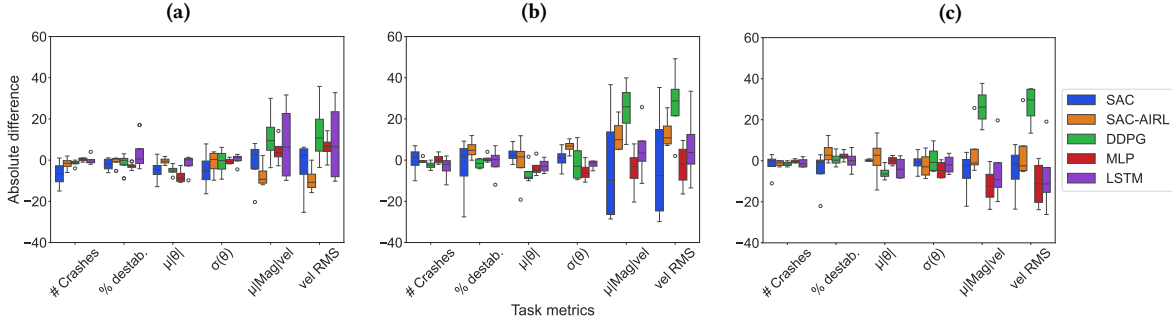


Figure 5: Absolute differences between baseline human performance metrics compared to AI-assistance in (a) Session 1 Task 2, (b) Session 2 Task 2 (different assistant model), and (c) Session 2 Task 3 (fine-tuned Session 2 Task 2 assistant).

5.2.1 Results. We first assessed whether subjects displayed any adaptation to the balancing task within or across sessions that might confound apparent performance improvements due to AI assistance. Following Vimal et al. [51], in which participants in the disorienting MARS task showed minimal learning across consecutive days, we take performance in Session 1 Task 1 vs. Session 2 Task 1 (which were separated by approximately 1 week) as a baseline “no learning” condition in which participants lost familiarity with the task. We then compare performance differences between Session 1 Task 1 vs. Session 2 Task 1 and between Session 2 Task 1 vs. Session 2 Task 2. If the performance differences between Session 2 Task 1 and Session 2 Task 2 are similarly non-significant compared to the performance differences between Session 1 Task 1 and Session 2 Task 1, this indicates that no significant adaptation to the task occurred between tasks within Session 2, and likewise is unlikely to have occurred between Session 2 Task 2 and Session 2 Task 3; therefore apparent differences in Session 2 Task 2 and Session 2 Task 3 performance are likely to be attributable to the nature of the AI assistance received.

We computed a score for each participant in each task of interest, given by Eq. 2,

$$s = \left(\frac{60 - \mu|\theta|}{60} \right) + \left(1 - \frac{C}{90} \right) + \left(1 - \frac{pD}{100} \right) + \frac{pA}{100} + \left(\frac{R}{\max_R} - \frac{C}{\max_C} \right), \quad (2)$$

where C is the count of crashes over the task ($3 \times 30s$ trials), pD is the percentage of deflections that were destabilizing, pA is the percentage of deflections that were anticipatory (see Fig. 2), R is the task-level count of recoveries from beyond 20° away from the DOB to within 20° of the DOB, and \max_R and \max_C are the maximum number of recoveries and crashes in the data, respectively.

Because these scores are not normally distributed, we ran a Wilcoxon Signed-Rank test between Session 1 Task 1 and Session 2 Task 1 scores, and between Session 2 Task 1 and Session 2 Task 2 scores. No statistically significant differences were found between either pairing, with similar p -values (.2627 between Session 1 Task 1 and Session 2 Task 1, and .3681 between Session 2 Task 1 and Session 2 Task 2). This indicates that there was no adaptation to the task significant enough to confound performance differences attributable to AI assistance.

Fig. 5 shows the absolute difference in performance metrics between human solo VIP performance and 3 versions of AI-assisted performance for each model type: using the original model weights

	SAC	SAC-AIRL	DDPG	MLP	LSTM
μ	69.17	58.83	132.17	28.33	28.41
σ	50.47	74.55	164.25	22.27	12.87

Table 3: Mean & SD of number of disagreement episodes logged during HITL study, by assistant model type.

in Session 1 (5a), the non-fine-tuned assistant from Session 2 (5b), and the Session 2 assistant fine-tuned on data from humans in Session 1 who interacted with the assistant of the same architecture (5c). We see that although, like in the digital twin studies, the RL assistants in Session 1 were better at reducing the absolute number of crashes than DL assistants, this distinction often vanished or reversed after the assistant models were fine-tuned on participant data and then re-evaluated in Session 2. In Session 2 Task 3, the DL models fine-tuned on Session 1 data were often now better on average at reducing metrics associated with velocity and oscillation such as RMS velocity, velocity magnitude, and standard deviation of position. This effect is absent in Session 2 Task 2, where subjects were assisted by a different type of model.

6 DISCUSSION

Like in the digital twins study, the SAC-AIRL assistant often helped the human subjects reduce crashes and oscillations. This is more pronounced in versions fine-tuned on HITL data, suggesting that models with a more human-like strategy contribute to this effect.

Assistance from the DDPG shows a strong tendency to increase RMS velocity and velocity magnitude values, and this is actually *more so* after HITL fine-tuning. This discrepancy is reflected in the number of disagreement episodes logged for each model type (Table 3). Human subjects registered a much higher mean number of disagreements with the DDPG model—and RL models more generally—than with the DL models. This further indicates that the DDPG and RL models behave in ways that may contradict human intuition and/or physical instinct. Through their data and training, DL models are embodying the problem space and performing the task in a more human-like way, including transitioning from destabilizing to corrective and anticipatory deflections at distances from the DOB that align with human behaviors (cf. Fig. 3).

Due to different human reaction times, it is not possible to know exactly when human subjects followed assistant suggestions, but we can calculate a heuristic estimate based on instances where a subject deflects in the direction suggested by the AI within a threshold of

Assistant	Perceived Performance Impact					
	Task 2			Task 3		
	+	~	-	+	~	-
Overall	50	15	35	55	30	15
SAC	25	25	50	50	25	25
SAC-AIRL	50	0	50	50	25	25
DDPG	75	0	25	50	25	25
MLP-GMB-0	75	25	0	75	25	0
LSTM-G-0.2	25	25	50	50	50	0

(a) +: improved (incl. slightly/significantly), ~: no change, -: decreased (incl. slightly/significantly).

Assistant	Reported Trust					
	Task 2			Task 3		
	+	~	-	+	~	-
Overall	25	35	40	30	45	25
SAC	25	0	75	25	50	25
SAC-AIRL	0	50	50	25	50	25
DDPG	0	75	25	25	25	50
MLP-GMB-0	50	25	25	25	50	25
LSTM-G-0.2	50	25	25	50	50	0

(b) +: high to complete, ~: moderate, -: low to none.

Table 4: Perceived performance impact of (4a), and reported level of trust in (4b) Session 2 Task 2 and Task 3 assistants (as %).

the AI making a suggestion. Using a threshold of 450ms, subjects followed AI suggestions approximately 44% of the time ($\sigma = 14\%$). This is significantly lower than the $\sim 80\%$ that can be seen in other domains (Sec. 5.1), suggesting particularities of the task need to be accounted for. The DDPG was the most followed assistant type at 53%, followed closely by SAC-AIRL (51%). The MLP was the least followed (32%). Interestingly, fine-tuned assistants were followed 7% less than non-fine-tuned assistants (4% less when comparing Session 2 Task 3 to only Session 2 Task 2), even though subjects rated these assistants as more trusted and preferable (see below).

Trust survey. Participants were asked to assess how the AI’s suggestions changed their performance, and to rate their trust in the AI. We report survey results after Session 2, where participants also expressed their preference for one of the two assistant models used: an assistant with no specific fine-tuning (*Task 2* assistant), or one fine-tuned using Session 1 human data (*Task 3* assistant).

Table 4a shows subjects’ perception of their assistant’s impact on their performance, overall and broken down by assistant type. Table 4b similarly shows the reported level of trust in each assistant. We see an overall trend toward better perceived impact on performance and more reported trust in the fine-tuned model when compared to the original, although interestingly some models to which participants ascribed a positive effect on performance (e.g., the MLP), were rated as less trusted after fine-tuning. At the assistant type level, these numbers should be taken in the context of small sample sizes ($N = 4$). When asked to pick a preference between the Task 2 and Task 3 assistants, 15% chose Task 2 and 70% chose Task 3 (15% no preference).

7 CONCLUSION AND FUTURE WORK

In this paper, we established a novel task of AI assistance in helping humans maintain balance in a disorienting condition. We first explored the space of possible AI assistance models using a high-throughput digital twins setting. The top performing models from this experiment were then used in a human-subject study to assess both performance impact and participants’ attitudes toward different assistants.

Given certain data and training methods, AIs that were capable of performing IP balancing alone were also able to assist real humans in reducing crashes and oscillation. SAC-AIRL, which learns rewards implicit in human performance data, appeared to be an effective disorientation countermeasure in both the digital twins

and human subject studies, by apparently embodying the problem space in a way that incorporates both physics and human signals. Although RL models on average make better assistants than DL models trained over human data, they do so by suggesting actions that often diverge significantly from the apparent model of the task captured in human actions. In the human subject studies: human subjects empirically perceived the RL models as performing the task incorrectly, and models that learned and embodied a human-like strategy through pretraining over human data, then were fine-tuned over more human data in the subject study, were able to significantly reduce factors related to oscillation and velocity.

Palmer et al. [32] illustrate trust dimensions in the use of autonomous or automated systems, such as *robustness* (handling perturbations/deviations appropriately), *benevolence* (supporting mission and operator), and *dynamism* (negotiating changes in environment). While our assistants’ suggested actions may be appropriately corrective (benevolent), respond to pilot-induced perturbations such as ignoring cues (robust), and transfer between the MARS and VIP tasks (dynamic), they need to also be understandable in terms of the pilot’s internal model of the situation to avoid corrections directly opposed to what the pilot expects (cf. the DDPG vs. MLP and LSTM in Fig. 5c). Our findings indicate that humans are in fact more receptive to assistance from an AI that demonstrates a more human-like, even if objectively suboptimal, balancing strategy.

Future study may investigate fine-tuning on data from a specific participant rather than an aggregate sample, to uncover person-specific patterns in task performance, or an investigation of modeling techniques that can account for the fact that human behavior is likely to change over time to account for assistance received from an AI agent, including one which is trained in real-time using live human feedback. Subsequent research may also involve transfer to more complicated conditions, like orientation in multiple roll planes or flight simulators, as well as investigating the transfer of AI assistance in the high-throughput VIP to the physical MARS.

There also remains the question of *how* to deliver an AI assistant’s cues to a human pilot. In this work we rendered visual indicators on the screen, but other modalities may include aurally rendered tones or vibrotactile cues (as in [54]) to indicate the direction and magnitude of the corrective action, or linguistic instructions. In a previous experiment Mannan and Krishnaswamy [26] presented evidence toward the utility of language understanding in task performance, and a multi-variable examination of intervention method and timing is another avenue of future study.

ACKNOWLEDGMENTS

HND was supported by a Brandeis Summer Internship Grant. VPV and PD were supported by Naval Aerospace Medical Research Unit Grant N0018923P0725. SM and NK would like to thank Sarath Sreedharan for his feedback on improving the framing and discussion of the paper. All authors would like to thank the many human subjects for contributing their valuable time toward the studies that led to this work, and the anonymous reviewers whose comments improved the final copy.

REFERENCES

- [1] Micah Allen, Nicolas Legrand, Camile Maria Costa Correa, and Francesca Fardo. 2020. Thinking through prior bodies: autonomic uncertainty and interoceptive self-inference. *Behavioral and Brain Sciences* 43 (2020).
- [2] Charles W Anderson. 1989. Learning to control an inverted pendulum using neural networks. *IEEE Control Systems Magazine* 9, 3 (1989), 31–37.
- [3] Shahar Arzy, Leila S Overney, Theodor Landis, and Olaf Blanke. 2006. Neural mechanisms of embodiment: Asomatognosia due to premotor cortex damage. *Archives of neurology* 63, 7 (2006), 1022–1025.
- [4] Shahar Arzy, Gregor Thut, Christine Mohr, Christoph M Michel, and Olaf Blanke. 2006. Neural basis of embodiment: distinct contributions of temporoparietal junction and extrastriate body area. *Journal of Neuroscience* 26, 31 (2006), 8074–8081.
- [5] Andrew G Barto, Richard S Sutton, and Charles W Anderson. 1983. Neuron-like adaptive elements that can solve difficult learning control problems. *IEEE transactions on systems, man, and cybernetics* 5 (1983), 834–846.
- [6] Bartosz Binias, Dariusz Myszor, Henryk Palus, and Krzysztof A Cyran. 2020. Prediction of pilot's reaction time based on EEG signals. *Frontiers in neuroinformatics* 14 (2020), 6.
- [7] Malcolm G Braithwaite, Simon J Durnford, John S Crowley, Norberto R Rosado, and John P Albano. 1998. Spatial disorientation in US Army rotary-wing operations. *Aviation, space, and environmental medicine* 69, 11 (1998), 1031–1037.
- [8] Peter Cariani. 1992. Some epistemological implications of devices which construct their own sensors and effectors. *Towards a practice of autonomous systems* (1992), 484–493.
- [9] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078* (2014).
- [10] Ronald Daiker, Kyle Ellis, Santosh Mathan, and W Redmond. 2018. Use of real-time, predictive human modeling for spatial disorientation detection and mitigation. In *Proceedings of the Modsim World 2018 Conference*.
- [11] Ezequiel Di Paolo and Evan Thompson. 2014. The enactive approach. *The Routledge handbook of embodied cognition* (2014), 68–78.
- [12] Huiying Du, Ge Zhu, and Jiali Zheng. 2021. Why travelers trust and accept self-driving cars: An empirical study. *Travel behaviour and society* 22 (2021), 1–9.
- [13] Razvan V Florian. 2007. Correct equations for the dynamics of the cart-pole system. *Center for Cognitive and Neural Studies (Coneural), Romania* (2007), 63.
- [14] Justin Fu, Katie Luo, and Sergey Levine. 2018. Learning Robust Rewards with Adversarial Inverse Reinforcement Learning. In *International Conference on Learning Representations*.
- [15] Randy Gibb, Bill Ercoline, and Lauren Scharff. 2011. Spatial disorientation: decades of pilot fatalities. *Aviation, space, and environmental medicine* 82, 7 (2011), 717–724.
- [16] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. 2018. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *International conference on machine learning*. PMLR, 1861–1870.
- [17] Peter A Hancock, Deborah R Billings, Kristin E Schaefer, Jessie YC Chen, Ewart J De Visser, and Raja Parasuraman. 2011. A meta-analysis of factors affecting trust in human-robot interaction. *Human factors* 53, 5 (2011), 517–527.
- [18] Randall B Hellman, Eric Chang, Justin Tanner, Stephen I Helms Tillery, and Veronica J Santos. 2015. A robot hand testbed designed for enhancing embodiment and functional neurorehabilitation of body schema in subjects with upper limb impairment or loss. *Frontiers in human neuroscience* 9 (2015), 26.
- [19] Daniel Hellmann, NN Giannakopoulos, R Blaser, L Eberhard, and HJ Schindler. 2011. The effect of various jaw motor tasks on body sway. *Journal of oral rehabilitation* 38, 10 (2011), 729–736.
- [20] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.
- [21] Jakob Hollenstein, Sayantan Auddy, Matteo Saveriano, Erwan Renaudo, and Justus Piater. 2022. Action Noise in Off-Policy Deep Reinforcement Learning: Impact on Exploration and Performance. *Transactions on Machine Learning Research* (2022).
- [22] Jin Joo Lee, Brad Knox, Jolie Baumann, Cynthia Breazeal, and David DeSteno. 2013. Computationally modeling interpersonal trust. *Frontiers in psychology* 4 (2013), 56004.
- [23] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. 2015. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971* (2015).
- [24] Siru Liu, Aileen P Wright, Barron L Patterson, Jonathan P Wanderer, Robert W Turer, Scott D Nelson, Allison B McCoy, Dean F Sittig, and Adam Wright. 2023. Using AI-generated suggestions from ChatGPT to optimize clinical decision support. *Journal of the American Medical Informatics Association* 30, 7 (2023), 1237–1245.
- [25] Terence J Lyons, William Ercoline, Kevin O'Toole, and Kevin Grayson. 2006. Aircraft and related factors in crashes involving spatial disorientation: 15 years of US Air Force data. *Aviation, space, and environmental medicine* 77, 7 (2006), 720–723.
- [26] Sheikh Mannan and Nikhil Krishnaswamy. 2022. Where Am I and Where Should I Go? Grounding Positional and Directional Labels in a Disoriented Human Balancing Task. In *Proceedings of the 2022 CLASP Conference on (Dis)embodiment*. Association for Computational Linguistics, Gothenburg, Sweden, 70–79. <https://aclanthology.org/2022.clasp-1.8>
- [27] Sheikh Mannan, Vivekanand Pandey Vimal, Paul DiZio, and Nikhil Krishnaswamy. 2024. Embodying Human-Like Modes of Balance Control Through Human-In-the-Loop Dyadic Learning. In *Proceedings of the AAAI Symposium Series*, Vol. 3. 565–569.
- [28] Bonnie M Muir. 1994. Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics* 37, 11 (1994), 1905–1922.
- [29] Charles Arthur Nagler and William Merle Nagler. 1973. Reaction time measurements. *Forensic science* 2 (1973), 261–274.
- [30] David G Newman. 2007. *An overview of spatial disorientation as a factor in aviation accidents and incidents*. Number B2007/0063. Australian Transport Safety Bureau Canberra City.
- [31] Stefanos Nikolaidis, Yu Xiang Zhu, David Hsu, and Siddhartha Srinivasa. 2017. Human-robot mutual adaptation in shared autonomy. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. 294–302.
- [32] Gari Palmer, Anne Selwyn, and Dan Zwillinger. 2016. *The "Trust V": Building and Measuring Trust in Autonomous Systems*. Springer US, Boston, MA, 55–77.
- [33] Alexander Sacha Panic, Heather Panic, Paul DiZio, and James R Lackner. 2017. Gravitational and somatosensory influences on control and perception of roll balance. *Aerospace medicine and human performance* 88, 11 (2017), 993–999.
- [34] Heather Panic, Alexander Sacha Panic, Paul DiZio, and James R Lackner. 2015. Direction of balance and perception of the upright are perceptually dissociable. *Journal of neurophysiology* 113, 10 (2015), 3600–3609.
- [35] Raja Parasuraman and Victor Riley. 1997. Humans and automation: Use, misuse, disuse, abuse. *Human factors* 39, 2 (1997), 230–253.
- [36] Rolf Pfeifer and Christian Scheier. 2001. *Understanding intelligence*. MIT Press.
- [37] James Pustejovsky and Nikhil Krishnaswamy. 2021. Embodied human computer interaction. *KI-Künstliche Intelligenz* 35, 3-4 (2021), 307–327.
- [38] Tom Quick and Kerstin Dautenhahn. 1999. Making embodiment measurable. *Proceedings of '4. Fachtagung der Gesellschaft für Kognitionswissenschaft' Bielefeld, Germany*. <http://supergoodtech.com/tomquick/phd/kogwis/webtext.html> (1999).
- [39] Antonin Raffin, Ashley Hill, Adam Gleave, Anssi Kanervisto, Maximilian Ernestus, and Noah Dormann. 2021. Stable-baselines3: Reliable reinforcement learning implementations. *The Journal of Machine Learning Research* 22, 1 (2021), 12348–12355.
- [40] Gary E Riccio, Eric J Martin, and Thomas A Stoffregen. 1992. The role of balance dynamics in the active perception of orientation. *Journal of Experimental Psychology: Human Perception and Performance* 18, 3 (1992), 624.
- [41] Paul Robinette, Wenchen Li, Robert Allen, Ayanna M Howard, and Alan R Wagner. 2016. Overtrust of robots in emergency evacuation scenarios. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 101–108.
- [42] Stéphane Ross and Drew Bagnell. 2010. Efficient reductions for imitation learning. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*. JMLR Workshop and Conference Proceedings, 661–668.
- [43] Garrett Sadler, Henri Battiste, Nhut Ho, Lauren Hoffmann, Walter Johnson, Robert Shively, Joseph Lyons, and David Smith. 2016. Effects of transparency on pilot trust and agreement in the autonomous constrained flight planner. In *2016 IEEE/AIAA 35th Digital Avionics Systems Conference (DASC)*. IEEE, 1–9.
- [44] Maha Salem, Gabriella Lakatos, Farshid Amirabdollahian, and Kerstin Dautenhahn. 2015. Would you trust a (faulty) robot? Effects of error, task type and personality on human-robot cooperation and trust. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*. 141–148.
- [45] Ben Shneiderman. 2020. Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human-Computer Interaction* 36, 6 (2020), 495–504.

- [46] Yuko Takada, Tetsuya Hisada, Naruo Kuwada, Masao Sakai, and Tomomitsu Akamatsu. 2009. Survey of severe spatial disorientation episodes in Japan Air Self-Defense Force fighter pilots showing increased severity in night flight. *Military medicine* 174, 6 (2009), 626–630.
- [47] Evan Thompson and Francisco J Varela. 2001. Radical embodiment: neural dynamics and consciousness. *Trends in cognitive sciences* 5, 10 (2001), 418–425.
- [48] Kathryn Tomczak, Adam Pelter, Corey Gutierrez, Thomas Stretch, Daniel Hilf, Bianca Donadio, Nathan L Tenhundfeld, Ewart J de Visser, and Chad C Tossell. 2019. Let Tesla park your Tesla: Driver trust in a semi-automated car. In *2019 Systems and Information Engineering Design Symposium (SIEDS)*. IEEE, 1–6.
- [49] Mark Towers, Jordan K. Terry, Ariel Kwiatkowski, John U. Balis, Gianluca de Cola, Tristan Deleu, Manuel Goulão, Andreas Kallinteris, Arjun KG, Markus Krimmel, Rodrigo Perez-Vicente, Andrea Pierré, Sander Schulhoff, Jun Jet Tai, Andrew Tan Jin Shen, and Omar G. Younis. 2023. Gymnasium. <https://doi.org/10.5281/zenodo.8127026>
- [50] Francisco J Varela, Evan Thompson, and Eleanor Rosch. 2017. *The embodied mind, revised edition: Cognitive science and human experience*. MIT press.
- [51] Vivekanand Pandey Vimal, Paul DiZio, and James R Lackner. 2017. Learning dynamic balancing in the roll plane with and without gravitational cues. *Experimental brain research* 235 (2017), 3495–3503.
- [52] Vivekanand Pandey Vimal, Paul DiZio, and James R Lackner. 2019. Learning and long-term retention of dynamic self-stabilization skills. *Experimental brain research* 237 (2019), 2775–2787.
- [53] Vivekanand Pandey Vimal, James R Lackner, and Paul DiZio. 2016. Learning dynamic control of body roll orientation. *Experimental brain research* 234 (2016), 483–492.
- [54] Vivekanand Pandey Vimal, Alexander Sacha Panic, James R Lackner, and Paul DiZio. 2023. Vibrotactile feedback as a countermeasure for spatial disorientation. *Frontiers in physiology* 14 (2023), 1249962.
- [55] Vivekanand Pandey Vimal, Han Zheng, Pengyu Hong, Lila N Fakhrazadeh, James R Lackner, and Paul DiZio. 2020. Characterizing individual differences in a dynamic stabilization task using machine learning. *Aerospace medicine and human performance* 91, 6 (2020), 479–488.
- [56] Kailas Vodrahalli, Roxana Daneshjou, Tobias Gerstenberg, and James Zou. 2022. Do humans trust advice more if it comes from ai? an analysis of human-ai interactions. In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*. 763–777.
- [57] Alan R Wagner, Jason Borenstein, and Ayanna Howard. 2018. Overtrust in the robotic age. *Commun. ACM* 61, 9 (2018), 22–24.
- [58] Yonglin Wang, Jie Tang, Vivekanand Pandey Vimal, James R Lackner, Paul DiZio, and Pengyu Hong. 2022. Crash prediction using deep learning in a disorienting spaceflight analog balancing task. *Frontiers in physiology* 13 (2022), 51.
- [59] Irina Zgonnikova, Arkady Zgonnikov, and Shigeru Kanemoto. 2016. Stick must fall: Using machine learning to predict human error in virtual balancing task. In *2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW)*. IEEE, 173–177.
- [60] Tom Ziemke. 2001. Are robots embodied. In *First international workshop on epigenetic robotics Modeling Cognitive Development in Robotic Systems*, Vol. 85. Kognitionsforskning, Lunds universitet, 701–746.