2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

The Turkish Ice Cream Robot: Examining Playful Deception in Social Human-Robot Interactions

Hyeonseong Kim[1], Roy El-Helou[2], Seungbeen Lee[3], Sungjoon Choi[1] and Matthew Pan[2]

***Abstract*— Playful deception, a common feature in human social interactions, remains underexplored in Human-Robot Interaction (HRI). Inspired by the Turkish Ice Cream (TIC) vendor routine, we investigate how bounded, culturally familiar forms of deception influence user trust, enjoyment, engagement, and perceived entertainment value during robotic handovers. We design a robotic manipulator equipped with a custom end-effector and implement five TIC-inspired trick policies that deceptively delay the handover of an ice cream-shaped object. Through a mixed-design user study with 91 participants, we evaluate the effects of playful deception and interaction duration on user experience. Results reveal that TIC-inspired deception significantly enhances enjoyment and engagement, though reduces perceived safety and trust, suggesting a structured trade-off across the multi-dimensional aspects. Our findings demonstrate that playful deception can be a valuable design strategy for interactive robots in entertainment and engagement-focused contexts, while underscoring the importance of deliberate consideration of its complex trade-offs. Videos and user study snapshots are available on `https://hyeonseong-kim98.github.io/turkish-ice-cream-robot/`**

I. INTRODUCTION

In Human–Robot Interaction (HRI), deceptive behaviours of robots are often treated as harmful, as they tend to reduce user trust [1]–[4]. Prior studies have mainly focused on avoiding deception through predictable behaviours and intent displays [5]–[7], or on repairing trust [8]–[10]. On the other hand, empirical investigations into the potential benefits of robot deception have been conducted in game contexts where deception is permitted or even encouraged [11]–[13].

However, beyond game-like contexts where the deception is explicitly allowed, some playful deceptions also enrich interpersonal experiences in everyday interaction. Playful deception can create enjoyable, light-hearted moments, so long as it is interpreted as part of a benign performance rather than as manipulation. For instance, a human handing over an object might playfully delay the delivery or briefly mislead the receiver through unexpected but harmless actions, forming humour that is only possible through physical embodiment.

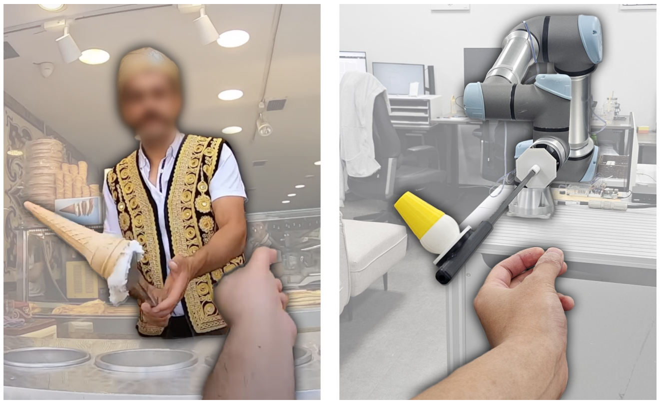

Fig. 1: **Conceptual illustration of the TIC robot.** Inspired by the playful techniques of Turkish ice cream vendors [14] (left), the robot is designed to humorously mislead the user while handing over an ice cream cone-shaped object (right).

The **Turkish ice cream (TIC)** vendor routine exemplifies this point: it relies on intentional misdirection to entertain while ultimately delivering the treat. The TIC interaction is a short street performance in which a vendor playfully prolongs the handover of an ice cream cone. Using a long spatula-like rod, the cone is presented and then briefly withdrawn, with light feints and showy gestures that tease the customer while keeping the exchange clearly playful. The routine is recognizably performative and always resolves with a successful handover, marking the deception as benign rather than malicious.

In this paper, we explore the potential role of playful deception as a design component in HRI through a study inspired by the TIC routine. To this end, as shown in Fig. 1, we design and implement a robotic handover system that reproduces TIC-inspired playful deceptive behaviours and conduct a user study with 91 participants to investigate the multi-dimensional effects of playful deception, such as enjoyment, trust, and perceived safety.

In summary, this paper makes three primary contributions:

- the design and implementation of a robotic system capable of reproducing TIC-inspired deceptive handovers, introducing a novel playful interaction scenario.
- a large-scale user study evaluating the multi-dimensional effects of playful TIC deception on user experience, revealing both its positive and negative impacts.
- the emphasis on the nuanced trade-offs that playful deception introduces and the importance of situating its use within appropriate contexts, offering guidance for both design and future research.

II. BACKGROUND

Robot deception in HRIs has often been regarded as problematic because it can undermine user trust and raise ethical concerns. Prior studies have focused on defining forms and analyzing the effect of robot deception [1]–[4], designing

---

[1] Hyeonseong Kim and Sungjoon Choi are with the Department of Artificial Intelligence, Korea University, Seoul, Republic of Korea (e-mails: [hyeonseong-kim, sungjoon-choi]@korea.ac.kr)

[2] Roy El-Helou and Matthew Pan are with the Ingenuity Labs Research Institute and the Department of Electrical and Computer Engineering, Queen's University, Kingston, Canada (e-mails: [19reh2, matthew.pan]@queensu.ca)

[3] Seungbeen Lee is with the Department of Artificial Intelligence, Yonsei University, Seoul, Republic of Korea (e-mail: seungblee@yonsei.ac.kr)

979-8-3315-8160-2/26/$31.00 ©2026 IEEE

motion strategies to avoid unintended misunderstandings [5]–[7], and developing methods to recover trust when it is damaged [8]–[10]. While these approaches generally assume that deception is inherently harmful, [15]–[17] suggest the potential benefit of robot deception in HRI scenarios. Some work has empirically shown that deception can improve engagement of the user and the perception about the robot's competence [11]–[13], but such findings are mostly limited to game settings where explicit rules allow deceptive strategies.

Although prior work has examined the multi-dimensional effects of benevolent deception in contexts such as patient and elderly care [18], its potential role in everyday HRI scenarios remains largely unexplored. In this work, we focus on the capability of playful deception in everyday interactions. This approach is grounded in *incongruity theory* [19], which posits that humour is triggered by the resolution of a perceived conflict between expectation and reality, and *benign violation theory* [20], which suggests that a norm violation elicits enjoyment when it is simultaneously appraised as non-threatening or safe.

III. RESEARCH QUESTIONS AND HYPOTHESES

This leads us to the following general research questions: **RQ1.** How does the integration of playful deception as a design element influence the multi-dimensional aspects of user experience in social HRI? **RQ2.** Can the effect of playful deception be controlled by a design parameter?

Identifying a scenario that naturally captures socially acceptable, safely bounded deception in everyday interactions is challenging. To address this, we draw inspiration from the TIC vendor routine, a culturally recognizable human-human interaction that relies on playful misdirection while maintaining clear safety and intent boundaries. The TIC scenario serves as a proxy for assessing how playful deception can enhance the user experience.

Based on this framework, we evaluate the effect of the playful deception and its controllability through the following specific hypotheses within the TIC context:

H1. (Enhancements) TIC-inspired playful deception increases users' enjoyment and engagement.

H2. (Drawbacks) TIC-inspired playful deception leads to a reduction in performance-related trust and perceived safety.

H3. The magnitude of these effects (H1 & H2) are influenced by the interactivity duration of the TIC trick policies.

Thus, the goal of this research is to design and implement a robotic handover system capable of reproducing TIC-inspired deceptive behaviours. Using this system, we explore the potential role of playful deception as a design element in HRI by analyzing its multi-dimensional effects on user experience and examining how the deceptive interaction design parameter, particularly interactivity duration, can influence those effects.

IV. THE TURKISH ICE CREAM ROBOT

A. Overview and Design Objectives

Our goal in this work is to investigate how playful deception inspired by the TIC routine can affect user experience in HRIs. To this end, we have designed a robotic handover system that reproduces key elements of the TIC routine while ensuring user safety and experimental control.

Specifically, we set the following design objectives:

1) Designing a hardware module that ensures safe interaction even in unintended collisions.
2) Capturing and reproducing the characteristic motions of TIC vendors while ensuring the deception remains socially acceptable and playful.
3) Developing multiple trick behaviours derived from observational analyses, allowing controlled manipulation of deceptive strategies and interaction time.

With these objectives, we built a *custom TIC End-effector module* that allows the robot to achieve swift and safe evasive motions (Section IV-B), developed *TIC trick policies* to reproduce key deceptive strategies observed of TIC vendors (Section IV-C), and integrated *implementation details*, including inverse kinematics, hand tracking, and system synchronization into a unified framework (Section IV-D). Figure 2 illustrates an overview of the entire system.

B. Custom TIC End-Effector Module

To replicate the equipment and characteristic gestures of TIC vendors, we developed a custom end-effector module that mounts to the manipulator and extends its reach. The module consists of a *twisting motor*, a *rod*, and a *cone* handover object (Fig. 2). Placing the cone at the rod tip not only reproduces the vendor's tool but also amplifies the manipulator's motion: small joint movements translate into larger, faster displacements, enabling expressive evasive gestures without requiring extreme joint speeds.

1) Twisting Motor: Because the manipulator's wrist joints are too slow and limited in range to reproduce TIC-style tricks, we employ a stepper motor to generate rapid twisting gestures. The motor receives the target rotation angle θ_{twist}, is mounted directly to the manipulator, and is coupled to the rod via a custom connector.

2) Rod: A carbon-fibre rod is coupled to the motor via neodymium magnets, reinforced with masking tape for stability during high-speed motion. This hybrid attachment provides rigidity while enabling safe detachment in the event of unintended contact. Measuring 55 cm, the rod extends the manipulator's reach: subtle wrist rotations produce large, agile displacements at the cone tip, increasing dodging speed without requiring large joint excursions.

3) Cone: The cone is magnetically mounted on the rod tip, enabling easy detachment and retrieval. To mimic the appearance of ice cream, it is divided into a white "ice cream" portion and a yellow "cone" portion. Participants were instructed to grasp only the yellow section, thereby reducing the graspable area, increasing capture difficulty, and reinforcing the playful misdirection in TIC interactions.

C. TIC Trick Policies

The original TIC performance involves a wide range of playful misdirections, including cone switching, conversational distractions, and gaze-based deception. However, our robotic platform is limited in both sensing and actuation modalities, making it infeasible to replicate the full richness of human vendor behaviours. To ensure both feasibility and safety, we focus on a subset of TIC behaviours where deception primarily arises from evasive physical motions.

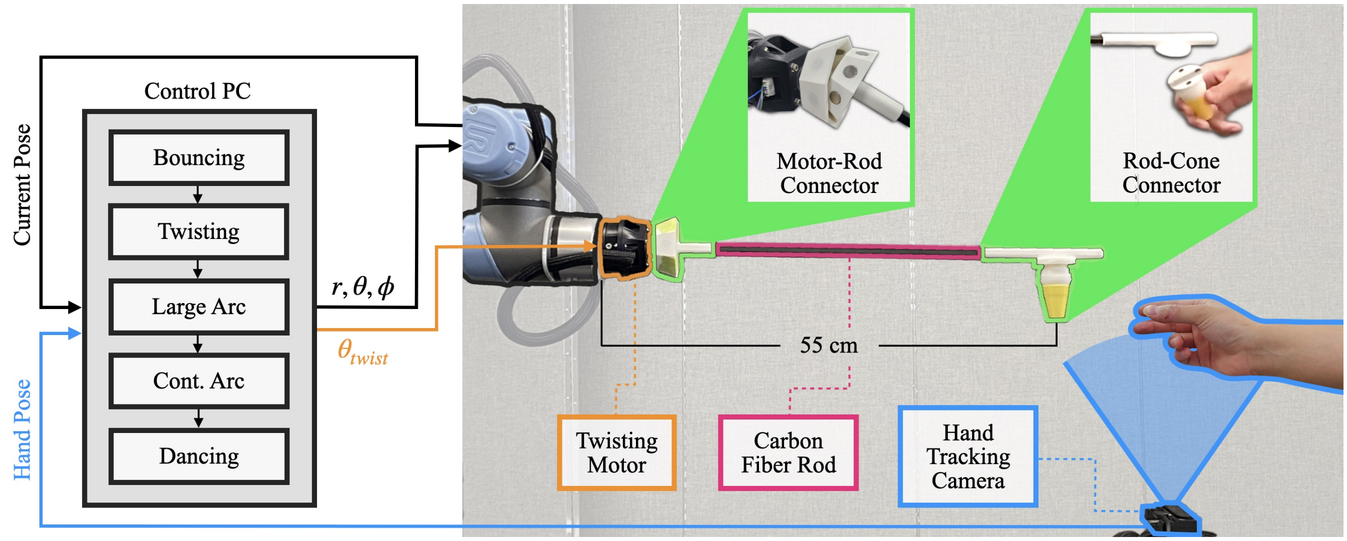

Fig. 2: **System overview.** The system takes the participant's hand pose and the robot's current pose as inputs. Five trick policies are deployed sequentially, each executed for a predefined interactivity duration. Based on the current policy, the controller computes (r, θ, ϕ) for the manipulator and θ_{twist} for the twisting motor. A rod-shaped end-effector amplifies cone motion, while the magnetic connectors (highlighted in green) enable safe and seamless handovers.

1) Extracting Features from Human TIC Interactions: To design these motions, we analyze observational videos of TIC performance online and extract key features that contribute to the playful nature of the tricks. From this analysis, we extracted four recurring features that capture essential aspects of the vendor's movement strategies: *least effort*, *near-miss*, *misdirection*, and *exaggeration*.

Least Effort refers to the vendor's use of minimal physical motion to maintain control of the cone while still preventing the user from grasping it. For example, the vendor may subtly retract the cone by just a few centimetres as the customer reaches forward, using only the wrist. This aligns with Zipf's principle of least effort [21] and often sustains user attention efficiently.

Near-Miss describes situations where the customer's hand comes close to grasping the cone but fails by a small margin. One common example is when the vendor quickly flips the cone upside down just before contact. This near-success amplifies engagement by introducing momentary tension, a mechanism also discussed in gambling theories [22].

Misdirection involves deliberately diverting the customer's attention from the actual movement of the cone. For instance, the vendor may use one hand to draw attention while smoothly transferring the cone to the other hand, exploiting the user's attention elsewhere. This technique, also used in magic [23], helps sustain unpredictability.

Exaggeration involves amplified gestures that emphasize the motion of the cone or the body of the vendor. A typical example includes twirling the cone in large arcs above the user's head. Our intent in realizing this feature is to enhance expressiveness and create a playful rhythm within the interaction.

These four features serve as the conceptual foundation for our robot's behaviour design, guiding both the structure of each trick and the dynamics of timing, effort, and attention within the interaction flow.

2) Trick Behaviour Policies: Based on the extracted TIC trick features, we design five distinct trick-behaviour policies.

TABLE I: Policy-characteristic mapping of implemented TIC tricks.

Policy	**Least Effort**	**Near-Miss**	**Misdirection**	**Exaggeration**
Bouncing	✓	✗	✓	✗
Twisting	✓	✓	✗	✗
Large Arc	✗	✓	✓	✓
Cont. Arc	✗	✗	✓	✗
Dancing	✗	✓	✗	✓

As shown in Tab. I, each policy (rows) has one or more characteristics of the TIC vendor trick (columns). Employing multiple policies introduces variation in the robot's motion, making it more difficult for the user to anticipate the next move. This unpredictability increases the likelihood of failed grasp attempts, thereby sustaining engagement and reinforcing the playful nature of the interaction. Fig. 3 visually describes the motion of each policy.

Bouncing makes the robot move the cone side-to-side in a rhythmic, hopping motion. Each bounce traces a semicircular trajectory, creating the illusion that the cone is "hopping" away. As the arc diameter increases with each bounce, the user must cover more distance to follow. Simultaneously, the arc motion reduces collision risk by discouraging direct linear hand movements toward the target.

Twisting rotates the cone to the opposite side of the approaching hand while the shaft position is fixed at a certain position. This is one of the fundamental policies that real TIC vendors use, in which they bring out customers' large actions while spending little energy.

Large Arc is the discrete action policy that dodges the hand by drawing a large arc path when the hand reaches the cone. The cone narrowly escapes the hand and then draws an exaggerated, large arc. Drawing a large arc has two advantages: it avoids collisions with the hand as the robot moves, and it confuses the user about the cone's destination.

Continuous Arc (Cont. Arc) maintains the robot's position opposite the user's hand along a fixed-center circular trajectory. This movement appears to be a repulsive magnetic force and misdirects the user to follow the cone in a circular

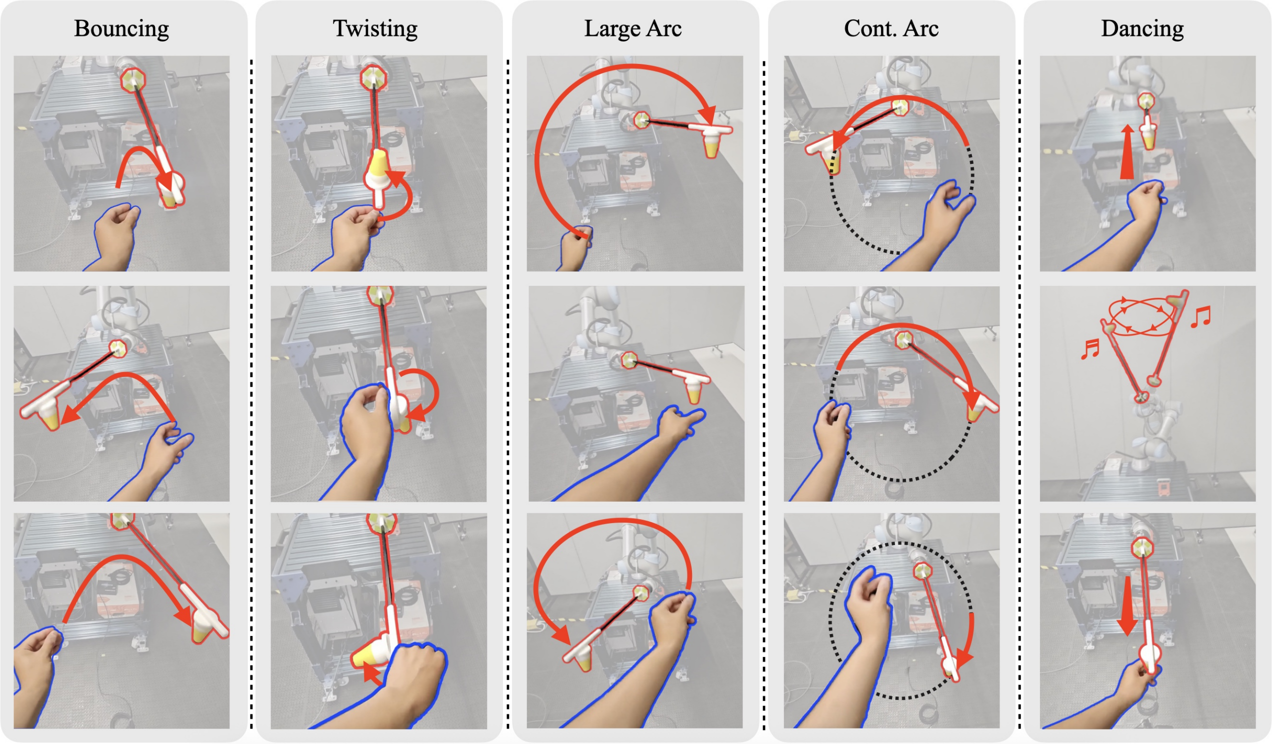

Fig. 3: **Visualization of the five distinct trick motion policies.** Each column represents a different policy, showing the corresponding motion trajectories executed during real-world demonstrations. From top to bottom, each row illustrates the temporal progression of the motion within each policy.

path. In contrast, the most effective user movement for reaching the cone is to go straight to it.

Dancing is a non-interactive policy that is intended to tease/taunt users by circularly waving the rod upright out of the reachable range of the users. The robot moves the cone backward when the hand reaches it, then traces a circle while twisting the cone. At the end of the policy, the robot extends the cone toward the user, finishing the interaction.

D. Implementation Details

1) Robot and Inverse Kinematics: A 6-DOF arm, UR5e manipulator [24], is controlled in a reduced 3-DOF task space anchored to a human-centered frame, parameterized as (r, θ, ϕ): a prismatic degree of freedom r along the forward–back (approach) axis, and two angular degrees of freedom θ (azimuth) and ϕ (elevation) that sweep the end effector laterally and vertically on a sphere of fixed radius. This spherical parameterization simplifies planning without compromising interaction quality, while reflecting real TIC routines where vendors achieve large, expressive motions with subtle wrist movements. Joint commands are computed via inverse kinematics, subject to joint and velocity constraints to ensure smooth, safe execution.

2) Hand Tracking: We use a Leap Motion Controller 2 [25] to track participants' hand pose at 120 Hz. The sensor is positioned below the hand, facing upward, to ensure reliable detection while minimizing collision risk (Fig. 2, right). Participants grasp the cone using a pinching gesture, with the pinch position—defined as the midpoint between the thumb and index fingertip—serving as a target grasp point. To mitigate sensor delay on interaction timing, we apply a velocity-based prediction with a 0.15 s lookahead and a low-pass filter to smooth noisy measurements.

3) System Integration: All sensory and actuation data, including the hand pose, twisting motor pose, and UR5e end-effector pose, are collected and synchronized in real time through the ROS2 framework. The UR5e manipulator is controlled at 200 Hz to ensure smooth and stable handover trajectories, while the twisting motor operates at 50 Hz to reproduce fast and precise rod rotations.

V. USER STUDY

A. Procedure

To evaluate the effects of playful deception and interaction timing in physical handovers, we conducted a mixed-design user study. All participants experienced two handover types: a **Straight Handover** (**SH**, baseline) and a **Deceptive Handover (DH)**, presented in randomized order. Participants were told that the study involved robot-to-human handover, but no mention was made of deception or TIC-inspired interaction.

In the **SH** condition, the robot handed the cone directly to the user without any deceptive motions. The cone was placed slightly to the left or right of the user's midline, depending on which hand the user extended, based on the prior finding that users generally prefer handovers to occur slightly on the receiver's reaching hand's side [26].

The **DH** condition was divided into three sub-conditions based on interactivity duration: *Short*, *Medium*, and *Long*. The robot executed a fixed sequence of five tricks in the following order: *bouncing (4 reps)*, *twisting (3s)*, *large arc (4 reps)*, *cont. arc (3s)*, and *dancing (3s)*. These parameters

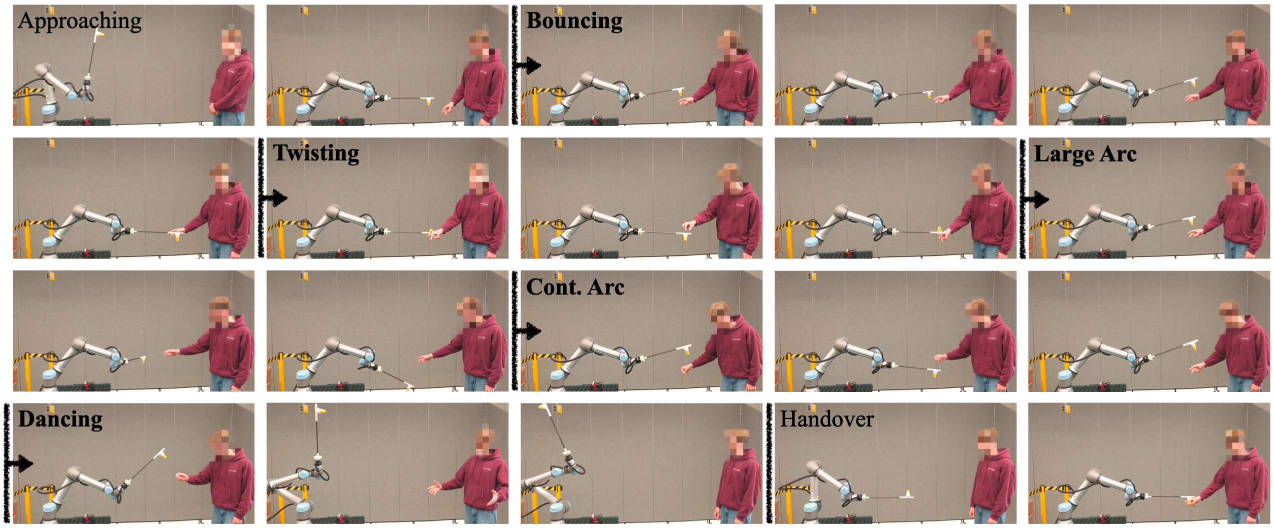

Fig. 4: **User study snapshots of the Deceptive Handover (DH).** The interaction proceeds from top left to bottom right, beginning with approaching (first two frames) and then executing five trick policies sequentially: *Bouncing*, *Twisting*, *Large Arc*, *Cont. Arc*, and *Dancing*. The bold-named frames mark the first image of each policy for clarity. Each policy was performed for its pre-assigned interaction duration. At the end, the cone approaches again to give the cone (last two frames), while the Straight Handover (SH) condition simply performs this final handover without deception. The corresponding motion sequences are available in the supplementary video.

defined the *Medium* condition, which served as the baseline. The *Short* and *Long* conditions were created by scaling all durations and repetitions by factors of 0.5 and 2, respectively, while keeping the behavioural sequence unchanged. This manipulation systematically varied interaction time while maintaining consistency in motion patterns across participants. Fig. 4 illustrates the snapshots of the DH condition in the user study setting.

Participants were randomly assigned to one of the three duration conditions in a between-subjects design, while the comparison between SH and DH was evaluated within-subjects. Each participant took approximately 30 minutes to complete the experiment.

B. Measures

We used a combination of standardized scales and custom questions to assess user perceptions, trust, enjoyment, and qualitative impressions. After each condition, participants completed two 7-point Likert-scale questionnaires (1-strongly disagree, 7-strongly agree): the Multi-Dimensional Measure of Trust (MDMT) v2 [27], assessing reliability, competence, ethics, transparency, and benevolence; and the ENJOY scale [28], measuring pleasure, relatedness, self-competence, challenge, and engagement. Participants also rated perceived deception on a 7-point Likert scale.

Furthermore, drawing on [29], we measured participants' Willingness-to-Pay (WTP) in Canadian dollars (CAD). In the context of the TIC routine, WTP serves as a proxy for the perceived value of the robot's performance, indicating whether the playful deception is recognized as an added source of entertainment.

VI. RESULTS

A total of 91 participants were recruited without compensation through on- and off-campus advertising—one participant's data was excluded due to equipment failure and is not reported here. Participants self-identified as men (n=54), women (n=35), and non-binary (n=1). Participants' ages ranged from 18 to 66 years ($M = 27.9$, $SD = 10.5$). This study was approved by Queen's University General Research Ethics Board (GELEC139-22, File No. 6036728).

A. Analysis

We conducted mixed-design repeated-measures ANOVAs (RM-ANOVAs) to examine changes in perceptions from SH to DH and to assess whether these changes depended on interactivity duration (short, medium, long). **Handover Type (SH vs. DH)** served as the within-subjects factor and **Interactivity Duration** as the between-subjects factor, testing both overall SH–DH differences and their interaction with length. Estimated marginal means with Bonferroni-adjusted comparisons clarify effect directions. To assess interactivity duration independently, one-way ANOVAs were run on DH scores alone. This design was appropriate because each participant experienced one SH and only one DH. Effect sizes are reported as partial eta squared (η^2), with 0.01, 0.06, and 0.14 denoting small, medium, and large effects, respectively [30].

B. Straight vs. Deceptive Handover Main Effects (H1 & H2)

To address H1 & H2, we tested whether perceptions differed between SH and DH. The RM-ANOVA showed a significant overall effect of handover type, $F(12, 76) = 27.31$, $p < .001$, $\eta^2 = .812$.

As summarized in Fig. 5, participants rated DH as more deceptive than SH, confirming the manipulation. DH also increased enjoyment-related outcomes (pleasure, engagement, challenge) and perceived robot competence, but decreased performance trust (reliability), moral trust (ethics, transparency, benevolence), perceived safety, and self-competence. We failed to detect significance in the relatedness measure (ENJOY). Bonferroni-corrected pairwise comparisons indicated these effects were consistent across most measures ($p < .05$). Full results are reported in Tab. II.

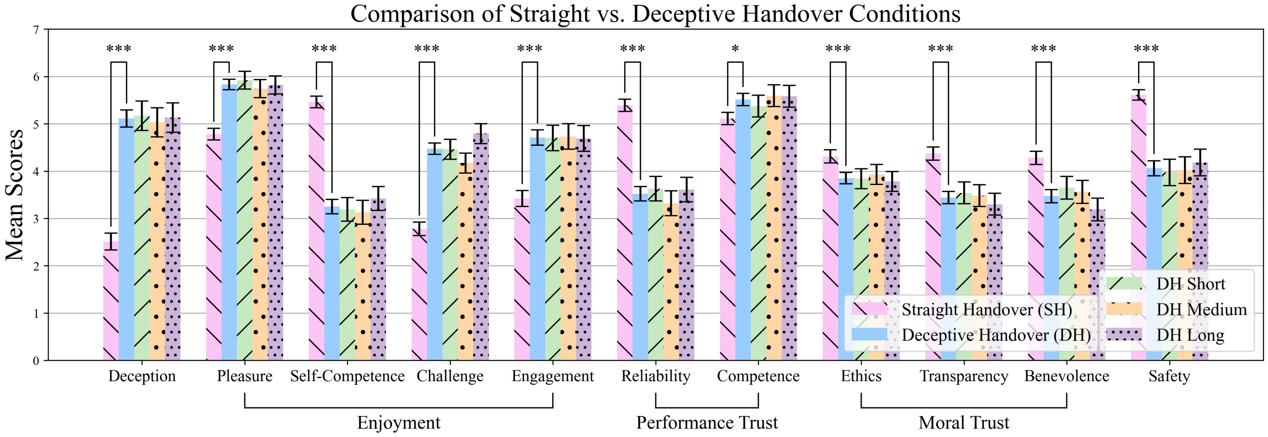

Fig. 5: **User study results comparing Straight (SH) vs. Deceptive Handover (DH).** DH Short, Medium, and Long indicate variations in interactivity duration, and DH represents the combined results across all durations. DH was perceived as more deceptive than SH, while increasing enjoyment-related outcomes (pleasure, challenge, engagement) and competence, but reducing trust, safety, and self-competence. Error bars represent ±Standard Error (SE), and asterisks denote significance levels with Bonferroni correction: $^{*}p < .05$, $^{***}p < .001$.

TABLE II: SH vs. DH ANOVA result details. Significant effects ($p < .05$) are shown with corresponding effect sizes (η^2).

Measure	$F(1, 87)$	p	η^2	**Effect Direction**
Deception				
Deception	92.28	$< .001$	0.52	DH > SH
ENJOY (Enjoyment)				
Pleasure	70.75	$< .001$	0.45	DH > SH
Self-Competence	139.16	$< .001$	0.62	SH > DH
Challenge	110.03	$< .001$	0.56	DH > SH
Engagement	63.59	$< .001$	0.42	DH > SH
MDMT (Performance Trust)				
Reliability	87.25	$< .001$	0.50	SH > DH
Competence	5.82	$< .05$	0.06	DH > SH
MDMT (Moral Trust)				
Ethics	14.77	$< .001$	0.15	SH > DH
Transparency	47.84	$< .001$	0.36	SH > DH
Benevolence	31.93	$< .001$	0.27	SH > DH
Godspeed (Perceived Safety)				
Perceived Safety	86.29	$< .001$	0.50	SH > DH

C. *Effects of Interaction Durations Across in DH (H3)*

To address H3, we examined whether interactivity duration influenced participants' perceptions of deceptive handovers. As shown in Fig. 5, the direction of effects is consistent across measures for SH vs. DH as compared to SH vs. any DH duration. However, their levels of significance do differ: competence and relatedness are non-significant across all durations. Ethics is non-significant for the short duration but significant at $p < .05$ for the long duration. Transparency ($p < .01$) and benevolence ($p < .05$) are significant for the short duration, while pleasure is significant at $p < .01$ for the moderate duration. All other measures and durations not mentioned are significant at the $p < .001$ level. Furthermore, within the DH condition, no significant differences emerged among the three duration groups on any of the 13 measures.

D. *Willingness to Pay*

WTP was reported by participants in CAD, but converted to US dollars here.[1] The RM-ANOVA showed no overall difference between SH and DH. For the short interaction group, WTP did not significantly differ between SH ($M = \$8.06$ USD, $SE = \$2.53$ USD) and DH ($M = \$6.90$ USD, $SE = \$2.31$ USD), $p = .58$. Similarly, for the moderate interaction group, no significant difference was found between SH ($M = \$4.51$ USD, $SE = \$2.39$ USD) and DH ($M = \$5.39$ USD, $SE = \$2.19$ USD), $p = .66$. In the long interaction group, WTP was significantly higher for DH ($M = \$9.87$ USD, $SE = \$2.20$ USD) than SH ($M = \$4.59$ USD, $SE = \$2.04$ USD), $p = .01$. This suggests that while deception does not consistently affect WTP across interaction durations, participants in the longest condition expressed greater willingness to pay. A one-way ANOVA across the three DH duration conditions revealed no significant differences.

E. *Retrieval Success vs. Perceived Safety*

We tested whether participants' success in retrieving the cone during the DH condition influenced perceived safety. Retrieval success (Successful, S vs. Non-Successful, NS) was treated as a between-subjects factor, with perceived safety ratings as the dependent variable. An independent-samples t-test showed no significant difference between S ($M = 3.90$, $SD = 1.36$) and NS ($M = 4.19$, $SD = 1.60$), $t(88) = 0.75$, $p > .05$.

F. *Presentation Order and Sequence Effects*

To test whether the order in which interactions were presented influenced participants' responses, ratings from the first and second interactions were compared for all post-condition measures. A linear mixed-effects model (LMM) was used for this specific check because presentation order was treated as a trial-level covariate rather than a fixed experimental condition. Including the participant as a random intercept accounts for individual differences in baseline responding. This approach allowed us to test whether the sequence of interactions influenced the measures independently of the conditions themselves. We also tested whether the overall order of conditions (SH → DH vs. DH → SH) affected responses.

LMM analysis revealed that reliability (+0.66 points, $p < .001$), ethics (+0.30, $p < .05$), and WTP (+\$3.60 USD, $p < .05$) were significantly higher in the first interaction,

---

[1] We use a CAD → USD exchange rate of 0.72.

while deception ratings were lower (-0.96, $p < .001$). Transparency ($+0.26$, $p = .055$) and relatedness ($+0.26$, $p = .052$) showed near-significant trends. The overall condition sequence influenced deception ratings: participants who experienced SH first rated deception 0.47 points higher than those who experienced DH first ($p < .05$). No other measures showed significant sequence effects.

VII. DISCUSSION

The study investigated how playful, TIC-inspired deception influences user experience in robot handovers. The results support H1 and H2; deceptive handovers increased enjoyment-related measures but reduced multiple dimensions of trust and perceived safety. Specifically, participants rated the DH condition as significantly higher in pleasure, challenge, and engagement (ENJOY subscales) than the SH, and perceived the robot as more competent. At the same time, DH scored significantly lower on performance trust (reliability), moral trust (ethics, transparency, benevolence), perceived safety, and self-competence, indicating that playful deception created a clear trade-off between enjoyment and predictability.

With respect to H3, the results provided limited support. We hypothesized that the design factor, interactivity duration, might moderate the effects of deception. Although the significance levels of several measures differed between SH and DH across interactivity durations, there were no systematic shifts in the overall pattern: the tendency observed in SH vs. DH — greater enjoyment and engagement but lower trust and perceived safety in DH — remained stable across short, medium, and long durations. Moreover, comparisons within the DH conditions themselves revealed no significant differences, suggesting that user perceptions were shaped more by the presence of deceptive behaviour than by its duration.

The WTP results add nuance to these findings: while deception did not affect WTP, participants in the long-interaction group reported significantly higher willingness to pay for the deceptive handover. This suggests that prolonged playful deception may increase the perceived entertainment value of the interaction. However, given the high degree of variability in responses, this effect should be interpreted cautiously.

Notably, participants' perception of safety was not contingent on whether they successfully retrieved the cone. Instead, safety judgments appear to have been shaped more by the robot's deceptive behaviour than by the functional outcome of the interaction. This highlights that, in deceptive HRI scenarios, perceived safety is less about task success and more about the predictability and transparency of robot behaviour. This result underscores the importance of considering how interaction style, not just performance outcomes, shapes user trust and comfort with social robots.

In addition to the primary condition effects discussed above, sequence analyses revealed a novelty effect: participants initially reported higher trust and WTP but shifted toward more critical judgments during the second interaction. This finding aligns with [31], [32]. Notably, this sensitivity was exclusive to trust and norm-related constructs—while enjoyment remained stable—reinforcing the enjoyment-predictability trade-off. Furthermore, SH-first sequences exhibited a contrast effect: the initial predictability established a baseline of reliability that made subsequent misdirection appear more deceptive by comparison.

The findings contribute to three broader insights for HRI. First, they demonstrate that deception is not inherently harmful; when bounded and framed as playful, it can reliably enhance enjoyment and sustain engagement. This aligns with earlier work showing that deceptive strategies can heighten engagement in game-like HRI contexts [12], [16], but extends those insights into embodied, physical handovers, a more everyday form of social interaction.

Second, the results suggest that the value of playful deception is context-dependent. In domains where safety, transparency, and reliability are critical—such as healthcare, industrial collaboration, or transportation—the observed reductions in trust and perceived safety would likely hinder acceptance. In contrast, in entertainment, hospitality, or retail settings, playful misdirection may function as a design feature that enhances engagement and memorability. The impact may also depend on how frequently users encounter the robot; similar to the TIC vendor routine, occasional interactions that briefly frustrate before delivering a reward may create more memorable and socially engaging experiences.

Third, the asymmetry between increased perceived robot competence and decreased user self-competence highlights that competence in HRI is relational rather than absolute. The robot's smooth, controlled tricks were interpreted as skillful, while simultaneously undermining the participants' sense of their own effectiveness. This echoes work on trust violation and repair [10], raising questions about how repeated exposure to playful deception might impact long-term user confidence and willingness to collaborate.

Finally, the results are consistent with benign violation theory [20], breaking expectations in a safe and bounded way produced positive surprise and enjoyment, but also introduced uncertainty about predictability and intent, which reduced trust and safety. For HRI design, at least in carefully structured scenarios, this means playful deception should be deployed selectively so that the benefits of entertainment and engagement do not come at unacceptable costs.

VIII. LIMITATIONS AND FUTURE WORK

This study has several limitations that shape the scope of its conclusions. First, the deceptive behaviours were implemented as a fixed set of justified trick policies. While these captured key features of TIC-style tricks, they may lack the natural timing and variability of adaptive performance, which can be treated as additional design parameters rather than the interactivity duration. Second, the participants experienced only a single deceptive and straight handover in a laboratory setting. As such, the results capture immediate, one-shot impressions rather than long-term dynamics; repeated exposure could either diminish novelty or erode trust more severely. Finally, the measure relied primarily on self-report questionnaires, and although they revealed clear patterns, complementary behavioural and physiological data could provide stronger evidence about the processes underlying engagement, safety perception, and trust.

Taken together, these limitations mean the results can be trusted as reliable evidence of how scripted, short-term playful deception affects immediate perceptions in controlled settings, but they should not be assumed to generalize fully to long-term or applied contexts. Future work should address these limitations by testing a broader range of design parameters, conducting longitudinal studies, and exploring field deployments to further complete our understanding of deception in HRI.

IX. CONCLUSION

This work examined whether playful, TIC-inspired deception alters user experience during robot handovers and whether these effects depend on interaction duration. We compared straight and deceptive handovers across short, moderate, and long interaction conditions. The results show that deceptive handovers consistently increased enjoyment, engagement, and perceived robot competence, while reducing performance trust, moral trust, perceived safety, and user self-competence. These effects remained stable across durations, indicating that the presence of deception—rather than the interaction duration—primarily shaped participants' perceptions. Accordingly, the findings directly address RQ1 by demonstrating how playful deception can function as an HRI design element, while providing only limited evidence that its effects are moderated by interaction timing.

The key takeaway is that playful deception may produce a structured trade-off: it can delight and sustain attention, but at the cost of predictability and trust. This positions deception not as inherently good or bad, but as a contextual design choice. In domains prioritizing entertainment or memorability, bounded misdirection may be valuable. In safety-critical applications, however, the associated declines in trust and safety would likely be unacceptable. These conclusions must be viewed in light of study limitations, scripted behaviours, single-trial laboratory settings, and reliance on self-report. Future work should explore a wider range of design parameters, repeated interactions, and field studies. Overall, our findings highlight deception as a multi-dimensional tool, one that, if carefully bounded, can balance engagement and trust in human-robot interaction.

ACKNOWLEDGMENTS

This work has been funded and supported by the Connected Minds CF-REF and Institute of Information & communications Technology Planning & Evaluation (IITP) under the artificial intelligence star fellowship support program to nurture the best talents (IITP-2026-RS-2025-02304828) grant funded by the Korea government(MSIT).

REFERENCES

[1] J. Danaher, "Robot betrayal: a guide to the ethics of robotic deception," *Ethics Inf. Technol.*, vol. 22, no. 2, pp. 117–128, 2020.
[2] A. Sharkey and N. Sharkey, "We need to talk about deception in social robotics!" *Ethics Inf. Technol.*, vol. 23, no. 3, pp. 309–316, 2021.
[3] H. S. Sætra, "Social robot deception and the culture of trust," *Paladyn, J. Behavioral Robot.*, vol. 12, no. 1, pp. 276–286, 2021.
[4] N. H. Saadon, B. Megidish, and H. Erel, "Scammed by robots: Can deception by two simple non-humanoid robots shape general attitudes toward robots?" in *Int. Conf. Human-Robot Interact.* IEEE, 2025, pp. 929–938.
[5] A. D. Dragan, K. C. Lee, and S. S. Srinivasa, "Legibility and predictability of robot motion," in *Int. Conf. Human-Robot Interact.* IEEE, 2013, pp. 301–308.
[6] A. D. Dragan, R. M. Holladay, and S. S. Srinivasa, "An analysis of deceptive robot motion." in *Robot.: Sci. Syst.*, 2014, p. 10.
[7] M. Chen, S. Nikolaidis, H. Soh, D. Hsu, and S. Srinivasa, "Planning with trust for human-robot collaboration," in *Int. Conf. Human-Robot Interact.*, 2018, pp. 307–315.
[8] A. L. Baker, E. K. Phillips, D. Ullman, and J. R. Keebler, "Toward an understanding of trust repair in human-robot interaction: Current research and future directions," *Trans. Interact. Intell. Syst.*, vol. 8, no. 4, pp. 1–30, 2018.
[9] S. S. Sebo, P. Krishnamurthi, and B. Scassellati, ""i don't believe you": Investigating the effects of robot trust violation and repair," in *Int. Conf. Human-Robot Intract.* IEEE, 2019, pp. 57–65.
[10] K. Rogers, R. J. A. Webber, and A. Howard, "Lying about lying: Examining trust repair strategies after robot deception in a high-stakes hri scenario," in *Int. Conf. Human-Robot Interact.*, 2023, pp. 706–710.
[11] E. Short, J. Hart, M. Vu, and B. Scassellati, "No fair!! an interaction with a cheating robot," in *Int. Conf. Human-Robot Interact.* IEEE, 2010, pp. 219–226.
[12] E. de Oliveira, L. Donadoni, S. Boriero, and A. Bonarini, "Deceptive actions to improve the attribution of rationality to playing robotic agents," *Int. J. Social Robot.*, vol. 13, no. 2, pp. 391–405, 2021.
[13] R. Esposito, A. Rossi, M. Ponticorvo, and S. Rossi, "Roboleaks: Non-strategic cues for leaking deception in social robots," in *Int. Conf. Human-Robot Interact.* IEEE, 2025, pp. 1111–1120.
[14] Youtube, "Gopro awards: Turkish ice cream tricks," 2025. [Online]. Available: https://www.youtube.com/watch?v=Jx_NjpeBqQw
[15] A. R. Wagner and R. C. Arkin, "Acting deceptively: Providing robots with the capacity for deception," *Int. J. Social Robot.*, vol. 3, no. 1, pp. 5–26, 2011.
[16] J. Shim and R. C. Arkin, "A taxonomy of robot deception and its benefits in hri," in *Int. Conf. Syst., Man, and Cybern.* IEEE, 2013, pp. 2328–2335.
[17] R. Esposito, A. Rossi, and S. Rossi, "Deception in hri and its implications: a systematic review," *Trans. Human-Robot Interact.*, vol. 14, no. 3, pp. 1–26, 2025.
[18] J. Shim and R. C. Arkin, "Other-oriented robot deception: How can a robot's deceptive feedback help humans in hri?" in *Int. Conf. Social Robot.* Springer, 2016, pp. 222–232.
[19] G. Forabosco, "Is the concept of incongruity still a useful construct for the advancement of humor research?" *Lodz Papers in Pragmatics*, vol. 4, no. 1, pp. 45–62, 2008.
[20] A. P. McGraw and C. Warren, "Benign violations: Making immoral behavior funny," *Psych. Sci.*, vol. 21, no. 8, pp. 1141–1149, 2010.
[21] G. K. Zipf, *Human behavior and the principle of least effort: An introduction to human ecology*. Ravenio books, 2016.
[22] R. Reid, "The psychology of the near miss," *J. Gambling Behavior*, vol. 2, no. 1, pp. 32–39, 1986.
[23] S. L. Macknik, S. Martinez-Conde, and S. Blakeslee, *Sleights of mind: What the neuroscience of magic reveals about our everyday deceptions*. Macmillan+ ORM, 2010.
[24] Universal Robots, "Ur5e manipulator robot," 2025. [Online]. Available: https://www.universal-robots.com/products/ur5-robot/
[25] Ultraleap, "Leap motion controller 2," 2025. [Online]. Available: https://www.ultraleap.com/products/
[26] P. Basili, M. Huber, T. Brandt, S. Hirche, and S. Glasauer, "Investigating human-human approach and hand-over," in *Human Centered Robot Syst.: Cognition, Interact., Technol.* Springer, 2009, pp. 151–160.
[27] B. F. Malle and D. Ullman, "A multidimensional conception and measure of human-robot trust," in *Trust Human-Robot Interact.* Elsevier, 2021, pp. 3–25.
[28] S. S. Davidson, J. R. Keebler, T. Zhang, B. Chaparro, J. Szalma, and C. M. Frederick, "The development and validation of a universal enjoyment measure: The enjoy scale," *Current Psych.*, vol. 42, no. 21, pp. 17 733–17 745, 2023.
[29] S. Ivanov and C. Webster, "Willingness-to-pay for robot-delivered tourism and hospitality services–an exploratory study," *Int. J. Contemporary Hospitality Management*, vol. 33, no. 11, pp. 3926–3955, 2021.
[30] J. Cohen, *Statistical power analysis for the behavioral sciences*. Routledge, 2013.
[31] I. Leite, C. Martinho, and A. Paiva, "Social robots for long-term interaction: a survey," *Int. J. Social Robot.*, vol. 5, no. 2, pp. 291–308, 2013.
[32] H. Gonzalez-Jimenez and D. Costa Pinto, "Can ai robots foster social inclusion? exploring the role of immersive augmentation in hospitality," *Int. J. Contemporary Hospitality Management*, vol. 36, no. 11, pp. 3889–3905, 2024.