

Discreet vs. Expressive Gestures for Controlling Physical Moving Objects in Public

Daniel Moroz

daniellebed@campus.technion.ac.il

Abstract

Gesture-based interaction offers an intuitive way to control robots and drones, but it introduces a social friction point: operators must perform visible, embodied actions that can feel awkward or attention-drawing in public. Prior work on mobile and head-worn devices shows that social context and how noticeable a gesture appears influence willingness to perform it in public (e.g., [3, 8, 10, 12]). In contrast, human-drone interaction and UAV gesture-metaphor work often yields larger, spatial, and metaphor-driven movements for commanding a visible flying robot (e.g., [1, 7]). We examine this tension by studying how gesture style (discreet vs. expressive) and social context (private vs. public with bystanders) shape operator preference, perceived control, and willingness to use gesture control for a clearly visible 2DOF pan-tilt robot whose on-board LEDs indicate target direction and progress. We attempted to collect brief bystander impressions via in-situ intercept prompts during public blocks; however, participation was low and incomplete, so we focus our analysis on operator self-report and system telemetry (time, completion, and gesture counts) captured by the study system.

All code and materials are available at: <https://github.com/Kiiwii10/discreet-vs-expressive-gestures>

1 Introduction

Gesture control is a promising modality for spatial computing, drones, and mobile robots, but it comes with a distinctive social risk: control is enacted through visible, embodied movements that can appear unusual, attention-seeking, or difficult to interpret. A consistent theme in mobile and wearable interaction is that users avoid gestures that feel socially awkward or draw unwanted attention, especially in the presence of an audience [3, 5, 10, 12]. For on-body and wearable interfaces, factors such as body location and the surrounding social setting can shape comfort and perceived acceptability [4, 8].

Public robot control adds an important twist: gestures are not only input to a system, but also a social signal to others. When the controlled device is a visible, moving robot—such as a drone—elicitation and metaphor-based work often yields larger, spatial, or interpersonal gestures for commanding it [1, 7]. Separately, work on flying robots shows that robot motion can communicate intent to observers and shape perceived safety [13]. This sets up a core design tension for public-facing deployment: do operators prefer discreet gestures to minimize embarrassment and social exposure, or do they prefer more expressive gestures because they feel more effective and make control more legible to observers?

To address this question, we compare two gesture styles—discreet and expressive—that implement the same command set, and we evaluate them under two social contexts: a private condition and a public condition with bystanders present. We measure operator preference, willingness to use, perceived control, and task telemetry. We also attempted brief intercept probes with nearby bystanders during public blocks, but participation was too limited and inconsistent for quantitative analysis; we report this as a practical constraint on in-the-wild bystander measurement.

Contributions.

- (1) A controlled comparison of discreet vs. expressive gesture sets implementing the same commands, evaluated under private vs. public-with-bystanders contexts.
- (2) Evidence characterizing trade-offs between social comfort, perceived control, and task efficiency across private vs. public contexts, grounded in operator self-report and system telemetry; we additionally report practical constraints encountered when probing bystanders in situ.
- (3) Design recommendations for gesture-based control of visible robots in public settings, grounded in operator outcomes and limited bystander observations.

2 Related Work

2.1 Social acceptability of gestures and wearables

Work on mobile gestures and around-device interaction shows that gesture form factor and social context shape willingness to perform gestures in public [9, 10, 12]. For head-worn displays and smart glasses, researchers have explored techniques such as hand-to-face input and willing-to-use-in-public gesture vocabularies [3, 11]. Survey and measurement work further proposes constructs and instruments for assessing social acceptability (e.g., the WEAR scale) [4, 5]. This literature motivates the expectation that discreet, less conspicuous gestures may better support public use, but it rarely considers cases where gestures must coordinate with a visible physical agent that moves through shared space.

2.2 Human-drone and human-robot interaction in public

Elicitation and metaphor-based approaches have been used to study how people want to command drones, often yielding spatial, upper-body, or interpersonal gestures [1, 7]. Related work on flying robots further highlights that motion can communicate intent and shape perceived safety [13]. Beyond interaction techniques, public acceptance of drones is shaped by broader concerns such as privacy, safety, and security [2, 14]. Perceptions of drone form factors and physical characteristics can also influence how drones are perceived [15]. Together, these findings suggest that expressive gestures may feel appropriate in this domain, but they also raise questions about bystander comfort and the social costs of conspicuous control.

2.3 Design framing: norms, audience, and interpretability

Systematic reviews argue that social norms influence comfort, trust, and acceptance in HRI and should be treated as design constraints [6]. For public robot control, the relevant audience extends beyond the operator. Gestures that appear more “legible” may reduce ambiguity for bystanders and help communicate that the robot is under intentional control, but these same gestures may increase conspicuousness and social exposure for the operator. This framing motivates studying gesture-based control as a multi-stakeholder interaction in which operator comfort and bystander interpretability may not align.

2.4 Summary and Research Questions

Prior work therefore points to competing design pressures: subtle gestures may maximize social acceptability for operators, while expressive, spatial gestures may better fit the mental model of commanding a visible moving robot and may increase perceived interpretability for observers. We investigate this tension by crossing gesture style (discreet vs. expressive) with social context (private vs. public with bystanders) and measuring operator responses, supplemented by limited bystander observations when feasible.

In sum, our research questions were:

RQ1: How does gesture style (discreet vs. expressive) affect operator preference, perceived control, and willingness to use?

RQ2: How does social context (private vs. public with bystanders) interact with gesture style to shape these outcomes?

RQ3 (attempted): How do bystanders perceive comfort, safety, and interpretability for discreet vs. expressive control?

Where our hypotheses were:

H1: Discreet gestures increase willingness to use in public, but may reduce perceived control/precision.

H2: Expressive gestures increase perceived justification/legibility (i.e., that the interaction “looks right” and is interpretable to others), particularly in public contexts.

Note: We attempted brief bystander intercept probes during public blocks, but most passersby declined, were time-constrained, or reported limited attention; the resulting sample was too small and incomplete for quantitative analysis. Accordingly, we treat RQ3 as a motivating design consideration and focus analyses on operator outcomes.

3 Method

3.1 Apparatus

Our system comprises (1) a wearable input device for capturing gestures, and (2) a clearly visible DIY 2DOF robot that executes basic navigation commands. We built a web server, ESP-based firmware, and a 3D-printable robot chassis (CAD) to support end-to-end, real-time gesture control. The robot provides real time guidance through six onboard LEDs that indicate where and when to move it during the task.

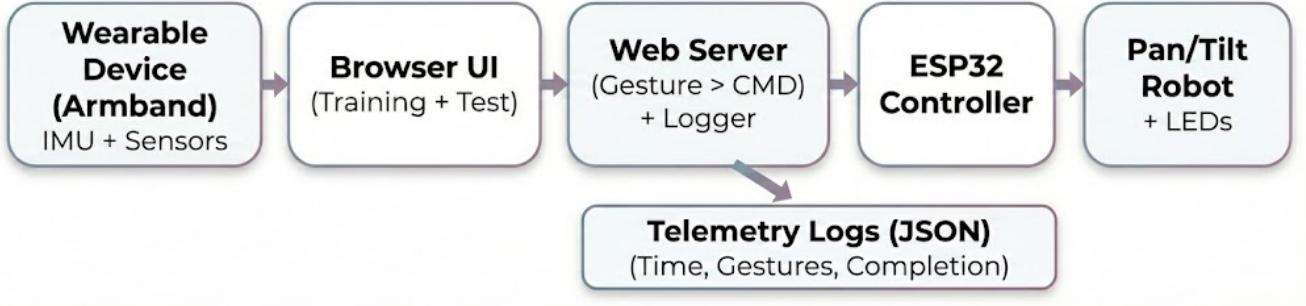
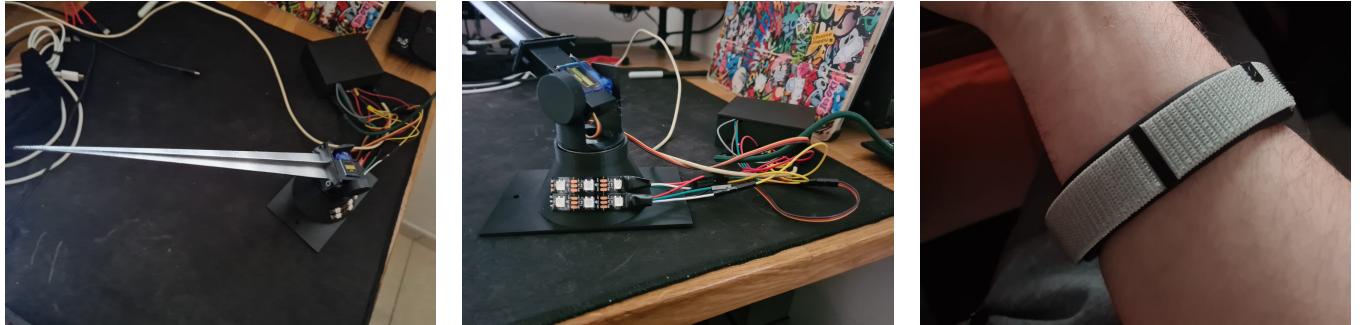


Figure 1: System overview and logging pipeline. The wearable provides gesture inputs to the browser-based UI, which forwards events to a web server that issues motion commands to an ESP32 controller driving the pan–tilt robot. We log timestamped events and per-target summaries for analysis (time, gestures, and completion).

Wearable input device. Participants wore a armband device that captured motion data for gesture recognition. The device includes 3 SNC neural sensors and 6-DoF IMU Accelerometer and Gyroscope sampled at 1000 Hz and transmitted wirelessly to the server with Bluetooth low-energy (BLE). We recorded raw sensor streams and recognized command events with timestamps for later analysis.

Robot platform. The robot is a 3D-printed, pan–tilt platform (2DOF) actuated by SG90 servos. Although not a free-flying drone, the robot is intentionally visually noticeable and moves in a way that is readily observable to nearby people, enabling us to study gesture control in a public-facing setting without flight safety risks.



(a) Pan–tilt robot (3D printed). (b) LED cue ring (direction/progress). (c) Wearable input device.

Figure 2: Study apparatus. We used a visible pan–tilt robot with on-board LED cues and a wearable gesture input device.

LED guidance. Six LEDs mounted on the robot, in a 2 by 3 grid, provided task guidance (e.g., indicating directionality or target orientation). During each trial, LEDs signaled:

- **Idle state:** a green light was used to indicate the target direction (middle LEDs for up/down and bottom row left/right for left/right).
- **Moving State:** only one light remains lit, in either solid white or yellow. The white light indicates movement, and the yellow light indicates movement in the target direction.
- **When reaching target** a solid blue light indicates successful arrival at the target position.

Specifically for the discreet gesture set, the top right and left LEDs were used to indicate succesful gesture recognition, where the top left LED indicated neutural tap/pinch gestures, and the top right LED indicated reverse tap gestures.

Study software + logging. We implemented a browser-based study interface backed by a web-server that (a) receives interpreted gesture inputs from the wearable device, (b) sends motion commands to the robot, and (c) logs timestamped events (gesture events, target achievements, and end-of-block summaries). For each block, the system produces a run log (JSON) containing per-target completion times and gesture counts, as well as aggregate metrics (total time, total gestures).

3.2 Gesture Sets

We compare two gesture styles that implement the same command vocabulary but differ in execution style.

- **Discreet:** low-amplitude wrist/hand movements performed close to the torso, designed to minimize visibility and social conspicuity. To execute them, users perform combinations of tap/pinch and reverse tap gestures (Table 1).
- **Expressive:** higher-amplitude arm gestures (involving elbow/shoulder) that visibly indicate direction, designed to maximize external legibility.

Command set. The command vocabulary consists of four discrete commands: move up, down, turn left, and turn right. Commands were triggered upon recognition of the corresponding gesture and held as long as a pinch hold (and position for expressive gestures) is being held.

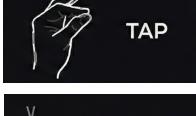
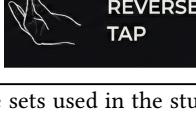
Discreet (composed)				Expressive (directional)	
Command	Step 1	Step 2	Step 3	Command	Gesture
Up	 TAP	 PINCH	—	Up	 PINCH + UP
	 REVERSE TAP	 PINCH	—		 PINCH + DOWN
Down	 TAP	 TAP	 PINCH	Left	 PINCH + LEFT
	 REVERSE TAP	 REVERSE TAP	 PINCH		 PINCH + RIGHT

Table 1: Gesture sets used in the study UI. Discreet gestures are short, close-to-body combinations of tap/pinch or reverse tap/pinch; expressive gestures are larger directional arm movements.

Gesture recognition and verification. Before completing study tasks, participants underwent a structured training phase for both gesture sets. We verified recognition reliability by having participants perform 5-6 random movement commands. We logged recognition errors (false positives/negatives) and response latency to contextualize subjective ratings of control and willingness.

3.3 Design

We use a 2×2 within-subjects design crossing:

- (1) **Gesture style:** discreet vs. expressive.
- (2) **Social context:** private vs. public with bystanders present.

Independent variables	Levels
Gesture style	Discreet; Expressive
Context	Private; Public w/ bystanders
Primary DVs (operator)	Preference; willingness; embarrassment; perceived control
Primary DVs (telemetry)	Time; gesture count; completion

Table 2: Study design overview.

3.4 Participants

We recruited 9 adults (ages 27–37; 4 female and 5 male). We analyze 6 participants (3 female and 3 male); 3 sessions were excluded due to missing or highly incomplete logs due to participant time constraint stoppage. Among the analyzed participants, due to frustration, 3

stopped early during the discreet condition while attempting target 3 or 4 (in both private and public blocks); we include these partial blocks and mark them explicitly in the target-progress plots.

Bystanders (attempted intercepts). In public blocks, we attempted brief intercept prompts with passersby immediately after they observed part of a block. In practice, most declined, were time-constrained, or reported that they did not observe enough to answer. As a result, we did not obtain a consistent bystander dataset and do not report quantitative bystander outcomes.

3.5 Procedure

Each session followed the sequence below.

- (1) **Consent and pre-survey:** demographics, prior experience with robots/drones and gesture interfaces, and baseline comfort using body movement in public.
- (2) **System introduction and safety briefing:** explanation of the robot's motion limits and the meaning of LED guidance cues.
- (3) **Training:** participants practiced both gesture sets with real-time feedback until meeting the verification criterion. Training order matched the upcoming block order to reduce interference.
- (4) **Task blocks (4 total):** participants completed navigation tasks under all four combinations of gesture style and context. After each block, participants completed brief self-report measures.
- (5) **Public blocks: attempted bystander intercepts:** we approached nearby passersby during public blocks to request brief impressions; participation was sparse and incomplete, and we therefore do not analyze bystander outcomes.
- (6) **Exit survey and interview:** comparative preferences, open-ended feedback on trade-offs, and suggestions for public-facing gesture design.

3.6 Task and command vocabulary

Participants completed repeated pose acquisition tasks: they issued commands to move the robot toward a sequence of target pan–tilt poses. Targets were indicated through the robot's LED cues, which provided directional guidance for how to adjust pose. A target was counted as complete when the robot entered a tolerance region around the target pose. Participants could correct overshoot and movements away from the target. Participants could end trials by signaling the experimenter or pressing the relevant button in the web interface to stop.

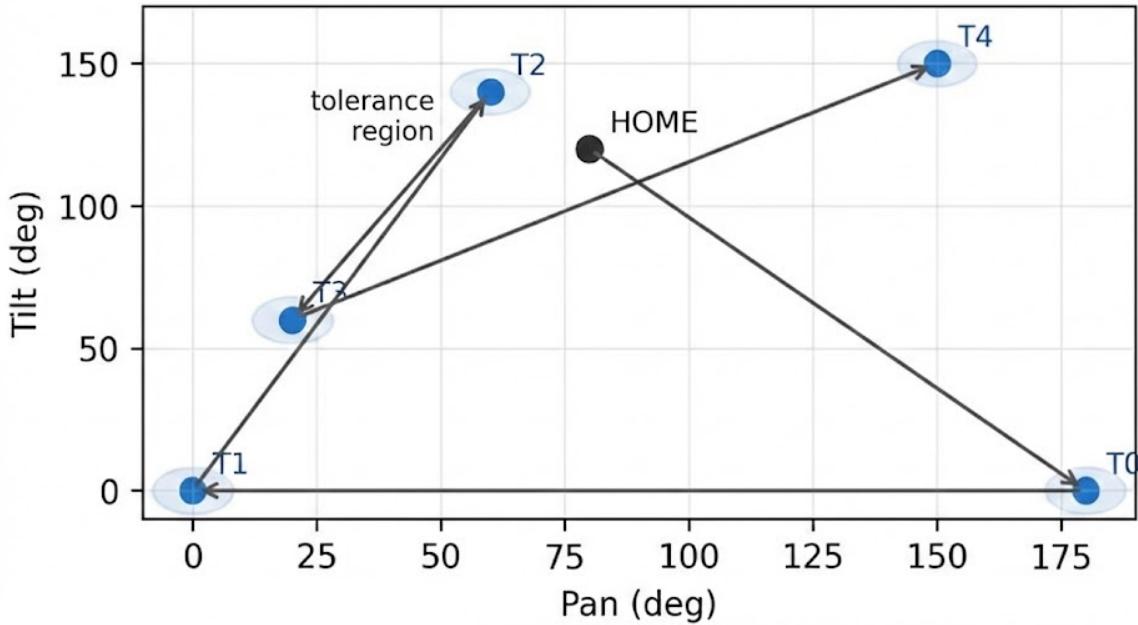


Figure 3: Task schematic. Participants used gestures to iteratively move the robot toward sequential pan–tilt targets. LEDs on the robot provided directional guidance toward the current target.

Operationalizing social context. In the private condition, the task took place in what participants described as a private setting (lab/office/closed room) with only the participant and experimenter present (no bystanders in view). In the public condition, the task took place in nearby public spaces with naturalistic passersby present.

3.7 Measures

Operator self-report. After each block, participants reported comfort, perceived performance, control/confidence, social acceptability, justification/legibility, and workload, plus an open-ended diagnostic response. After all blocks, participants completed comparative items (e.g., preference in public vs private, precision, embarrassment, justification) and acceptability checklists (people/places). See Appendix A for full item text.

Behavioral + telemetry (from logs). From the per-block run logs we compute: targets completed and completion rate, total elapsed time, total gesture count, and cumulative time/gestures to reach each sequential target (targets 1–5). For blocks that ended early, we also include partial progress on the attempted target where the block stopped (visualized separately).

Bystander perception (planned; limited collection). We intended to analyze bystander comfort observing, perceived safety, and understandability/legibility of the mapping from gesture to robot action, but recruitment during public blocks yielded too few consistent responses for quantitative analysis.

3.8 Analysis plan

We focus on within-participant effects of gesture style and social context. We report condition summaries (median and IQR) and paired within-subject comparisons. To test moderation by context, we compute an interaction-style contrast via a difference-of-differences: $(E-D)_{\text{public}} - (E-D)_{\text{private}}$. Where appropriate, we supplement effect sizes with nonparametric paired inference (e.g., sign-flip/permuation tests) and bootstrap confidence intervals.

4 Results

4.1 Manipulation checks

Across sessions, the study system logged distinct event types for the two gesture sets (Table 1). Discreet blocks produced click-like events (LeftClick/RightClick), while expressive blocks produced directional events (ArrowUp/Down/Left/Right). All blocks used the same target plan (5 sequential targets) and comparable controller parameters (step size and tolerance), enabling a direct comparison across conditions.

4.2 Operator outcomes

We analyze $n=6$ participants (24 total blocks: 2 gesture sets \times 2 contexts). For blocks that ended early, we include partial progress on the attempted target and visualize it explicitly rather than excluding it.

Task performance (telemetry). Expressive gestures achieved complete task success across blocks (median targets completed = 5 [IQR 5–5]; median completion rate = 1.0 [1.0–1.0]). Discreet gestures were less reliable (median targets completed = 4 [3–5]; median completion rate = 0.8 [0.6–1.0]), and 6/12 discreet blocks were stopped early (3 participants; both contexts), always while attempting target 4. Figure 4 summarizes targets completed and completion rates by gesture set.

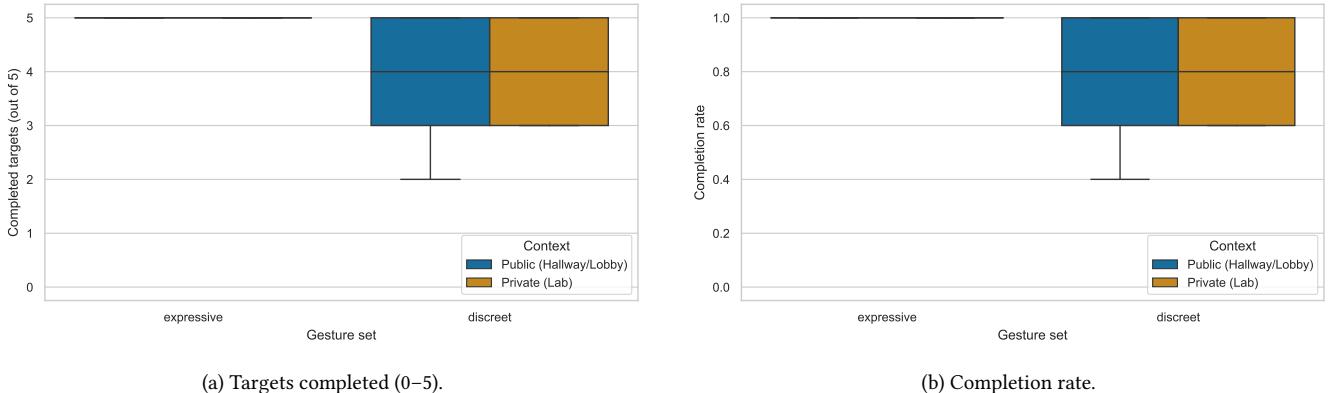


Figure 4: Task completion outcomes by gesture set. Points show medians; bars show IQR.

Efficiency (time and gestures). Expressive gestures required substantially fewer actions and less time to make progress, both in private and in public settings, and reached all five targets in every block. In contrast, discreet performance bifurcated: half of discreet blocks completed all five targets, while the other half were stopped by target 4. Figure 5 shows cumulative progress to targets 1–5, with early-stopped blocks shown as hatched bars at the target where the block ended.

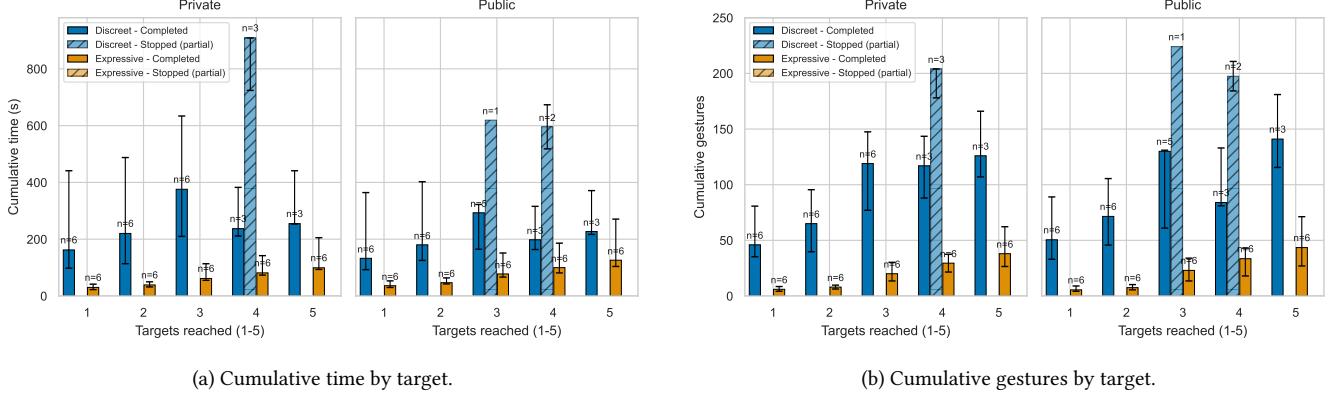


Figure 5: Efficiency outcomes by gesture set and progress through the 5-target sequence (median with IQR). Hatched bars denote early-stopped blocks at the target where the block ended.

Figure 6 further illustrates the per-target gesture counts, showing that discreet gestures required substantially more attempts to reach later targets.

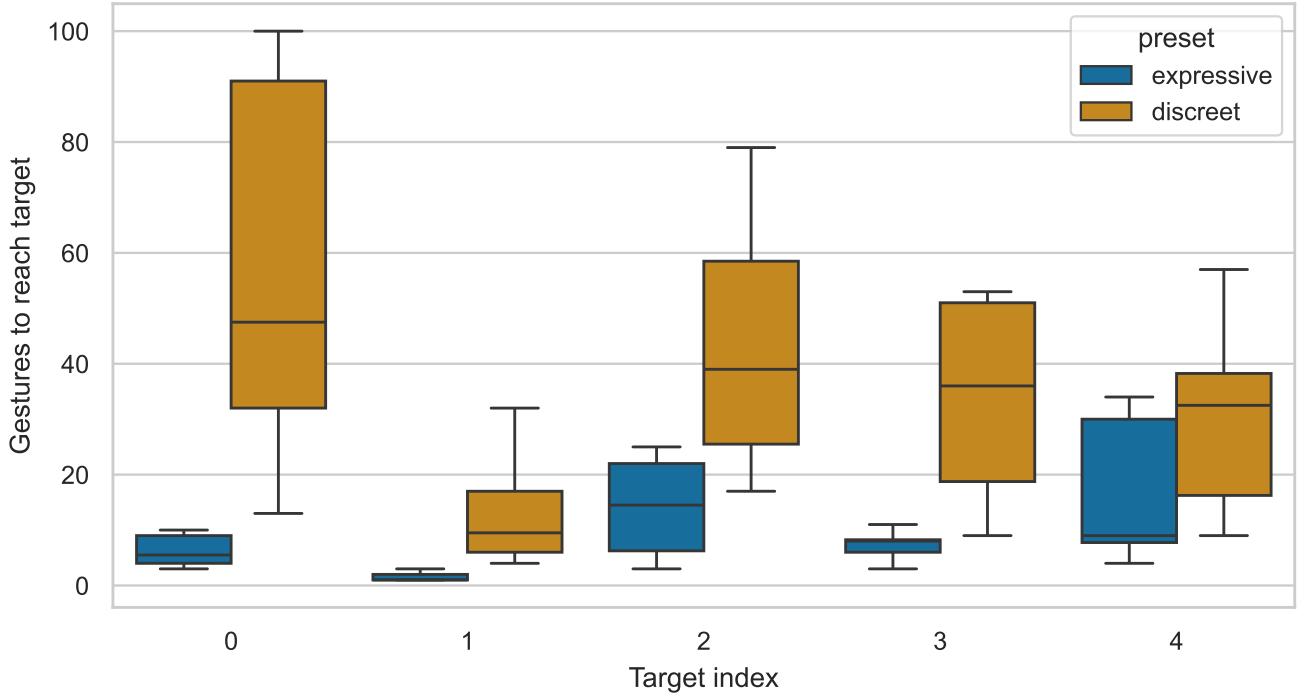


Figure 6: Gestures per target by target index (median with IQR). Discreet gestures required more attempts, particularly for later targets.

Subjective ratings after each block. Figure 7 summarizes block questionnaire ratings. Participants rated expressive as higher-performing and more controllable, with lower workload. In particular, expressive increased perceived performance (median 5.0 vs. 2.0 for discreet) and

control/confidence (median 5.75 vs. 2.62), and reduced workload (median 2.38 vs. 4.25; lower is better). Expressive also increased perceived justification/legibility, and this advantage was larger in public blocks (paired mean difference in justification mean: +2.42 in public vs. +1.25 in private). In contrast, comfort and social acceptability were similar across gesture sets in public blocks, suggesting that (in this study) performance and control dominated any social benefit of discreetness.

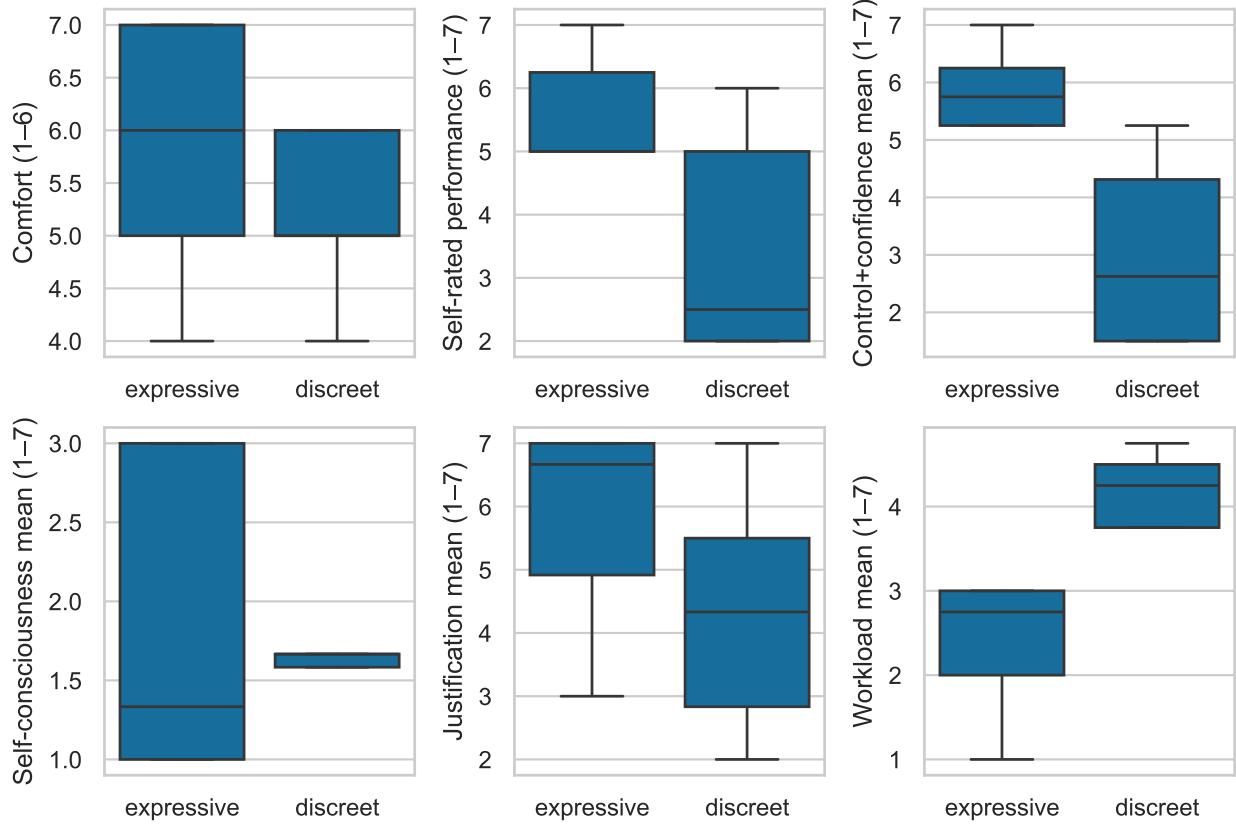


Figure 7: Block questionnaire ratings by gesture set. Points show medians; bars show IQR.

Final comparative preferences and acceptability. In the final comparative items (Likert: 1=Discreet, 7=Expressive), participants preferred expressive gestures in both public and private contexts (mean preference $\approx 6.0\text{--}6.5$), and rated expressive as more precise (mean 6.75) and more justified (mean 5.25). Expressive was rated as slightly more embarrassing (mean 4.25), suggesting a mild social cost that was outweighed by gains in efficiency and perceived control (Figure 8).

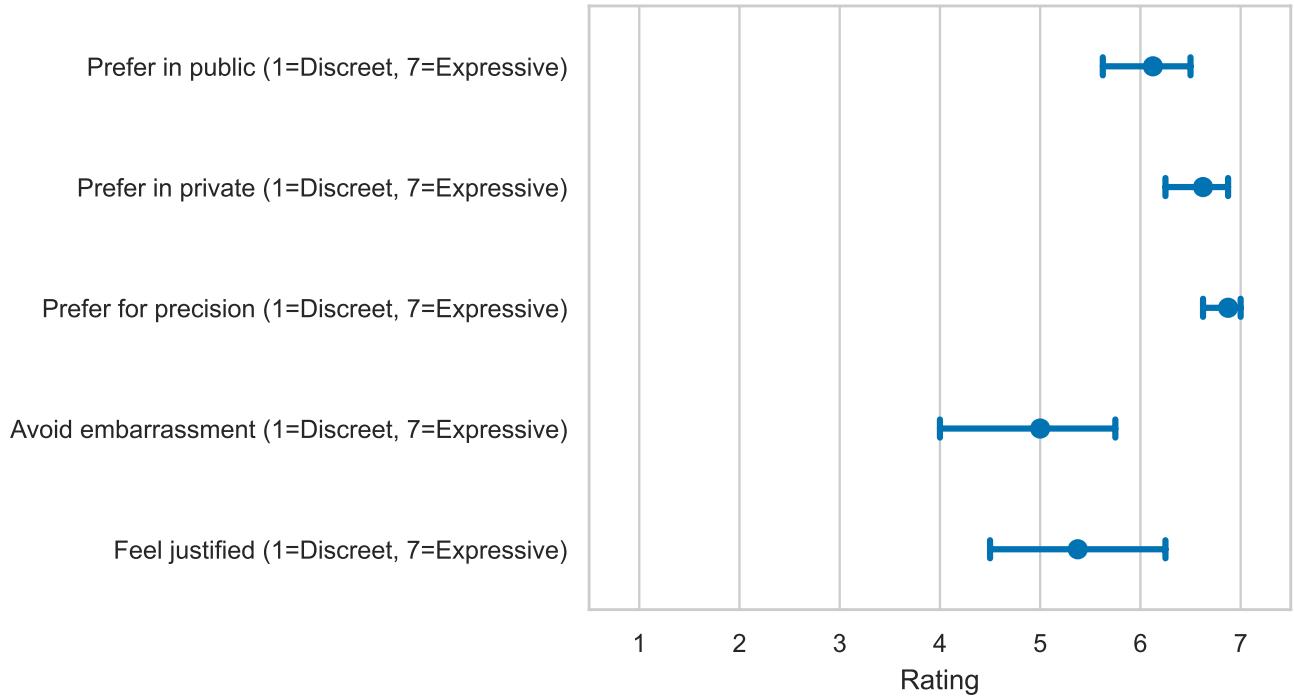


Figure 8: Final comparison ratings (1=Discreet, 7=Expressive): mean with 95% CI.

Acceptability checklists indicated broad willingness to use both sets across audiences and locations, including around strangers and in public places (Figure 9). Notably, high acceptability did not imply comparable *performance*: discreet gestures were often acceptable in principle, but participants still reported mis-recognition and loss of control in practice.

4.3 Bystander outcomes

We attempted brief intercept prompts with passersby during public blocks, but participation was sparse and inconsistent. As a result, we do not report quantitative bystander measures. However, a practical observation emerged: many bystanders did not watch long enough to form a confident interpretation of the mapping.

4.4 Qualitative themes

Open-ended diagnostics reinforce the quantitative pattern that recognition reliability and perceived control were the dominant drivers of preference. Across participants, we observed three recurring themes:

- **Discreet gestures traded social subtlety for reliability.** Participants described discreet control as “not intuitive” and reported missed recognitions that led to repeated attempts and frustration.
- **Expressive gestures felt more direct and predictable.** Participants described expressive control as “more intuitive” and “quicker” to correct the robot’s pose.
- **Public context did not dominate experience in this setting.** At least one participant explicitly noted that the crowd “didn’t disturb,” consistent with the small differences on the social acceptability scale.

5 Discussion

5.1 Summary of findings

Across both contexts, expressive gestures outperformed discreet gestures on objective telemetry (completion, time, and gesture count) and were rated as more controllable, higher-performing, and less mentally/physically demanding. While discreet gestures were designed to reduce public awkwardness, we did not observe a corresponding improvement in comfort or social acceptability during public blocks. Instead, the primary determinant of preference was *control quality*: when discreet recognition failed or required repeated retries, its social subtlety did not compensate for the loss of predictability.

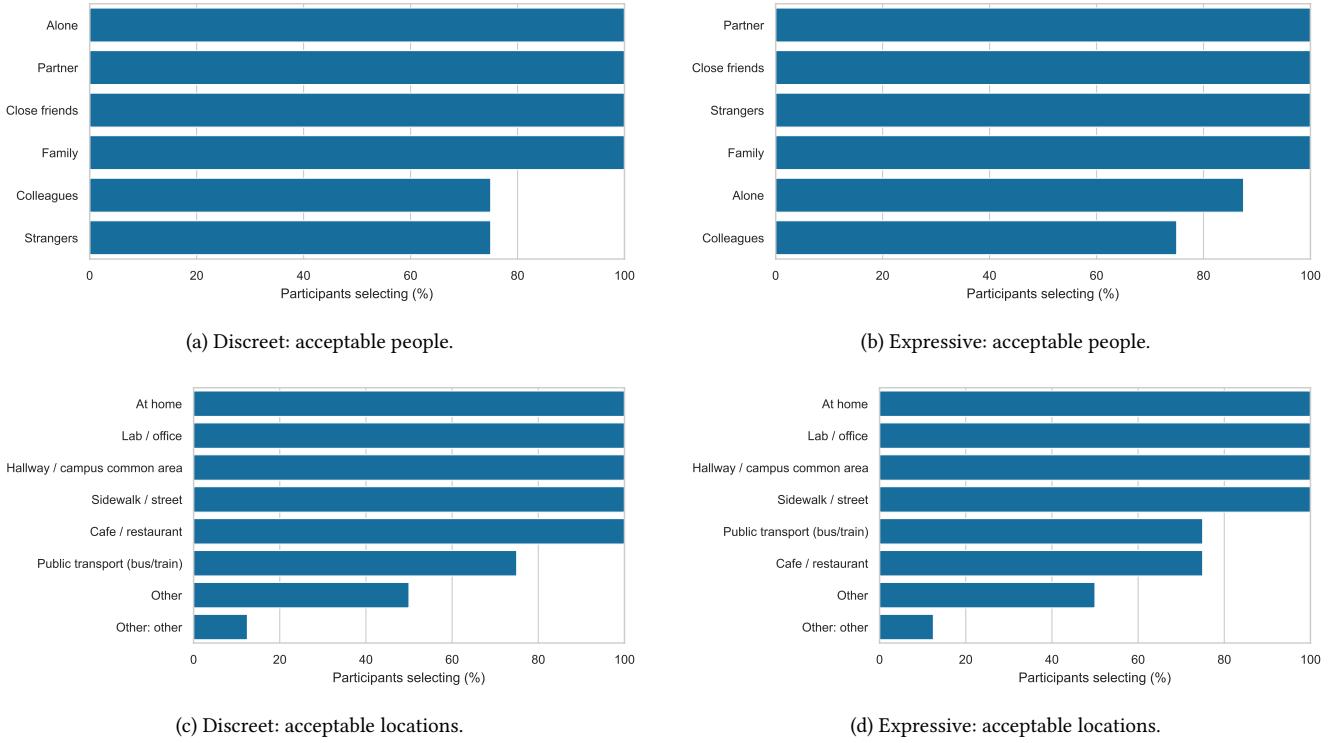


Figure 9: Acceptability checklists (percentage selecting each category). Denominator = participants with a response for this item.

5.2 Why expressive “won” even in public

Prior work on mobile and wearable gestures often finds that subtle gestures are preferred around others. Our results suggest that when controlling a *visible moving object*, the calculus can shift: expressive gestures may feel more appropriate to the task, and may help users form a clear mental model of the mapping between action (gesture) and outcome (robot motion). In our study, participants rated expressive as more justified/legible, and this advantage was especially pronounced in public blocks. Importantly, the robot itself served as an attention magnet; once a moving robot is present, the marginal social cost of a larger gesture may be small compared to the benefit of completing the task quickly and confidently.

5.3 Design implications

1 - Make reliability the first-class social feature. Social acceptability is fragile when recognition is unreliable. Participants’ frustration with missed recognitions in the discreet set indicates that even subtle gestures can become *socially risky* if they require repeated attempts or lead to visible trial-and-error. For public deployment, designers should prioritize robust recognition (or explicit confirmation) for subtle gestures, and provide immediate, unambiguous feedback for both recognition and command execution.

2 - Provide an escalation ladder: subtle by default, expressive on demand. Rather than choosing between discreet and expressive, systems can support both as complementary modes: use subtle gestures for routine adjustments and allow a deliberate “escalation” (e.g., a hold or mode switch) for expressive, highly legible commands when users need clarity or recovery. This aligns with our finding that expressive improved recovery and confidence, while discreet may still be desirable when users want to minimize movement.

3 - Shift legibility from the body to the robot when possible. If expressive gestures are intended to help bystanders understand what is happening, the same goal may be achieved with robot-side cues: directional LEDs, motion previews, or brief confirmation animations. This can reduce the need for large public gestures while preserving interpretability for observers.

5.4 Measuring bystanders in the wild

Our difficulty collecting bystander intercept data highlights a methodological challenge: real-world bystanders often lack the time, attention, or perceived relevance to answer questions about an interaction they only glimpsed briefly. Future work can complement in-situ probes with alternative approaches such as (a) short video vignettes shown to recruited participants, (b) passive measures of attention (e.g., gaze or dwell time) with strong privacy protections, or (c) structured public demonstrations with opt-in observation areas.

6 Limitations

During the study, we encountered several limitations:

- **Sample size and diversity:** limited generalizability due to a small sample; participants were affiliated with a university setting.
- **Environment setting:** the private setting varied (lab/office/room) based on participant constraints; the public setting was typically a high-traffic indoor campus area and may not generalize to outdoor or non-academic environments.
- **Robot capabilities:** due to practical and regulatory constraints, we used a DIY pan–tilt robot rather than a free-flying drone. This limits motion dynamics and may affect perceived risk and realism.
- **Wearable recognition limits:** the wearable device and recognition pipeline exhibited variable accuracy, which strongly influenced perceived control (especially for discreet gestures). The device also required tight fastening, which some participants found uncomfortable.
- **Measurement constraints:** reliance on self-report measures can introduce bias; future work should complement them with additional objective measures (e.g., failure modes, recovery time, and observer attention).

7 Ethics and Safety

All participants provided informed consent prior to participation. Safety measures were implemented to ensure that the robot operated within safe parameters, including speed limits and emergency stop functions. Participants were briefed on safety procedures and instructed to stop the robot immediately if they felt uncomfortable at any point during the study.

8 Conclusion

We compared discreet and expressive gesture sets for controlling a visible pan–tilt robot across private and public-with-bystanders contexts. In this study, expressive gestures consistently produced better task outcomes (higher completion, lower time, and fewer gestures) and were perceived as more controllable and effective, with lower workload. We did not observe a strong social comfort advantage for discreet gestures in the public condition; instead, recognition reliability and perceived control were the dominant drivers of preference. These results suggest that for public robot control, designers should treat robustness and clear feedback as foundational to social acceptability, and consider hybrid designs that enable subtle default control with optional expressive escalation when legibility or recovery is needed.

References

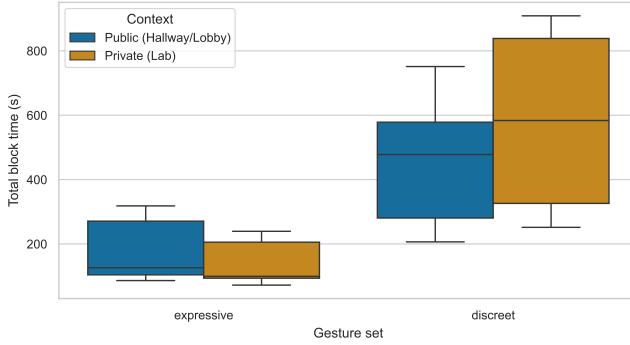
- [1] Jessica R. Cauchard, Jane L. E, Kevin Y. Zhai, and James A. Landay. 2015. Drone & Me: An Exploration into Natural Human-Drone Interaction. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. Association for Computing Machinery, 361–365. doi:10.1145/2750858.2805823
- [2] Victoria Chang, Pramod Chundury, and Marshini Chetty. 2017. Spiders in the Sky: User Perceptions of Drones, Privacy, and Security. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, 6765–6776. doi:10.1145/3025453.3025632
- [3] Yi-Ta Hsieh, Antti Jylhä, Valeria Orso, Luciano Gamberini, and Giulio Jacucci. 2016. Designing a Willing-to-Use-in-Public Hand Gestural Interaction Technique for Smart Glasses. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. Association for Computing Machinery, 4203–4215. doi:10.1145/2858036.2858436
- [4] Norene Kelly and Stephen Gilbert. 2016. The WEAR Scale: Developing a Measure of the Social Acceptability of a Wearable Device. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. Association for Computing Machinery, 2864–2871. doi:10.1145/2851581.2892331
- [5] Marion Koelle, Swamy Ananthanarayan, and Susanne Boll. 2020. Social Acceptability in HCI: A Survey of Methods, Measures, and Design Strategies. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery. doi:10.1145/3313831.3376162
- [6] Steven Lawrence, Melanie Jouaiti, Jesse Hoey, Christopher L. Nehaniv, and Kerstin Dautenhahn. 2025. The Role of Social Norms in Human–Robot Interaction: A Systematic Review. *ACM Transactions on Human-Robot Interaction* 14, 3, Article 56 (2025), 44 pages. doi:10.1145/3722120
- [7] Kevin P. Pfeil, Seng Lee Koh, and Joseph J. LaViola. 2013. Exploring 3D Gesture Metaphors for Interaction with Unmanned Aerial Vehicles. In *Proceedings of the 18th International Conference on Intelligent User Interfaces (IUI '13)*. Association for Computing Machinery, 257–266. doi:10.1145/2449396.2449429
- [8] Halley P. Profitta, James Clawson, Scott Gilliland, Clint Zeagler, Thad Starner, Jim Budd, and Ellen Yi-Luen Do. 2013. Don't Mind Me Touching My Wrist: A Case Study of Interacting with On-Body Technology in Public. In *Proceedings of the 17th International Symposium on Wearable Computers (ISWC '13)*. Association for Computing Machinery, 89–96. doi:10.1145/2493988.2494331
- [9] Julie Rico and Stephen Brewster. 2009. Gestures All Around Us: User Differences in Social Acceptability Perceptions of Gesture Based Interfaces. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '09)*. Association for Computing Machinery, 64:1–64:2. doi:10.1145/1613858.1613936
- [10] Julie Rico and Stephen Brewster. 2010. Usable Gestures for Mobile Interfaces: Evaluating Social Acceptability. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. Association for Computing Machinery, 887–896. doi:10.1145/1753326.1753458
- [11] Marcos Serrano, Barrett Ens, and Pourang Irani. 2014. Exploring the Use of Hand-to-Face Input for Interacting with Head-Worn Displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. Association for Computing Machinery, 3181–3190. doi:10.1145/2556288.2556984
- [12] Calkin Sueru Montero, Jason Alexander, Mark T. Marshall, and Sriram Subramanian. 2010. Would You Do That?: Understanding Social Acceptance of Gestural Interfaces. In *Proceedings of the 12th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '10)*. Association for Computing Machinery, 275–278. doi:10.1145/1851600.1851647
- [13] Daniel Szafir, Bilge Mutlu, and Terrence Fong. 2014. Communication of Intent in Assistive Free Flyers. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction (HRI '14)*. Association for Computing Machinery, 358–365. doi:10.1145/2559636.2559672
- [14] Ning Wang, Nico Mutzner, and Karl Blanchet. 2023. Societal acceptance of urban drones: A scoping literature review. *Technology in Society* 75 (2023), 102377. doi:10.1016/j.techsoc.2023.102377
- [15] Anna Wojciechowska, Jeremy Frey, Esther Mandelblum, Yair Amichai-Hamburger, and Jessica R. Cauchard. 2019. Designing Drones: Factors and Characteristics Influencing the Perception of Flying Robots. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3, Article 111 (2019), 19 pages. doi:10.1145/3351269

A Supplementary Materials

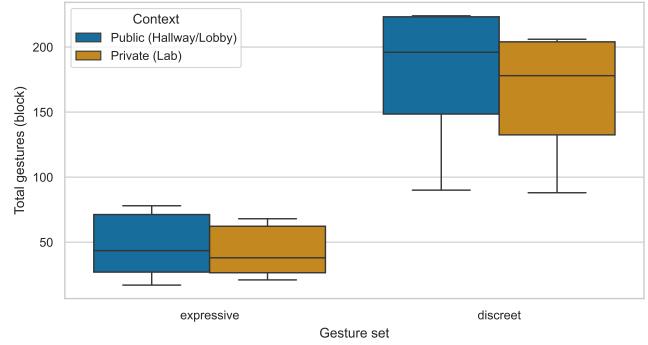
All code and materials are available at: <https://github.com/Kiiwii10/discreet-vs-expressive-gestures>

B Supplementary Telemetry Figures

This appendix provides additional visualizations of task telemetry to complement the main results.

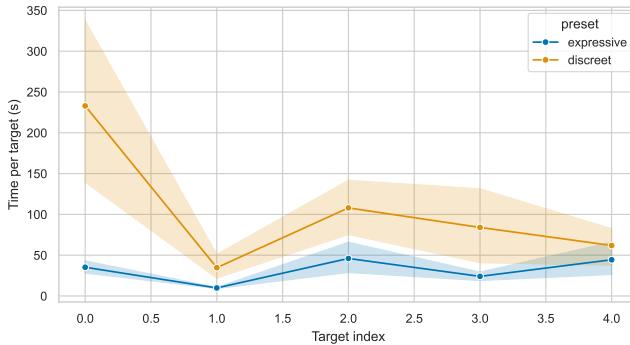


(a) Total block time (s).

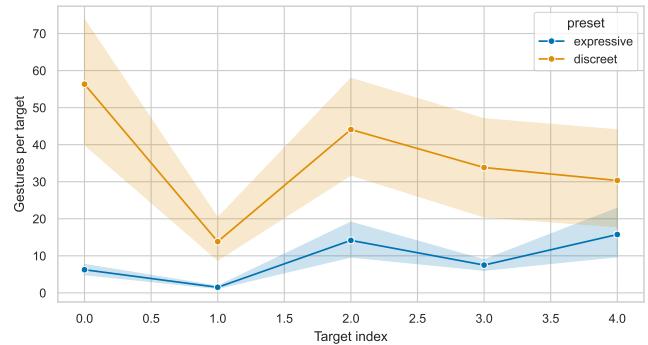


(b) Total gestures per block.

Figure 10: Aggregate efficiency metrics by gesture set (median with IQR).

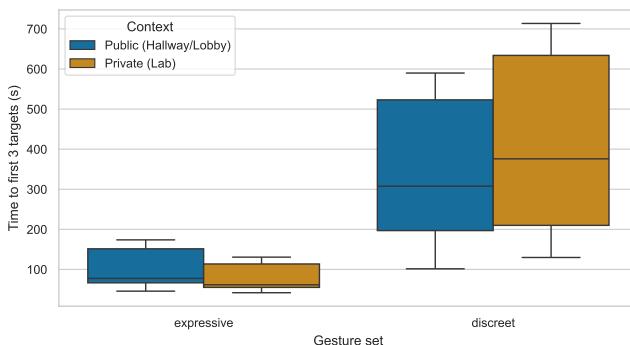


(a) Time per target (mean with 95% CI).

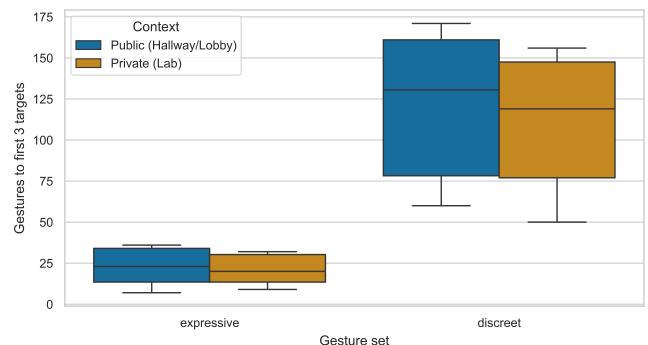


(b) Gestures per target (mean with 95% CI).

Figure 11: Per-target efficiency trends by gesture set (mean with 95% CI).



(a) Time for first 3 targets (s).



(b) Gestures for first 3 targets.

Figure 12: Early-block efficiency (first 3 targets) by gesture set (median with IQR). Restricting to the first 3 targets avoids confounding from early-stopped blocks.

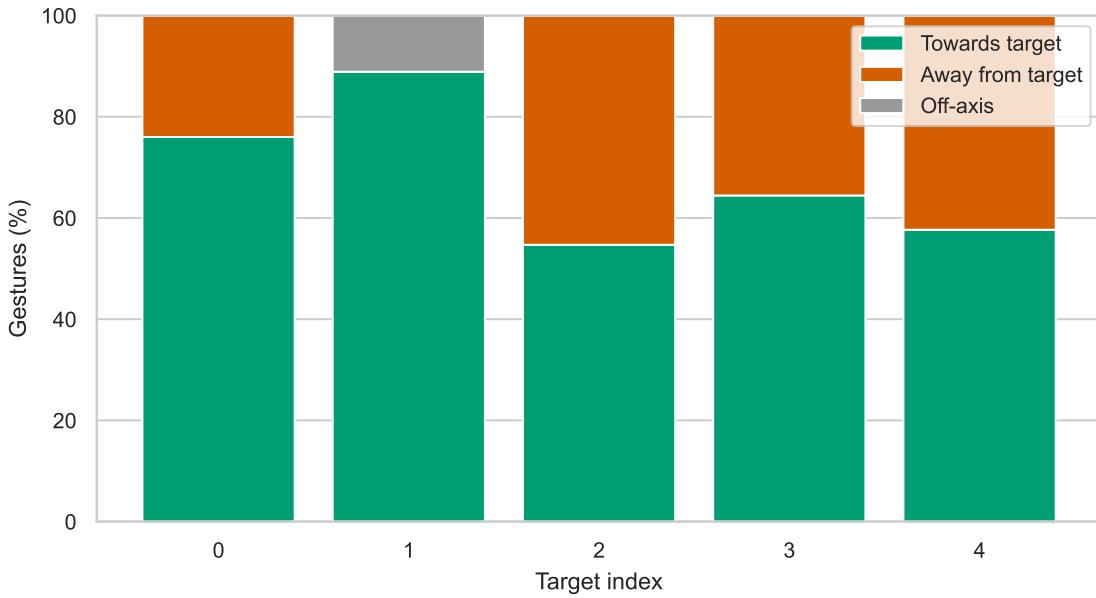


Figure 13: Expressive only: share of Arrow-key gestures classified as towards/away/off-axis relative to the start-of-target delta. This illustrates directional accuracy during expressive control.

C Questionnaire Items

Question	Response type	Context
Participant ID	Short text	Consent stage at session start.
Study condition order (Discreet first / Expressive first)	Single choice	Consent stage at session start.
Environment type (Private lab / Public hallway)	Single choice	Consent stage at session start.
Age	Number (13–99)	Pre-study background.
Gender (optional)	Single choice (Male/Female/Other)	Pre-study background.
Dominant hand	Single choice (Right/Left/Both)	Pre-study background.
Prior experience (check all that apply)	Multi-select checklist (Gesture control; Drones/RC; Robotics; Wearables; None)	Pre-study background.
I generally feel comfortable doing noticeable actions in public.	Likert 1–7 (Strongly disagree to Strongly agree)	Pre-study baseline.
I often worry about what strangers think of my behavior in public.	Likert 1–7 (Strongly disagree to Strongly agree)	Pre-study baseline.
I try to avoid drawing attention to myself in public spaces.	Likert 1–7 (Strongly disagree to Strongly agree)	Pre-study baseline.

Table 3: Consent and pre-study questions (asked before any test blocks).

Question	Response type	Context
If you saw someone else doing what you just did, what would you think they were doing? (2–6 words)	Short text	After each block, private condition only.
How did you feel performing these gestures here, right now? I felt in control of the robot's movement.	Scale 1–7 (Embarrassed to Comfortable) Likert 1–7 (Strongly disagree to Strongly agree)	After each block. After each block.
The robot responded the way I expected.	Likert 1–7 (Strongly disagree to Strongly agree)	After each block.
I felt confident I executed the gestures correctly.	Likert 1–7 (Strongly disagree to Strongly agree)	After each block.
When I made a mistake, I could recover quickly.	Likert 1–7 (Strongly disagree to Strongly agree)	After each block.
I felt insecure while performing these gestures.	Likert 1–7 (Strongly disagree to Strongly agree)	After each block.
I worried that others were judging me.	Likert 1–7 (Strongly disagree to Strongly agree)	After each block.
I held back or made my movements smaller to avoid attention.	Likert 1–7 (Strongly disagree to Strongly agree)	After each block.
Because the robot's movement was visible, the gestures felt socially justified.	Likert 1–7 (Strongly disagree to Strongly agree)	After each block.
I think bystanders could tell I was controlling the robot.	Likert 1–7 (Strongly disagree to Strongly agree)	After each block, public condition only.
I think bystanders could predict what the robot would do from my gestures.	Likert 1–7 (Strongly disagree to Strongly agree)	After each block, public condition only.
Mental demand	Likert 1–7 (Very low to Very high)	After each block.
Physical demand	Likert 1–7 (Very low to Very high)	After each block.
Effort	Likert 1–7 (Very low to Very high)	After each block.
Frustration	Likert 1–7 (Very low to Very high)	After each block.
I would rate my performance in this block as	Scale 1–7 (Failure to Perfect)	After each block.
What specifically made this block feel comfortable or uncomfortable? (1–2 sentences)	Long text	After each block.

Table 4: Block questions (asked after each block of targets for Discreet or Expressive).

Question	Response type	Context
Gesture Set: Discreet – In front of whom would you feel comfortable using it?	Multi-select checklist (Alone; Partner; Close friends; Family; Colleagues; Strangers)	End of session.
Gesture Set: Discreet – In what locations would you feel comfortable using it?	Multi-select checklist (Home; Lab/office; Hallway/common area; Sidewalk/street; Public transport; Cafe/restaurant; Other + text)	End of session.
Gesture Set: Expressive – In front of whom would you feel comfortable using it?	Multi-select checklist (Alone; Partner; Close friends; Family; Colleagues; Strangers)	End of session.
Gesture Set: Expressive – In what locations would you feel comfortable using it?	Multi-select checklist (Home; Lab/office; Hallway/common area; Sidewalk/street; Public transport; Cafe/restaurant; Other + text)	End of session.
In a public setting, I would prefer to use:	Scale 1–7 (Prefer Discreet to Prefer Expressive)	End of session.
In a private setting, I would prefer to use:	Scale 1–7 (Prefer Discreet to Prefer Expressive)	End of session.
For precision, I would prefer to use:	Scale 1–7 (Prefer Discreet to Prefer Expressive)	End of session.
To avoid embarrassment in public, I would prefer to use:	Scale 1–7 (Prefer Discreet to Prefer Expressive)	End of session.
To feel socially justified when controlling the visible robot, I would prefer to use:	Scale 1–7 (Prefer Discreet to Prefer Expressive)	End of session.

Table 5: Final questions (asked after completing both gesture sets; assume a visible robot).