Dance2Hesitate: A Multi-Modal Dataset of Dancer-Taught Hesitancy for Understandable Robot Motion

Srikrishna Bangalore Raghu
University of Colorado Boulder
Boulder, Colorado, United States
srikrishna.bangaloreraghu@colorado.edu

Anna Soukhovei
University of Colorado Boulder
Boulder, Colorado, United States
anna.soukhovei@colorado.edu

Divya Sai Sindhuja Vankineni
University of Colorado Boulder
Boulder, Colorado, United States
divya.vankineni@colorado.edu

Alexandra Bacula
Pacific Lutheran University
Tacoma, Washington, United States
alexandra.bacula@plu.edu

Alessandro Roncone
University of Colorado Boulder
Boulder, Colorado, United States
alessandro.roncone@colorado.edu

Abstract

In human-robot collaboration, a robot's expression of *hesitancy* is a critical factor that shapes human coordination strategies, attention allocation, and safety-related judgments. However, designing *hesitant* robot motion that generalizes is challenging because the observer's inference is highly dependent on embodiment and context. To address these challenges, we introduce and open-source a multi-modal, dancer-generated dataset of *hesitant* motion where we focus on specific context-embodiment pairs (i.e., manipulator/human upper-limb approaching a Jenga Tower, and anthropomorphic whole body motion in free space). The dataset includes (i) kinesthetic teaching demonstrations on a Franka Emika Panda reaching from a fixed start configuration to a fixed target (a Jenga tower) with three graded hesitancy levels (slight, significant, extreme) and (ii) synchronized RGB-D motion capture of dancers performing the same reaching behavior using their upper limb across three hesitancy levels, plus full human body sequences for extreme hesitancy. We further provide documentation to enable reproducible benchmarking across robot and human modalities. Across all dancers, we obtained 70 unique whole- body trajectories, 84 upper limb trajectories spanning over the three hesitancy levels, and 66 kinesthetic teaching trajectories spanning over the three hesitancy levels. The dataset can be accessed here: https://brsrikrishna.github.io/Dance2Hesitate/.

CCS Concepts

• **Human-centered computing → Empirical studies in HCI**; **Empirical studies in interaction design**.

Keywords

Human-Robot Interaction, Expressive Motion, Hesitancy

Permission to make digital or hard copies of all or part 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 components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
D-TUR '26, Edinburgh, UK
© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.

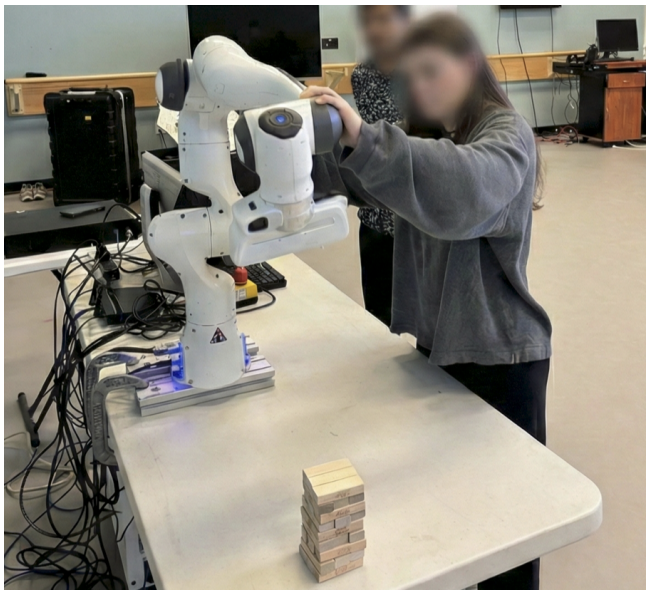


Figure 1: A dancer performing kinesthetic teaching on the Franka Emika Panda manipulator to generate a *hesitant* trajectory towards the Jenga tower

.

ACM Reference Format:
Srikrishna Bangalore Raghu, Anna Soukhovei, Divya Sai Sindhuja Vankineni, Alexandra Bacula, and Alessandro Roncone. 2026. Dance2Hesitate: A Multi-Modal Dataset of Dancer-Taught Hesitancy for Understandable Robot Motion. In *Proceedings of HRI '26 Designing Transparent and Understandable Robots (D-TUR '26).* ACM, New York, NY, USA, 6 pages.

1 Introduction

As robots become more widespread in our everyday lives, ensuring successful collaboration with humans has become exceedingly important [18]. Coherent human–robot collaboration requires the robot to transparently communicate its intention and sentiment, while the human attempts to form an accurate mental model of the robot (i.e., an understanding of its capabilities, intentions, and limitations) [21]. While speech is a dominant channel for conveying information, humans regularly complement verbal communication

with nonverbal cues [11]. Employing robots that exhibit similar nonverbal signals can enhance coordination, cohesion, and shared understanding within human–robot teams [1, 2, 7]. Nonverbal communication is inherently multimodal, encompassing gaze, haptics, facial expressions, and other bodily cues [11]. In robotics, a widely applicable modality is motion, where spatiotemporal variations in the robot joints can serve as readable signals of its internal state [5, 9].

Expressive robot motion has been studied extensively in the context of affective expression, where robots convey emotions such as happiness, sadness, or fear through stylistic variations in movement [12, 20]. In this work, we focus on *hesitancy*, which falls under the umbrella of *functional expressivity* (i.e., the robot's ability to communicate its functional capabilities and limitations) [16]. In pursuit of exploring this direction of research, we frame "uncertainty" as the functional capability of the robot to express the unreliability and riskiness of successfully completing its task, and *hesitancy* as the functional expression, i.e., what is used to express uncertainty. *Hesitancy* is particularly consequential in human–robot collaboration since it directly shapes how humans coordinate with the robot, readjust their own attention, and make safety-critical judgments [4, 10]. When a robot is uncertain about its success, conveying it through motion can help humans calibrate trust, anticipate potential failure, and intervene appropriately, thereby reducing unsafe surprises and improving teamwork [15, 22].

However, achieving this is extremely challenging due to two major reasons. Due to regular interactions between humans in everyday life, we tend to relate to a humanoid robot's form factor and easily interpret its nonverbal cues [6], which does not translate to most robots with non-anthropomorphic morphologies (e.g., manipulators, mobile bases, or UAVs). As a result, the same underlying internal state can produce motion cues that are ambiguous or even misinterpreted when expressed through different embodiments, since each platform offers different degrees of freedom and kinematic constraints [7, 9]. Second, expressive motion is inherently context-specific: how a behavior such as hesitancy is perceived depends strongly on the task, environment, and interaction dynamics [10, 15]. For example, a pause or a slow approach may signal caution, uncertainty, or deference in one setting, but may be interpreted as malfunction, inefficiency, or obstruction in another [9, 15].

Together, these embodiment and context dependencies make it difficult to design expressive behaviors that generalize beyond a single robot or scenario, motivating controlled experimental tasks and multi-modal datasets that isolate expressive intent while holding the functional goal constant. In this work, we take a deliberate first step toward addressing this challenge by studying hesitancy in two carefully chosen context-embodiment pairs that emphasize both task-grounded and embodiment-grounded expression: (i) a fixed-goal manipulation setting in which a Franka Emika Panda approaches a Jenga tower from a canonical start pose, and (ii) anthropomorphic whole-body motion in free space that captures broader posture and weight-shift cues. We position this release as a controlled first step that establishes reproducible benchmarks and a cross-modal structure that future work can extend to additional robot platforms and scenarios. To achieve this, we recruited dancers to ensure reliable, repeatable, and intentionally modulated demonstrations of graded hesitancy under a fixed task context. Dancers' movement expertise helps reduce unwanted variability and enables clearer separation between hesitancy levels, which is valuable for establishing a benchmark and extracting kinematic signatures of hesitancy, resulting in the following contributions:

(1) **A multi-modal, dancer-generated dataset of hesitant motion** that captures functional expressivity across three modalities of data, consisting of kinesthetic teaching demonstrations on a Franka Emika Panda reaching from a fixed start configuration to a fixed target (a Jenga tower) with three graded hesitancy levels (slight, significant, extreme), alongside RGB-D motion capture of dancers performing the same reaching behavior with upper-limb keypoints at all three levels and full human-body sequences for extreme hesitancy.
(2) **A unified data release with standardized formats and tooling**, including MP4 videos, ROSBags, CSV/NPZ robot logs, synchronized RGB videos, and 2D/3D keypoints with confidence scores, enabling reproducible benchmarking across robot and human modalities.

2 Related Work

2.1 Robots expressing affect via motion

A long-standing theme in Human-Robot interaction (HRI) is that *how* a robot moves can be as communicative as *what* it does [3, 5, 7]. Movement timing, acceleration, spatial path shape, etc., contribute to shaping human interpretations of a robot's attitude, competence, and social stance [14, 24]. Empirical studies have shown that robot motion can measurably impact humans' affective and subjective responses. For example, Kulić and Croft study how different properties of articulated robot motion relate to physiological and subjective responses, reinforcing that motion parameters can modulate affective reactions [14]. This motivated computational representations of "movement quality," often borrowing from human movement science and performing arts [13]. A prominent example is the work proposed by Knight and Simmons, where they leverage Laban-Effort inspired motion features for mobile robots and show that such features can support classifying distinct "manners" of moving from trajectories collected with expert input, pointing toward principled feature spaces for expressivity rather than ad-hoc parameter tweaking [12]. More recent work increasingly aims to *learn* controllable affective style rather than handcraft it. One line of work treats affect as a trajectory optimization objective by learning style- or emotion-specific cost functions from user feedback and then optimizing these costs alongside task constraints [23]. A key limitation of learning separate costs for each affect label is inefficiency and poor generalization across emotions. Addressing this, Sripathy et al. leverage the Valence–Arousal–Dominance (VAD) affect space to learn a single mapping from trajectories to VAD, enabling a unified cost to generate motions throughout a continuous affect space [19]. In parallel, representation-learning approaches propose editing affective motion in a latent space. Suguitan and Hoffman introduce a classifying variational autoencoder that can modify a base affective movement toward a desired affect class, enabling post-hoc affect modulation without re-planning from scratch [20]. Taken together,

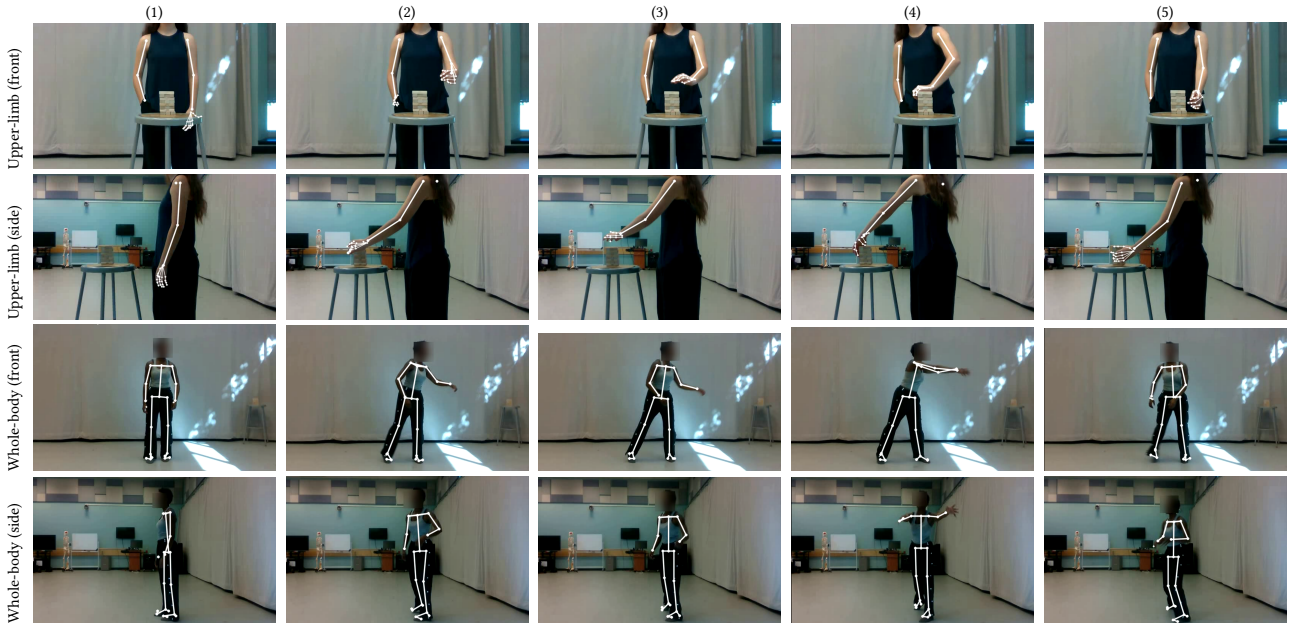


Figure 2: Five sequential frames (1–5) from the postprocessed overlays for each view: upper-limb and whole-body motions captured from front and side cameras.

these works indicate that affective expression via motion is moving from handcrafted heuristics to data-driven, continuous, and controllable style representations that can be composed with task objectives.

2.2 Robots expressing hesitancy

Hesitancy is a particularly important subcase of expressive robot behavior because it often functions as an interactional signal. People usually express hesitation to negotiate right-of-way, communicate uncertainty, soften commitments, or coordinate in shared spaces [8]. Moon et al. [15] characterize "hesitation gestures" that occur during human-human reaching conflicts and translate them into robot motion responses for human-robot resource conflicts. They derive a characteristic Acceleration-based Hesitation Profile (AHP) from recorded human hesitation motions and evaluate whether such hesitation-like responses are recognizable and distinguishable from abrupt stopping. This formalization frames *hesitancy* as a structured, legible social cue grounded in human motion patterns. Subsequent work extends hesitancy beyond reaching conflicts into a mechanism for communicating a robot's internal uncertainty or confidence via the manner of action execution. Hough and Schlangen [10] explicitly investigate how a simple robot can ground and convey different degrees of internal uncertainty through non-verbal task actions (i.e., how it acts, not just the action outcome), and show that observers can infer the robot's internal uncertainty more reliably when the system is designed to express it in action execution. More broadly, recent work on uncertainty communication in collaborative decision-making explores embodied behaviors as one modality among others (e.g., GUIs vs. embodied cues) and reports that such uncertainty visualizations can meaningfully affect human decision-making and perceived transparency in high-stakes tasks [17].

Taken together, prior work suggests that hesitancy is both a readable social signal and a useful mechanism for communicating uncertainty through motion. However, it also highlights a practical challenge, that expressive cues are tightly coupled to embodiment and context. This motivates a dataset that holds the functional goal constant while allowing expressive intent to vary, enabling systematic analysis of spatiotemporal signatures of hesitancy and providing training data for robot-side models.

3 Methods

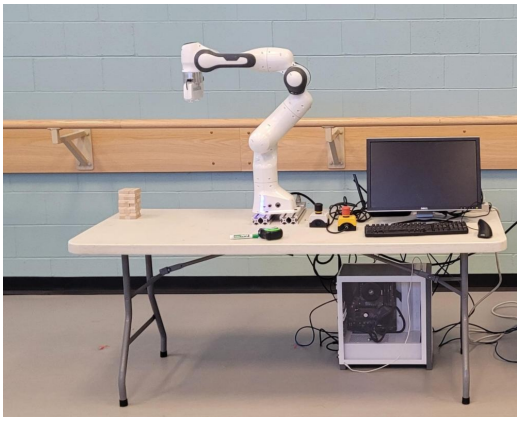
We collected a multi-modal dataset of *hesitant* motion demonstrations from recruited dancers, intending to capture how hesitancy is expressed both on a robotic manipulator and through human motion. Concretely, the dataset spans:

(1) **Kinesthetic teaching** of a Franka Emika Panda manipulator reaching from a fixed start configuration to a fixed target (a Jenga tower), and
(2) **RGB-D motion capture** of dancer reaching motions using either the upper limb or the full body.

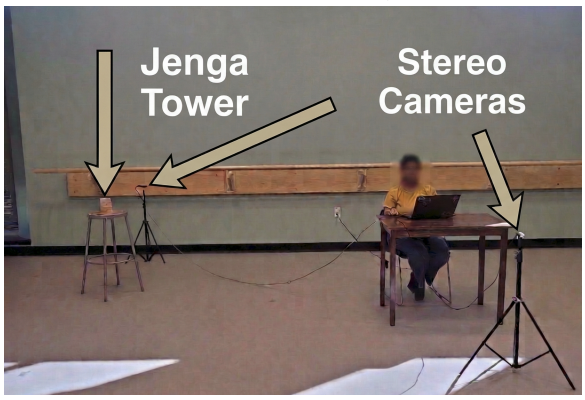
Across modalities, hesitancy was recorded at three levels: *slight*, *significant*, and *extreme*. For the full-body condition, dancers performed a single trajectory expressing *extreme* hesitancy to emphasize whole-body hesitation cues.

3.1 Robot Kinesthetic Teaching (Franka Emika Panda)

We used the Franka Emika Panda robotic manipulator in a kinesthetic teaching setting (gravity-compensation / compliant guidance), where dancers physically guided the robot from its default



(a) The kinesthetic teaching setup.



(b) The motion capture setup.

Figure 3: Photos of the data collection setup featuring the kinesthetic teaching station (a) and the motion capture station (b).

configuration towards the Jenga tower, as seen in Fig.1 and Fig.3a. To isolate expressive differences while holding the task goal constant, the robot always began from the same start configuration, and the Jenga tower remained fixed in the workspace. Demonstrations were labeled with three hesitancy levels, and the resulting kinesthetic teaching dataset is balanced, with approximately one third of trajectories each corresponding to *slight*, *significant*, and *extreme* hesitancy. The dancers were allowed to watch the other demonstrations for approximately 10 minutes to familiarize themselves with the robot's motion. During each demonstration, we logged robot measurements from the Franka ROS topic providing full robot state (`franka_states`). For ease of access and downstream learning, we additionally export the following time-series variables to CSV:

- Joint angles: $\mathbf{q}(t)$
- Joint velocities: $\dot{\mathbf{q}}(t)$
- Joint efforts/torques: $\boldsymbol{\tau}_j(t)$
- End-effector position: $\mathbf{p}_{ee}(t)$
- End-effector rotation/orientation: $\mathbf{R}_{ee}(t)$

Kinesthetic teaching data are released in three complementary formats: (i) ROS bags, (ii) five CSV files (one per variable group above), and (iii) NPZ files containing joint-angle-only trajectories $\mathbf{q}(t)$. In each CSV, *columns correspond to participants* and *rows correspond to timesteps*, enabling straightforward batching across demonstrations.

3.2 Human Motion Capture with RGB-D Cameras

To capture how dancers express hesitancy through their own movement, we recorded RGB-D motion data using two RealSense cameras (Fig.3b), collecting synchronized color and depth streams and extracting 2D and 3D keypoints. We consider two human conditions: upper-limb reaching toward the Jenga tower (recorded at all three hesitancy levels) and a full-body sequence emphasizing *extreme* hesitancy.

3.2.1 2D keypoints (OpenPose). We extract 2D keypoints using the OpenPose BODY_25 model. Each keypoint provides pixel coordinates (x, y) and a confidence score $c \in [0, 1]$ per frame. To ensure consistent downstream processing, we apply a minimum confidence threshold `MIN_CONF` = 0.30 and treat lower-confidence detections as missing for that frame. For visualization, we provide two post-processed overlay modes:

(1) **Body overlay (no face):** arms, torso, legs, and feet are rendered, while face-related keypoints are not drawn.
(2) **Arm-only overlay:** left/right shoulder–elbow–wrist chains are rendered; when available, we also record hand keypoints (21 per hand) for fine-grained manipulation cues.

In addition to MP4 videos (Fig.2), we store raw 2D keypoints frame-by-frame as CSV files with schema `(frame, joint_index, x, y, confidence)`, and we additionally provide equivalent streams in ROS bags for robotics toolchain compatibility.

3.2.2 3D keypoints (Depth). To obtain 3D keypoints, we align each depth frame to its corresponding color image so that the depth value at pixel (x, y) matches the 2D keypoint location. Depth is recorded in millimeters. Using the camera intrinsics (focal lengths and principal point), we back-project 2D keypoints into 3D camera coordinates:

$$\mathbf{X} = (X, Y, Z) = \pi^{-1}(x, y, Z; \mathbf{K}), \tag{1}$$

where $\mathbf{K}$ denotes intrinsics and Z is the aligned depth at (x, y). By default, 3D points are expressed in the camera coordinate frame (origin at the camera). When a calibrated two-camera setup is available, 3D positions may alternatively be triangulated from both views. We store the RGB-D outputs in three forms: (i) timestamped 16-bit PNG depth images (lossless), (ii) 3D keypoints as CSV with schema `(frame, joint_index, X, Y, Z, confidence)` in millimeters, and (iii) optional JSON and ROS bag topics for 3D keypoints. Table 1 summarizes the modalities and their stored artifacts.

For RGB-D motion capture, we target a sampling rate of 30 fps with time-aligned color, depth, and keypoints. Each depth frame and keypoint record includes timestamps to facilitate synchronization during downstream analysis. All downstream measures are computed in a confidence-aware manner: for any per-frame quantity requiring a set of keypoints (e.g., segment lengths or reach distances), we only compute values when all required joints have confidence $\geq$ `MIN_CONF`. Short confidence drops are left blank (or

Table 1: Summary of recorded modalities and file formats.

Modality	Source	Content	Units	Stored artifacts
Robot kinesthetic teaching	Franka Panda (`franka_states`)	$\mathbf{q}, \dot{\mathbf{q}}, \boldsymbol{\tau}_j, \mathbf{p}_{ee}, \mathbf{R}_{ee}$	SI	ROS bags; 5 CSVs (one per variable group); NPZ ($\mathbf{q}$ only)
Upper-limb reaching (human)	RealSense + OpenPose	2D keypoints + confidence; raw color video	px	Videos; `keypoints_2d.csv`; ROS bags
Upper-limb reaching (human)	RealSense depth	3D keypoints + confidence; depth frames	mm	`keypoints_3d.csv`; timestamped 16-bit depth PNG; ROS/JSON
Full-body extreme hesitancy	RealSense + OpenPose	BODY_25 2D/3D keypoints + confidence; raw video	px/mm	Videos; `keypoints_2d.csv`, `keypoints_3d.csv`; depth PNG

lightly smoothed for visualization), and we do not hallucinate or infill missing keypoints in the released raw files.

3.3 Participants and Protocol

We recruited 14 dancers (2 M and 12 F, 10 dancers with 10+ years of dance training, and 4 dancers with 1-10 years of dance training). The specific genre of the dancers' expertise spanned Hip-Hop, Ballet, Contemporary, and Jazz. All 14 dancers participated in generating the motion-capture video dataset, and 11 out of 14 dancers participated in creating the kinesthetic teaching dataset. Across all dancers, we obtained 70 unique whole-body trajectories, 84 upper limb trajectories spanning over the three hesitancy levels, and 66 kinesthetic teaching trajectories spanning over the three hesitancy levels. For motion-capture visualization and analysis, we focus on body/arm keypoints and do not render facial keypoints in overlays.

4 Conclusion and Future Work

We presented an open-source, multi-modal dataset of *hesitant* motions collected by recruiting dancers that span kinesthetic teaching demonstrations on a Franka Emika Panda and RGB-D motion capture of human reaching behaviors (upper-limb across three hesitancy levels and full-body for extreme hesitancy). By fixing the functional objective (reaching from a canonical start toward a fixed Jenga tower) and varying only the intended hesitancy, the robot and upper limb dataset isolates hesitancy as a form of *functional expressivity* expressed through spatiotemporal motion cues. Looking forward, this dataset enables several research directions.

On the robotics side, kinesthetic teaching trajectories can support learning hesitancy-aware representations and controllers, including supervised recognition of hesitancy from robot state, as well as conditional generative modeling (e.g., diffusion-based synthesis) to produce a continuous "hesitancy slider" that interpolates between hesitancy levels while satisfying task constraints. More broadly, hesitant motion planning can be framed as optimizing feasibility and safety while regularizing trajectories toward a learned hesitancy prior. On the human side, the RGB-D recordings enable benchmark tasks for hesitancy recognition from 2D keypoints, 3D keypoints, and RGB, as well as analyses of which kinematic features most strongly correlate with perceived hesitancy. Finally, the dataset provides a foundation for cross-embodiment modeling. Future work can investigate shared latent embeddings of hesitancy that align human and robot trajectories without explicit retargeting, and can extend whole-body hesitant behaviors to humanoids. As an additional direction for future work, it would be valuable to study how experts from different communities express hesitancy under the same task constraints. In particular, comparing demonstrations from dancers with those from puppeteers, actors, and other performance practitioners could reveal which kinematic cues are consistent across training backgrounds. Such comparisons may help disentangle "universal" hesitancy signatures (e.g., pauses, delayed commitment) from domain-specific embellishments, and inform data collection strategies that better balance expressivity, realism, and generalization to everyday users. We hope this release accelerates research on motion as a mode of communication by providing reproducible resources for studying, modeling, and generating hesitancy as a task-relevant nonverbal signal.

5 Acknowledgments

We would like to thank Prof. Madeline Harvey at Colorado State University for her support and time. We thank the dancers at Colorado State University for their time and creativity, leading to the creation of this dataset.

References

[1] Henny Admoni, Thomas Weng, Bradley Hayes, and Brian Scassellati. 2016. Robot Nonverbal Behavior Improves Task Performance in Difficult Collaborations. In *Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 51–58. doi:10.1109/HRI.2016.7451733

[2] Cynthia Breazeal, Cory D. Kidd, Andrea Lockerd Thomaz, Guy Hoffman, and Matt Berlin. 2005. Effects of Nonverbal Communication on Efficiency and Robustness in Human-Robot Teamwork. In *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 708–713. doi:10.1109/IROS.2005.1545011

[3] Anca Dragan and Siddhartha Srinivasa. 2013. Generating legible motion. (2013).

[4] Anca D Dragan, Shira Bauman, Jodi Forlizzi, and Siddhartha S Srinivasa. 2015. Effects of robot motion on human-robot collaboration. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*. 51–58.

[5] Anca D. Dragan, Kenton C. Lee, and Siddhartha S. Srinivasa. 2013. Legibility and Predictability of Robot Motion. In *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI '13)*. 301–308. doi:10.1109/HRI.2013.6483603

[6] Julia Fink. 2012. Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction. In *Social Robotics (ICSR 2012) (Lecture Notes in Computer Science, Vol. 7621)*. Springer, 199–208.

[7] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. A Survey of Socially Interactive Robots. *Robotics and Autonomous Systems* 42, 3–4 (2003), 143–166. doi:10.1016/S0921-8890(02)00372-X

[8] Justin W Hart, Brian Gleeson, Matthew Pan, AJung Moon, Karon MacLean, and Elizabeth Croft. 2014. Gesture, gaze, touch, and hesitation: Timing cues for

collaborative work. In *HRI Workshop on Timing in Human-Robot Interaction, Bielefeld, Germany*, Vol. 21.

[9] Guy Hoffman and Wendy Ju. 2014. Designing Robots with Movement in Mind. *Journal of Human-Robot Interaction* 3, 1 (2014), 89–122. doi:10.5898/JHRI.3.1. Hoffman

[10] Julian Hough and David Schlangen. 2017. It's Not What You Do, It's How You Do It: Grounding Uncertainty for a Simple Robot. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction (HRI '17)*. 274–282. doi:10.1145/2909824.3020214

[11] Mark L. Knapp, Judith A. Hall, and Terrence G. Horgan. 2021. *Nonverbal Communication in Human Interaction* (9 ed.). Kendall Hunt Publishing Company.

[12] Heather Knight and Reid Simmons. 2014. Expressive motion with x, y and theta: Laban Effort Features for mobile robots. In *2014 IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 267–273. doi:10.1109/ROMAN.2014.6926264

[13] Heather Knight and Reid Simmons. 2016. Laban head-motions convey robot state: A call for robot body language. In *2016 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2881–2888.

[14] Dana Kulić and Elizabeth A. Croft. 2007. Physiological and subjective responses to articulated robot motion. *Robotica* 25, 1 (2007), 13–27. doi:10.1017/S0263574706002955

[15] Ajung Moon, Chris A. C. Parker, Elizabeth A. Croft, and H. F. Machiel Van der Loos. 2013. Design and Impact of Hesitation Gestures during Human-Robot Resource Conflicts. *Journal of Human-Robot Interaction* 2, 3 (2013), 18–40. doi:10.5898/JHRI.2.3.Moon

[16] Srikrishna Bangalore Raghu, Clare Lohrmann, Akshay Bakshi, Jennifer Kim, Jose Caraveo Herrera, Bradley Hayes, and Alessandro Roncone. 2025. Employing Laban Shape for Generating Emotionally and Functionally Expressive Trajectories in Robotic Manipulators. In *2025 34th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1201–1208.

[17] Sarah Schömbs, Saumya Pareek, Jorge Goncalves, and Wafa Johal. 2024. Robot-assisted decision-making: Unveiling the role of uncertainty visualisation and embodiment. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–16.

[18] Francesco Semeraro, Alexander Griffiths, and Angelo Cangelosi. 2023. Human–robot collaboration and machine learning: A systematic review of recent research. *Robotics and computer-integrated manufacturing* 79 (2023), 102432.

[19] Arjun Sripathy, Andreea Bobu, Zhongyu Li, Koushil Sreenath, Daniel S. Brown, and Anca D. Dragan. 2022. Teaching Robots to Span the Space of Functional Expressive Motion. arXiv:2203.02091 [cs.RO] doi:10.48550/arXiv.2203.02091

[20] Mark Suguitan and Guy Hoffman. 2020. MoveAE: Modifying Affective Robot Movements Using Classifying Variational Autoencoders. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI '20)*. doi:10.1145/3319502.3374807

[21] Aaquib Tabrez, Matthew B. Luebbers, and Bradley Hayes. 2020. A Survey of Mental Modeling Techniques in Human–Robot Teaming. *Current Robotics Reports* 1, 4 (2020), 259–267. doi:10.1007/s43154-020-00019-0

[22] Ning Wang, David V Pynadath, and Susan G Hill. 2016. Trust calibration within a human-robot team: Comparing automatically generated explanations. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 109–116.

[23] Allan Zhou and Anca D. Dragan. 2018. Cost Functions for Robot Motion Style. arXiv:1809.00092 [cs.RO] doi:10.48550/arXiv.1809.00092

[24] 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*. 22–31.