# A Goal-Conditioned Deep Reinforcement Learning Approach to Gaze Modulation for Attention Regulation

Nipuni H. Wijesinghe[1], Maleen Jayasuriya[1], David Hinwood[2], Janie Busby Grant[3], Damith Herath [1]

***Abstract*— Effective human robot interaction (HRI) requires dynamic attention regulation, the ability to both capture and release human attention based on context, yet current robotic systems lack this fundamental social capability. This paper presents a goal conditioned reinforcement learning (GCRL) framework that enables robots to adaptively modulate human gaze engagement and consequently attention across distinct levels through coordinated personalized behavioural control. Unlike prior approaches requiring separate models for each engagement level, this unified GCRL simultaneously learns to adjust multiple robot platform specific behaviours (for example head movements, navigation, gestures, and vocal parameters on a Softbank Pepper robot) based on real time gaze feedback. Central to our approach is the Dynamic Gaze Engagement Index 2.0 (DGEI 2.0), which combines stationary gaze entropy with temporal consistency metrics to differentiate sustained engagement from scattered attention, providing fine grained feedback to calculate appropriate rewards for personalized attention modulation. Controlled experiments with 30 participants demonstrate the framework's effectiveness: gaze engagement precisely tracked target levels (0, 50, 100), while subjective measures revealed significant differences across the intensity conditions ($p < 0.001$, $\eta = 0.599$ for attention/awareness). High intensity behaviours achieved $90\%$ attention agreement rates, while low intensity modes successfully suppressed visual engagement to near zero levels. These results establish a novel validated framework for bidirectional attention regulation in HRI, providing a key enabler for robots to navigate the complex social dynamics of when to engage versus when to fade into the background, a capability essential for long term human robot coexistence.**

## I. Introduction

As robots move from controlled industrial environments into complex social domains such as healthcare, education, and collaborative workspaces, their capacity for adaptive social interaction becomes paramount. Beyond functional competence, these contexts demand robots capable of dynamically modulating social behaviours to accommodate individual differences and situational demands. Central to this challenge is the ability to regulate social presence, a key determinant of acceptance, trust, and interaction quality in agent mediated interactions [1], [2].

Collaborative Robotics Laboratory [1] Faculty of Science and Technology, University of Canberra [2] IA-Cobotics Lab, School of engineering, RMIT, Melbourne [3] Faculty of Health, University of Canberra, Australia



Attention is a foundational antecedent of social presence, yet current robotic systems lack the nuanced attention regulation that underpins natural interaction [3]. Unlike traditional systems with static behavioural patterns, effective social interaction requires the dynamic ability to both capture and release human attention as the context demands. This capacity is critical across applications: in healthcare, robots must command attention during critical communication while remaining unobtrusive in routine care [4]; in education, they must sustain engagement without overwhelming cognitive load [5]; and in collaboration, they must adapt to highly variable attentional thresholds across users.

Adaptive, personalized attention regulation, performed intuitively by humans through subtle behavioural cues, remains absent in current robotic systems. Although prior work has advanced gaze tracking [6], saliency driven control [7], and multimodal engagement strategies [8], these approaches largely assume that robots should continually maximize human attention. Such a unidirectional paradigm overlooks the nuanced dynamics of natural social interaction, where modulation , not mere maximization of engagement is essential for sustaining long term human–robot relationships.

Gaze offers a powerful mechanism for measuring attention. Gaze patterns, though an imperfect proxy of attention, constitute measurable behavioural markers that exhibit significant correlation with attentional allocation [9]. Information theoretic measures such as gaze entropy further enrich this assessment, capturing the complexity of visual attention distribution beyond simple fixation counts [10].

Our previous work has demonstrated the feasibility of gaze based attention modulation using reinforcement learning [11]. However, these early efforts suffered limitations such as separate algorithms were required for different modes (attention seeking versus attention avoiding), the behavioural action space remained restrictive, typically just lighting, volume, and movement, limiting the applicability of the system to different robot platforms.

We address these limitations through a GCRL framework that enables multi level attention modulation within a single policy. By conditioning on engagement levels (low, medium, high), the system autonomously coordinates a broad behavioural repertoire, head movement, navigation, gesture, and vocal modulation, to achieve targeted attentional states.

At the core of this work is an updated DGEI, which integrates gaze frequency with entropy based measures of attentional distribution. By combining Stationary Gaze Entropy with temporal consistency metrics, DGEI 2.0 differentiates scattered glances from sustained engagement, providing fine grained feedback for real time adaptation tailored to individual attentional patterns [10], [12].

The main contributions of this paper are:

- Enhanced Dynamic Gaze Engagement Index (DGEI 2.0): a fine grained, realtime measure of human gaze engagement. Integrates gaze frequency, stationary gaze entropy, and temporal consistency. Provides a more nuanced assessment of visual attention than traditional fixation based metrics.
- A unified GCRL framework for multi level gaze modulation. Enables a single trained policy to autonomously regulate attentional engagement through personalized, coordinated extended behavioural adaptation. Validated in controlled HRI studies with 30 participants.
- Empirical validation of the link between gaze modulation and human attention regulation, demonstrating with standardized measures that gaze modulation directly affects perceived attentional engagement.

This research marks a foundational step toward socially intelligent robots capable of natural, context sensitive attention regulation, an ability essential for authentic social presence. As robots enter human centred environments where attention is both a limited cognitive resource and the currency of interaction, such adaptive capabilities become critical for effective social integration. Paper website `https://collaborativeroboticslab.github.io/agcdrlatgmfar_2026`

## II. Related Work

### A. Social presence and Attention in HRI

The centrality of attention in human cognition fundamentally shapes the landscape of HRI, serving as a foundational antecedent to social presence that ultimately determines acceptance, trust, and interaction quality [2]. Contemporary HRI research operates predominantly within a unidirectional framework that positions robots as persistent attention seekers. This orientation has driven development of sophisticated mechanisms for attention capture and maintenance, encompassing gaze tracking architectures [6], saliency based attention control systems [7], and integrated multimodal strategies that orchestrate speech, kinematic behaviors, and expressive signaling [8]. Within socially assistive contexts, sustained robotic attentiveness demonstrably enhances trust formation and social presence perceptions [4], while educational and companionship applications have successfully deployed attention seeking strategies as engagement catalysts [13].

Yet this persistent focus on attention maximization neglects a fundamental constraint: human attention represents a limited cognitive resource whose excessive taxation precipitates disengagement, frustration, and cognitive fatigue [5]. Authentic social interaction demands flexible attention regulation, the capability to strategically attract, sustain, and crucially, release attentional focus according to contextual demands. This requirement becomes increasingly salient as AI and large language model integration expand robots' conversational sophistication [14].

Our work addresses this critical gap by inverting the traditional perspective: rather than examining how robots perceive human attention, we investigate how robots can strategically shape human attention to achieve desired personalized engagement states. Reinforcement learning offers the best tool for this challenge, enabling unsupervised adaptation through continuous interaction while learning personalized behavioural policies tailored to individual responses, capabilities essential for attention regulation in HRI.

### B. Reinforcement Learning for Behavior Modulation

Reinforcement learning uniquely suits adaptive attention regulation by learning optimal behavioural policies through trial and error without explicit programming. RL's broader robotics applications validate this: adaptive locomotion [15], [16], fault tolerant control [17], and hierarchical behavior coordination [18] demonstrate its real time potential for adaptive context driven robot behaviours. RL can thus be effectively leveraged to enable robots to discover personalized strategies accommodating individual differences in attentional thresholds [19], essential given that optimal behaviors vary across individuals and contexts.

Prior work has successfully established the feasibility of gaze based attention modulation using RL [11]. However, these foundational efforts suffered key limitations, firstly previous implementations faced three main limitations: first, separate Q-learning algorithms were required for different operational modes (attention seeking versus attention avoiding); second, only two extreme operational modes (low and high gaze engagement) were considered; and third, the behavioural action space was restricted to lighting, volume, and movement, where lighting contributed minimally. The current work directly addresses these limitations and significantly improves on these limitations. The effectiveness of RL for personalized attention modulation fundamentally depends on access to reliable, real time feedback that captures the target phenomenon. For attention regulation, gaze patterns provide this critical signal, offering measurable behavioural markers strongly correlated with attentional allocation.

### C. Measuring and Modelling Gaze in HRI

Within HRI contexts, gaze emerges as a primary channel for both assessing and influencing attentional states and by extension, social presence itself. Extensive investigation has documented how robots leverage human gaze patterns to infer cognitive focus, and reciprocally, how robotic gaze allocation signals intent and engagement capacity. These bidirectional gaze dynamics have demonstrated value across diverse social scenarios, with evidence confirming that strategic gaze deployment strengthens trust, rapport, and fundamentally shapes perceptions of the robot's social presence [20], [21].

A wide range of methods have been proposed for measuring and modelling human gaze, each balancing accuracy, robustness, and computational cost. Early approaches, such as gaze guidance degree calculation, segmented visual layouts into regions and quantified their gaze attracting strength, providing insights into attention distribution [22]. Polynomial and neural network models mapped gaze parameters to gaze points, accommodating head movements for robust estimation in dynamic settings [23]. Corneal reflection techniques, using facial images and eyeball geometry, enabled calibration free estimation of 3D gaze vectors with reasonable accuracy [24]. Other studies analysed gaze trajectories and event detection to identify attentional shifts [25]. More recent multimodal systems, such as IntelliGaze AI, integrate gaze tracking with head pose and biometric signals to monitor engagement in online learning contexts [26].

While these approaches focus on gaze estimation accuracy/ similarity, they often overlook deeper measures of attentional complexity or distribution. In contrast, information theoretic measures like gaze entropy capture variability and structure in visual scanning, offering richer insights into attention allocation. Gaze entropy has a long history in marketing and consumer research, predicting purchasing decisions via fixation sequences and area of interest distributions, often combined with dwell time and fixation frequency [10], [27]. Its recent adoption in HRI is promising: both Stationary Gaze Entropy (SGE) and Gaze Transition Entropy (GTE) have assessed cognitive workload, attention shifts, and trust dynamics during robot interactions, achieving predictive accuracies around 70% for task performance [28]. Although challenges remain in defining areas of interest and accounting for individual differences, gaze entropy metrics provide a flexible, interpretable framework for understanding visual attention in complex interactive environments [29]–[31].

Building on these foundations, we propose a unified framework where robots dynamically modulate engagement through gaze entropy driven RL, transitioning attention regulation from passive response to active control. This enables context aware social robots that adaptively calibrate their presence, a fundamental requirement for seamless integration in human environments.

## III. Design and Methodology

In this paper, we present a methodology that employs a gaze conditioned reinforcement learning (GCRL) framework using gaze feedback. Given a target level of eye gaze engagement (low, medium, or high), the GCRL framework adapts the robot's action space to achieve the required engagement (refer Fig. 1). This adaptation is personalized to each individual as it is driven by a gaze based reward function, ensuring that the robot's behaviours are suited to the human's gaze responses. To measure gaze engagement, we use an upgraded DGEI 2.0.

The choice of low, medium, or high gaze levels is context dependent and can be determined via a rule based mechanism/ a context assessment module [32]. Context determination is beyond the scope of this paper; here, we focus on how robot behaviors being optimally adapted to individuals' attentional thresholds through real time gaze driven modulation.

The system's implementation is adaptable to various gaze tracking hardware and robot specific actions based on morphology. For the experiments detailed in Section IV, we employed L2CS Net [33] for gaze data collection and utilized a SoftBank Robotics Pepper robot whose head movement, navigation, gestures, and voice comprised the action space for eliciting required attention dynamics. The GCRL algorithm, trained online using this configuration, was subsequently validated through a human subject study to evaluate the proposed approach (Section IV).

### A. Upgraded Dynamic Gaze Engagement Index (DGEI 2.0)

To quantify participants' attentional engagement toward the robot in real time, we developed an improved DGEI 2.0, which extends earlier formulations [11]. Unlike the original version, which only combined attention ratio and stationary gaze entropy, DGEI 2.0 introduces additional components, *temporal consistency*, *gaze transition entropy*, and *data sufficiency scaling*, yielding a more robust engagement score on a continuous 0-100 scale.

*1) DGEI 2.0 (Gaze Score) Calculation:* The underlying logic assigns high scores when gaze is both frequent and focused (high gaze count (GC), low entropy), reflecting sustained engagement. Conversely, scenarios characterized by high GC with high entropy, or low GC regardless of entropy, are assigned low scores, indicating either dispersed attention or minimal engagement. This formulation ensures that the DGEI captures both the quantity and quality of visual attention in a single, interpretable metric. The method operates in sliding time windows $W$ (For the purpose of the experiments in Section IV, W=3 seconds) and outputs a bounded score at each step:

- **Attention ratio (AR):** proportion of gaze directed at the robot AOIs.
- **Stationary Gaze Entropy (SGE):** dispersion of gaze across AOIs.
- **Gaze Transition Entropy (GTE):** variability in gaze shifts between AOIs.
- **Temporal consistency (C):** stability of continuous fixations on the robot.
- **Data sufficiency scaling (D):** reliability adjustment under sparse gaze samples.

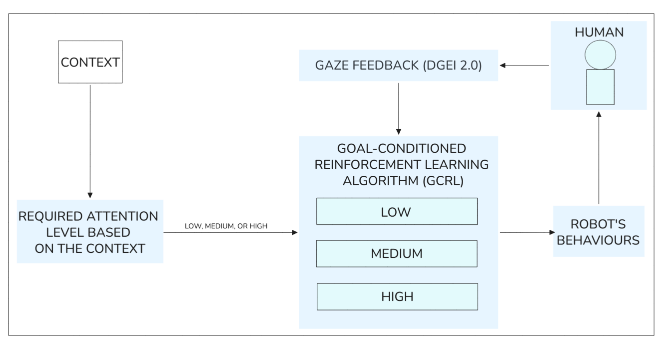


Fig. 1: Overview of the technical methodology

Let the calibrated set of Areas Of Interest (AOIs) be

$$\mathcal{A} = \{a_1, \ldots, a_N\}.$$

The sequence of gaze samples is $\{X_1, \ldots, X_T\}$, with the robot AOI subset denoted as $\mathcal{A}_R \subseteq \mathcal{A}$.

Attention Ratio (AR), Stationary Gaze Entropy (SGE), Gaze Transition Entropy (GTE), Weighted Entropy Fusion (comb), Entropy Normalization(norm), Temporal Consistency (C), and Data Sufficiency Scaling (D) were calculated as follows:

$$\mathrm{AR} = \frac{\#\{X_t \in \mathcal{A}_R\}}{T}, \quad \mathrm{AR} \in [0,1].$$

$$H_{\mathrm{SGE}} = -\sum_{i=1}^{N} p_i \log_2 p_i, \quad p_i = \frac{\#\{X_t = a_i\}}{T}.$$

$$H_{\mathrm{GTE}} = -\sum_{i=1}^{N} \pi_i \sum_{j=1}^{N} P_{ij} \log_2 P_{ij},$$

where $P_{ij} = \Pr(X_{t+1} = a_j \mid X_t = a_i)$ and $\pi_i = p_i$.

$$H_{\mathrm{comb}} = \alpha H_{\mathrm{SGE}} + (1-\alpha) H_{\mathrm{GTE}}, \qquad \alpha \in [0,1].$$

$$H_{\mathrm{norm}} = \frac{H_{\mathrm{comb}}}{\log_2 N} \in [0,1].$$

For foundational information on gaze entropy, readers may refer to its original formulations in prior work on gaze based entropy measures [29], [30].

$$C = 1 - \frac{\text{transitions between states}}{T-1}, \qquad C \in [0,1].$$

$$D = \min\left(1, \frac{T}{T_{\min}}\right).$$

Final Score Fusion:

$$S_{\mathrm{raw}} = w_H(1 - H_{\mathrm{norm}}) + w_A \mathrm{AR} + w_T C + w_D D,$$

with $w_H + w_A + w_T + w_D = 1$.

Rescaled to get the gaze score ranging from 0-100:

$$\mathrm{GazeScore}_t = 100 \cdot S_{\mathrm{raw}}.$$

Temporal Smoothing:

$$\mathrm{GazeScore}_t^{\mathrm{EMA}} = \gamma\, \mathrm{GazeScore}_t + (1-\gamma)\, \mathrm{GazeScore}_{t-1}^{\mathrm{EMA}}, \quad \gamma \in (0,1]. \tag{1}$$

Gaze data in this study were obtained using L2CS Net [33], a multi loss two branch CNN architecture designed for 3D gaze estimation in dynamic environments. Importantly, the proposed framework is not tied to any particular gaze estimation technique and is compatible with data from a range of systems, including gaze tracking glasses and other vision based methods. L2CS Net was selected for this study as it provided higher accuracy compared to prior work using MediaPipe [11]. The DGEI 2.0 serves as a core component of the gaze driven modulation framework.

### *B. Goal Conditioned Deep Q-Network for Gaze Regulation*

*1) GCRL Variables:* We formulate the gaze optimization task as a finite horizon Markov Decision Process (MDP), defined by the tuple $\langle S, A, P, R, \gamma \rangle$, where:

- *State Space ($S$):* The state representation is restricted to a categorical encoding of human gaze engagement levels, discretized into 3 categories: low (0–33%), medium (34–66%), and high (67–100%).
- *Action Space ($A$):*The robot's action space was defined across four behavioural dimensions: head movement (HM; 0–2, where 0 = downward, 1 = intermediate, 2 = direct), navigation positioning (N; 0–4, spanning –0.15 m backward to +0.15 m forward), gestural expression (0–10, where 0 = minimal/no gestures, 10 = exaggerated, rapid, full body gestures; values increase gradually between these extremes), and vocal volume (0–10; where 0 = no volume, 10 = highest volume; values increase gradually between these extremes; utterances ranged from "hi" and "hello there" to more attention-directing phrases such as "look at me"). This space comprises 81 discrete actions ($3^4$), each representing a unique combination of incremental modifications to the four parameters: head movement ($\delta_h$), navigation ($\delta_n$), gesture ($\delta_m$), and vocal volume ($\delta_v$). Each parameter can vary by $-1, 0, +1$ units per time step.
- *Goal Conditioning:* Explicit goal conditioning is incorporated through a one hot encoded goal vector $g \in \{0,1,2\}$, denoting the desired target gaze engagement category. This formulation enables the agent to learn a unified policy capable of pursuing diverse engagement objectives within a single neural architecture.

*2) Reward Function Design:* In contrast to prior implementations that required distinct reward functions for different gaze modulation modes, the proposed approach employs a single unified reward function within the GCRL framework, extending these prior work [11].

Let $G_t$ denote the gaze engagement value at time step $t$, and let $g$ represent the desired attentional goal level selected within the GCRL framework. Each goal $g$ is associated with a bounded target range $[t_{\min}, t_{\max}]$, whose midpoint defines the ideal engagement level, since deviations in either direction are equally undesirable within the target band.:

$$G^* = \frac{t_{\min} + t_{\max}}{2}. \tag{2}$$

We define the distance to the goal at time step $t$ as:

$$D_t = |G_t - G^*|. \tag{3}$$

The reward function evaluates progress toward the goal by comparing the current distance $D_t$ with the previous distance $D_{t-1}$. The reward $R_t$ is defined as:

$$R_t = \begin{cases} 50, & G_t \in [t_{\min}, t_{\max}] \\ \dfrac{|D_{t-1} - D_t|}{2}, & D_t < D_{t-1} \\ -|D_{t-1} - D_t|, & D_t > D_{t-1} \\ -2, & \text{otherwise.} \end{cases} \tag{4}$$

When the gaze engagement lies within the target range, a high positive reward reinforces stability around the desired attentional level. If the system reduces the distance to the goal relative to the previous timestep, a proportional positive reward is assigned to encourage corrective behavior. Conversely, if the distance increases, a penalty proportional to the deviation is applied. A small negative reward is issued when no progress is observed, discouraging stagnation. Consequently, both gaze attraction and gaze moderation behaviors emerge from the same underlying objective function, eliminating the need for mode-specific reward definitions and enabling scalable, unified presence regulation within the GCRL framework.

*3) Online Training:* The GCRL algorithm was trained using an online approach over 7.5 hours with human participants, continuously adapting to real time gaze engagement data. After policy updates, the trained systems were evaluated in a user study with fig. 5 (Section IV).

*4) Actions for Each State at Each Goal:* The system dynamically modulates the robot's behavior based on the user's gaze engagement, classified as low, medium, or high. In low gaze mode (LGM), behavior intensity is reduced to avoid overwhelming the user; in medium gaze mode (MGM), behaviors are maintained at moderate levels to sustain engagement; and in high gaze mode (HGM), behaviors are amplified to reinforce attention.

The specific behaviors to modulate depend on the robot's morphology. In the experiments reported in Section IV, Pepper robot's head movement (HM), navigation (N), gestures (G), and voice (V) were adjusted according to gaze engagement, with actions categorized as Increase (+1), Decrease (-1), or Keep (0) to optimize interaction [34]. Gaze engagement was discretized into three levels, with 0 representing the lowest and 2 the highest level of engagement.

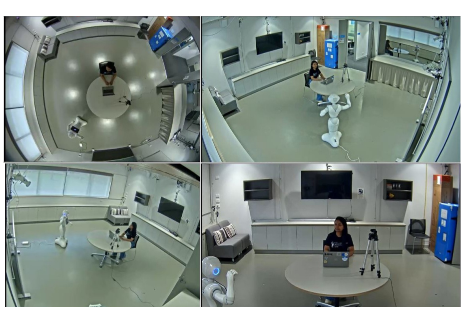

Fig. 2: Experimental Setup for Experiment 01 and 02

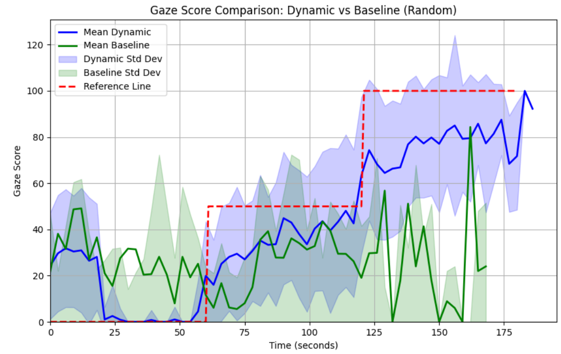


Fig. 3: Mean and standard deviation of DGEI 2.0 values over time during Experiment 01.

## IV. Experimental Results and Discussion

Two experimental studies were conducted to rigorously evaluate the proposed system, involving a cohort of 30 participants and following the experimental framework established in literature [11] (Fig. 2): Experiment 01: Technical Evaluation, and Experiment 02: Gaze Modulation and Attention.

### A. Experiment 01: Technical Evaluation of Gaze Modulation

This experiment assessed the effect of the GCRL on human gaze engagement. In this within subject study, thirty participants experienced two conditions: a *dynamic mode*, implementing the GCRL to cycle through LGM, MGM, and HGM (60 seconds each), and a *baseline mode*, where robot was in it's autonomous life mode, behaviors were random and non adaptive. Each six minute session was recorded with DGEI 2.0 as participants watched a six minute video, seated slightly to the side but facing the robot (see Fig. 2).

The reference engagement pattern was set at 0 (LGM), 50 (MGM), and 100 (HGM) across successive intervals. Observed gaze engagement in dynamic mode closely followed this trajectory (Fig. 3). Each shift produced the expected change in gaze, with an initial adjustment followed by stable engagement at the new level. Notably, medium levels of gaze were reliably sustained during the MGM phase, suggesting that the robot could modulate attention without simply driving participants to extremes.

An additional observation was an initial spike in gaze engagement during the first 10–20 seconds of each transition. This reflects participants' heightened attention as the robot shifted from its previous set of action levels into the new behavioural mode. Importantly, after this spike, gaze engagement stabilised at the intended reference level, underscoring both the responsiveness of participants and the ability of the GCRL to regulate attention consistently over time.

In contrast, gaze engagement during the baseline condition was erratic underscoring the effectiveness of adaptive modulation in regulating attention. Overall, Experiment 01 demonstrates that the GCRL framework can dynamically adjust to individual gaze feedback. The DGEI 2.0 metric provided a reliable basis for reward shaping, validating its

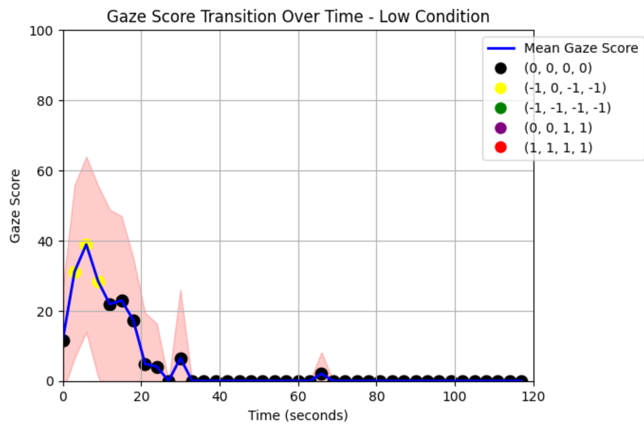


Fig. 4: Mean and standard deviation of the DGEIs against time during LGM. The four digit vectors are action deltas.

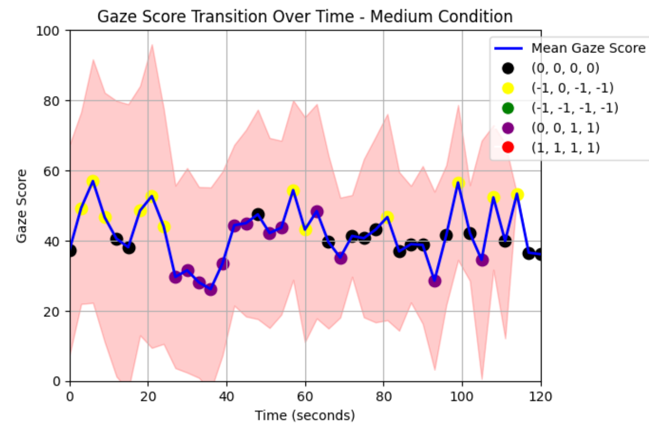


Fig. 6: Standard deviation, mean of the DGEIs against time during MGM. The four digit vectors are action deltas.

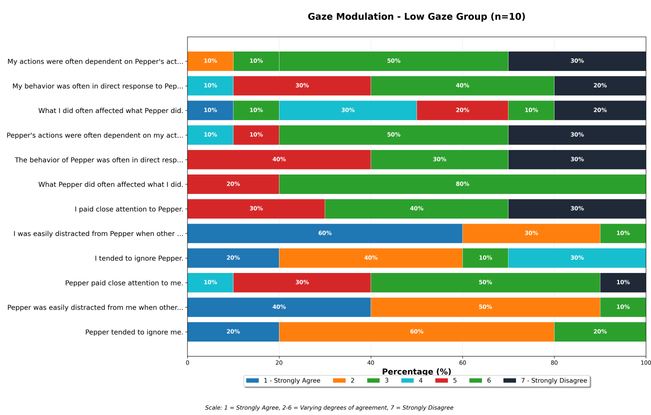


Fig. 5: LGM Questionnaire Results

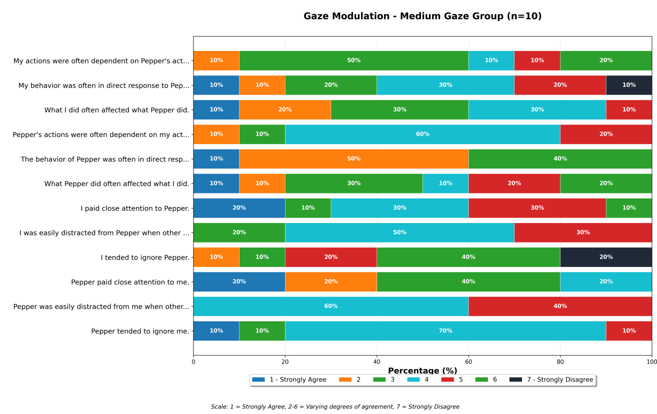


Fig. 7: MGM Questionnaire Results

role in guiding gaze responsive behaviors and supporting the proposed evaluation framework.

### *B. Experiment 02: Investigating the Relationship between Gaze Modulation and Attention*

The second experiment adopted a between subjects design, with participants randomly assigned to one of three conditions: LGM, MGM, or HGM. Each participant watched a five minute video, which served as a neutral activity to focus attention. The robot remained inactive during the first minute, before engaging in the assigned behavioural module for the following three minutes. Real time DGEI 2.0 data were continuously recorded throughout the session. Following the interaction, participants completed a brief questionnaire based on the Perceived Attentional Engagement dimension of the Networked Minds Social Presence Inventory (NM-SPI) [3]. This SP measure was rated on a 7-point Likert scale, with selected items specifically targeting attentional engagement. In total, 30 participants took part in this study, with 10 allocated to each condition.

*1) Results and Discussion:* In the LGM condition, 10 participants (4 females, 6 males) were tested. Following a brief transition phase, gaze engagement remained at zero throughout the session, indicating complete visual disengagement from Pepper. All participants focused exclusively on the video, with no gaze toward the robot. As shown in Fig. 4, the engagement trace aligns with the reference line, confirming that LGM suppressed visual attention. The plotted action set frequencies further reveal that under low gaze conditions, the robot predominantly maintains its current behaviors, in contrast to reduced output under high engagement.

Post-experiment responses from the LGM condition (see Fig. 5) indicate a low sense of interactivity with Pepper. Most participants selected neutral or disagree responses on items assessing mutual responsiveness. Half expressed neutrality regarding behavioural dependence on Pepper, while 30% disagreed. Similarly, 40% were neutral and up to 30% disagreed with statements about Pepper's responsiveness. The sole item with unanimous agreement was "What Pepper did often affected what I did," suggesting unidirectional influence rather than reciprocal engagement. These results imply that LGM effectively suppressed attentional engagement.

In the MGM condition, 10 participants (4 females, 6 males) were tested. As illustrated in Fig. 6, gaze engagement remained at a moderate level throughout the 3 minute session. Fluctuations likely reflect natural shifts in attention, with participants initially focused on Pepper, then momentarily attending to the video before re-engaging. The

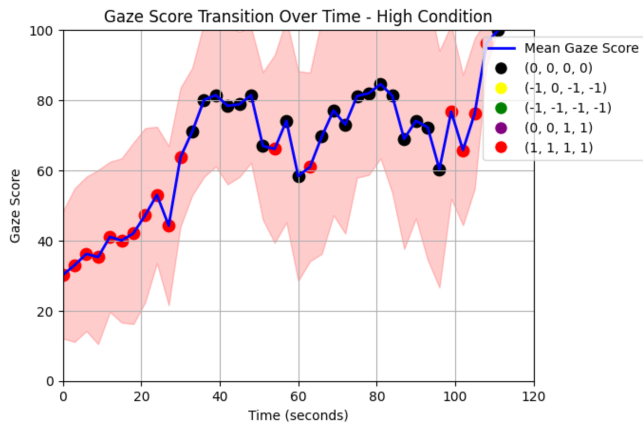


Fig. 8: Standard deviation, mean of the DGEIs against time during HGM

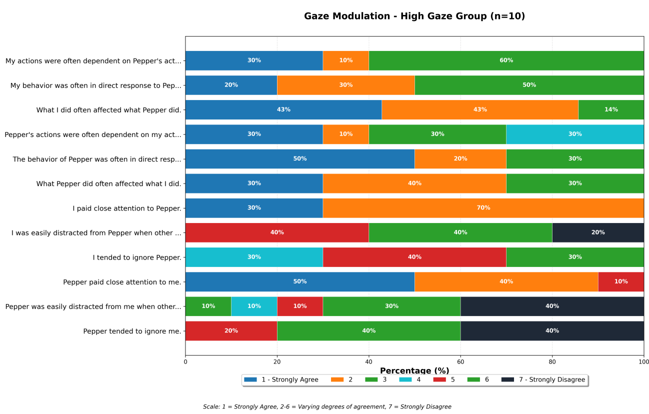


Fig. 9: High Mode Questionnaire Results

plotted action set frequencies show that under medium gaze conditions, the robot primarily maintains its existing behaviors, contrasting with increased activity under low gaze and reduced output under high gaze modulation.

Post-experiment responses from the MGM condition (see Fig. 7) reflect moderate attentional engagement and perceived interactivity with Pepper. 60% of participants agreed they paid close attention, while only 20% reported distraction, suggesting balanced engagement. Half disagreed with the statement "I tended to ignore Pepper," indicating that Pepper remained salient yet unobtrusive. Interaction related items showed 60–70% agreement on mutual influence, though neutral responses persisted, implying that reciprocity was perceived but not uniformly strong. Overall, MGM modulation sustained attention and interaction without inducing overstimulation, positioning it as a balanced strategy for human–robot engagement.

In the HGM condition, 10 participants (4 females, 6 males) were observed. As shown in Fig. 8, the interaction began with a brief transition phase, followed by a sustained period of elevated gaze engagement throughout the three minute session. Minor fluctuations likely reflect natural shifts in attentional focus, with participants intermittently attending to the video before re-engaging with Pepper. The plotted action set frequencies indicate that under high gaze conditions, the robot primarily maintains its current behaviors, whereas under low and medium conditions, it increases behavioural activity to modulate engagement.

Participants in the HGM condition (see Fig. 9) reported strong attentional engagement and heightened interactivity with Pepper. 90% agreed they paid close attention, while only 10% reported distraction. None indicated ignoring Pepper, and over 80% agreed that Pepper attended to them. Interaction related items further supported this effect, with 70–90% agreement on mutual influence. For instance, 80% agreed that "Pepper's actions were often dependent on my actions," and 90% affirmed that "What Pepper did often affected what I did." These findings suggest that HGM effectively amplified attentional focus and fostered a robust sense of bidirectional engagement, enhancing perceived social presence.

2) *Statistical Analysis:* Questionnaire data were analyzed to assess differences across three experimental conditions (LGM, MGM, HGM; $n = 10$ each). Descriptive statistics revealed distinct patterns. Attention/Awareness scores were $M = 3.22$ (LGM), $M = 3.98$ (MGM), and $M = 4.38$ (HGM); behavioural Interdependence scores were $M = 5.60$ (LGM), $M = 3.45$ (MGM), and $M = 2.15$ (HGM); Total NMPSI scores were $M = 4.41$ (LGM), $M = 3.72$ (MGM), and $M = 3.29$ (HGM).

Shapiro–Wilk tests indicated approximate normality, with minor deviations for High Attention ($W = 0.788$, $p = 0.010$) and Low Interdependence ($W = 0.839$, $p = 0.043$). Given ANOVA's robustness to such deviations, one way ANOVAs were conducted and revealed significant group differences for Attention/Awareness ($F(2, 27) = 20.203$, $p < .001$, $\eta^2 = 0.599$), behavioural Interdependence ($F(2, 27) = 55.862$, $p < .001$, $\eta^2 = 0.805$), and Total NMPSI ($F(2, 27) = 13.409$, $p < .001$, $\eta^2 = 0.498$).

Due to normality violations, Mann–Whitney $U$ tests were used for post hoc comparisons. All pairwise contrasts were significant for Attention/Awareness and behavioural Interdependence. For Total NMPSI, Low vs. Medium ($U = 83.0$, $p = 0.014$) and Low vs. High ($U = 94.5$, $p = 0.001$) were significant; Medium vs. High was not ($U = 70.0$, $p = 0.139$). Cronbach's alpha confirmed excellent reliability: $\alpha = 0.934$ (Attention/Awareness), $\alpha = 0.933$ (Interdependence). Overall, adaptive gaze modulation significantly shaped perceptions of attention and interdependence, though in opposite directions. High intensity behaviors enhanced attentional engagement, while low intensity behaviors promoted interdependence and yielded the highest overall interaction quality. These findings highlight the need to balance attentional capture with social reciprocity.

Experiment 02 further validated the GCRL based gaze modulation framework, demonstrating its direct influence on human engagement and confirming its effectiveness in shaping perceived attention in real world HRI scenarios.

## V. Conclusion

This work introduced a unified GCRL framework for real time, gaze driven behavioural adaptation in social robots. By conditioning a single Deep Q-Network on target engagement levels, the system overcomes limitations of separately trained models while enabling scalable adaptation across diverse robot morphologies, demonstrated through the Pepper robot implementation incorporating head movement, navigation, gesture, and vocal modulation.

Empirical validation with 30 participants confirmed the framework's effectiveness. Experiment 01 showed adaptive modulation precisely tracked target engagement levels, outperforming non adaptive baselines. This validates the reward feedback provided by our Dynamic Gaze Engagement Index 2.0 (DGEI 2.0), which combines stationary gaze entropy with temporal consistency metrics to differentiate sustained engagement from scattered attention. Experiment 02 demonstrated that gaze modulation directly shaped perceived attention: high intensity behaviours achieved 90% attention agreement rates while low intensity modes suppressed engagement to near zero. Comparison against the separated Q-learning models of prior work or rule-based heuristics is left for future work.

These findings establish GCRL based gaze modulation as an effective mechanism for bidirectional attention regulation in HRI. By enabling robots to both capture and release human attention as context demands, the framework advances beyond traditional attention maximization paradigms. Future work will incorporate context assessment and evaluate against heuristic finite state machine baselines. Refer video: `https://youtu.be/f4WwvZE8WU8`.


## ACKNOWLEDGMENT

This research was supported by an Australian Government Research Training Program (RTP) Scholarship `https://doi.org/10.82133/C42F-K220`


## REFERENCES


[1] V. Ngo, "Humanizing ai for trust: the critical role of social presence in adoption," *AI SOCIETY*, 2025.

[2] N. H. Wijesinghe, Y. Peng, S. Konrad, M. Jayasuriya, J. B. Grant, and D. Herath, "Reframing social presence in human-robot interaction," in *International Conference on Human–Robot Interaction*, 2025.

[3] F. Biocca, C. Harms, and J. Gregg, "The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity," in *4th annual international workshop on presence, Philadelphia, PA*, pp. 1–9, 2001.

[4] J. Bravo, J. Cacace, D. Dal Moro, D. Serrano, and M. Dalmau-Moreno, "Improving robot social perception in human-robot-interaction using multi-modal cues," in *2024 7th Iberian Robotics Conference (ROBOT)*, pp. 1–6, IEEE, 2024.

[5] M. Han, Y. Wang, X. Liu, X. Cheng, H. Niu, and T. Liu, "Attention demands modulate brain electrical microstates and mental fatigue induced by simulated flight tasks," *Journal of Neural Engineering*, vol. 21, no. 6, p. 066024, 2024.

[6] R. Ramnauth, F. Shic, and B. Scassellati, "Gaze behavior during a long-term, in-home, social robot intervention for children with asd," *arXiv preprint arXiv:2501.02583*, 2025.

[7] M. Staffa and P. Barra, "Multi-monitor system for adaptive image saliency detection based on attentive mechanisms," in *International Conference on Human-Computer Interaction*, pp. 607–617, Springer, 2023.

[8] T. Schreiter, L. Morillo-Mendez, R. T. Chadalavada, A. Rudenko, E. Billing, M. Magnusson, K. O. Arras, and A. J. Lilienthal, "Advantages of multimodal versus verbal-only robot-to-human communication with an anthropomorphic robotic mock driver," in *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pp. 293–300, IEEE, 2023.

[9] M. Lombardi, E. Maiettini, A. Wykowska, and L. Natale, "Gaze estimation learning architecture as support to affective, social and cognitive studies in natural human-robot interaction," *arXiv preprint arXiv:2410.19374*, 2024.

[10] M. Wedel, R. Pieters, and R. van der Lans, "Eye tracking methodology for research in consumer psychology," in *Consumer Psychology*, Taylor & Francis, 2023.

[11] N. H. Wijesinghe, M. Jayasuriya, D. Hinwood, J. B. Grant, and D. Herath, "Adaptive gaze modulation in social robots: A reinforcement learning approach to attention regulation," in *2025 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2025.

[12] B. K. Behe, P. T. Huddleston, and I. S. Muraro, "Seeing through the forest: The gaze path to purchase," *PLoS ONE*, 2020.

[13] E. Torta, J. van Heumen, R. H. Cuijpers, and J. F. Juola, "How can a robot attract the attention of its human partner? a comparative study over different modalities for attracting attention," in *Social Robotics: 4th International Conference, ICSR 2012, Chengdu, China, October 29-31, 2012. Proceedings 4*, pp. 288–297, Springer, 2012.

[14] S. Stankevich and W. Dudek, "Interpreting and learning voice commands with a large language model for a robot system," *arXiv preprint arXiv:2407.21512*, 2024.

[15] T. A. Fahmy and S. A. Maged, "Teaching quadruped to walk using fault adaptive deep reinforcement learning algorithm," in *2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)*, pp. 129–134, IEEE, 2021.

[16] D. Waidner and M. Strand, "Learning and transfer of movement gaits using reinforcement learning," in *Climbing and Walking Robots Conference*, pp. 383–395, Springer, 2021.

[17] R. Evans, "Reinforcement learning in a neurally controlled robot using dopamine modulated stdp," *arXiv preprint arXiv:1502.06096*, 2015.

[18] M. G. Azar, M. N. Ahmadabadi, A. M. Farahmand, and B. N. Araabi, "Learning to coordinate behaviors in soft behavior-based systems using reinforcement learning," in *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, pp. 241–248, IEEE, 2006.

[19] A. Gonzalez-Santocildes, J.-I. Vazquez, and A. Eguiluz, "Adaptive robot behavior based on human comfort using reinforcement learning," *IEEE Access*, 2024.

[20] L. Cavedon, C. Kroos, D. Herath, D. Burnham, L. Bishop, Y. Leung, and C. J. Stevens, "'c' mon dude!': Users adapt their behaviour to a robotic agent with an attention model," *International Journal of Human-Computer Studies*, vol. 80, pp. 14–23, 2015.

[21] S.-R. Lu, J.-H. Lo, Y.-T. Hong, and H.-P. Huang, "Implementation of engagement detection for human–robot interaction in complex environments," *Sensors*, vol. 24, no. 11, p. 3311, 2024.

[22] N. Kayahara and S. Yamada, "Gaze guidance degree calculation system, gaze guidance degree calculation program, storage medium, and gaze guidance degree calculation method," Apr. 26 2011. US Patent 7,931,602.

[23] C. Jian-nan, Z. Chuang, Y. Yan-tao, L. Yang, and Z. Han, "Eye gaze calculation based on nonlinear polynomial and generalized regression neural network," in *2009 Fifth International Conference on Natural Computation*, vol. 3, pp. 617–623, IEEE, 2009.

[24] S. Kojima and S. OSUGA, "Gaze-tracking device, computable readable medium, and method," Sept. 22 2020. US Patent 10,783,663.

[25] K. Sakata, S. Maeda, R. Yonetani, H. Kawashima, T. Hirayama, and T. Matsuyama, "Gaze target determination device and gaze target determination method," Mar. 25 2014. US Patent 8,678,589.

[26] A. Donkada and W. Kolluru, "Intelligaze ai: An advanced real-time multimodal system for adaptive attention and integrity monitoring in digital learning environments using integrated gaze dynamics, head pose estimation, and customizable biometric scoring algorithms," *preprint*, 10 2024.

[27] Z. Liu, W.-C. Yeh, K.-Y. Lin, C.-S. H. Lin, and C.-Y. Chang, "Machine learning based approach for exploring online shopping behavior and preferences with eye tracking," *Computer Science and Information Systems*, vol. 21, no. 2, pp. 593–623, 2024.

[28] Y. Zhang, A. Yadav, S. K. Hopko, and R. K. Mehta, "In gaze we trust: Comparing eye tracking, self-report, and physiological indicators of dynamic trust during HRI," in *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, 2024.

[29] K. Krejtz, T. Szmidt, A. T. Duchowski, and I. Krejtz, "Entropy-based statistical analysis of eye movement transitions," in *Proceedings of the Symposium on Eye Tracking Research and Applications*, ETRA

'14, (New York, NY, USA), p. 159–166, Association for Computing Machinery, 2014.

[30] K. Krejtz, A. Duchowski, T. Szmidt, I. Krejtz, F. González Perilli, A. Pires, A. Vilaro, and N. Villalobos, "Gaze transition entropy," *ACM Trans. Appl. Percept.*, vol. 13, Dec. 2015.

[31] M. Jayasuriya, G. Hu, D. D. K. Le, K. Ang, S. Sankaran, and D. Liu, "A 3d vector field and gaze data fusion framework for hand motion intention prediction in human-robot collaboration," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 5637–5643, IEEE, 2024.

[32] N. H. Wijesinghe, M. Jayasuriya, J. B. Grant, and D. Herath, "Real-time social presence modulation of embodied ai-based robots: An audio-centric approach," in *2025 IEEE 21st International Conference on Automation Science and Engineering*, 2025.

[33] A. A. Abdelrahman, T. Hempel, A. Khalifa, A. Al-Hamadi, and L. Dinges, "L2cs-net : Fine-grained gaze estimation in unconstrained environments," in *2023 8th International Conference on Frontiers of Signal Processing (ICFSP)*, pp. 98–102, 2023.

[34] A. K. Pandey and R. Gelin, "A mass-produced sociable humanoid robot: Pepper: The first machine of its kind," *IEEE Robotics & Automation Magazine*, vol. 25, no. 3, pp. 40–48, 2018.