



Doctoral Thesis

**Implicit Influence on Low-Moral Behaviors
Through Social Robots**

Kanghui Du

September 2025

Department of Social Informatics
Graduate School of Informatics
Kyoto University

Doctoral Thesis
submitted to Department of Social Informatics,
Graduate School of Informatics,
Kyoto University
in partial fulfillment of the requirements for the degree of
DOCTOR of INFORMATICS

Thesis Committee: Takayuki Kanda, Professor
Michinori Hatayama, Professor
Takayuki Ito, Professor

Implicit Influence on Low-Moral Behaviors Through Social Robots*

Kanghui Du

Abstract

As robots are increasingly deployed in public environments, they will be increasingly deployed to assist with everyday tasks, including navigation, customer support, and social regulation. While much attention has been given to robots’ ability to provide friendly and functional services, their role in influencing and preventing low-moral behaviors (e.g., obstruction, harassment, or norm violations) remains underexplored. Without the ability to discourage such behaviors, robots may fail to gain acceptance or perform effectively in complex real-world settings.

This thesis addresses this gap by investigating how robots can influence low-moral behaviors in public spaces without provoking negative reactions. We propose that subtle, human-inspired robot behaviors—rather than overt commands—can nudge people toward socially appropriate conduct. Through a series of field studies and cross-cultural experiments, we develop and evaluate models of “suggestive avoidance,” legible motion signaling, and subtle social feedback that align with human expectations while respecting social norms.

Moreover, we identify that robots themselves may unintentionally exhibit behaviors perceived as low-moral, such as stopping abruptly in shared walking paths, causing discomfort or disruption. To address this, we design behavior-corrective strategies that communicate the robot’s intention, such as legible stopping cues, thereby reducing misinterpretations and fostering trust.

In our first study (Chapter 3), we investigate how robots can learn to recognize low-moral actions in public spaces despite the rarity of such events. Because these behaviors—such as deliberately blocking a robot—occur infrequently, collecting large-scale training data is impractical. To address this challenge, we introduce

*Doctoral Thesis, Department of Social Informatics, Graduate School of Informatics, Kyoto University, KU-I-DT6960-33-7210, September 2025.

a novel dataset consisting primarily of rare low-moral actions captured through depth sensors. This dataset enables the exploration of low-data learning scenarios in moral behavior recognition. We evaluate the performance of multiple classifiers using two different data modalities: raw depth maps and extracted human skeletons. Experimental results show that both approaches achieve comparable accuracy, with Top-1 accuracy around 55% and Top-5 accuracy near 90%. Furthermore, we demonstrate that transfer learning significantly boosts recognition performance for both data types. This study highlights the potential of using compact and efficiently annotated datasets to train moral behavior recognition systems, forming the basis for downstream robot behavior modulation.

In our second study (Chapter 4), we explore how robots can subtly influence low-moral behavior using suggestive avoidance—a human-inspired strategy in which the robot does not confront the person directly, but instead expresses discomfort through bodily orientation and movement. Drawing on examples of human avoidance, we design a navigation behavior that implicitly communicates that the robot is being bothered. Through a field deployment and video-coded behavioral analysis, we demonstrate that suggestive avoidance significantly reduces repeated obstruction and is perceived as more humanlike compared to direct avoidance.

In our third study (Chapter 5), we examine the generalizability of the suggestive avoidance model by replicating it in a different country. Conducting a cross-cultural field experiment in a Spanish university, we investigate whether the same avoidance behavior elicits similar responses and perceptions in a different sociocultural context. Our analysis reveals that suggestive avoidance still effectively works in a different environmental setting.

In our fourth study (Chapter 6), we shift focus to a different but related problem: when robots themselves exhibit behavior that causes discomfort—specifically, sudden stops in the presence of trailing pedestrians. We observe that such stops, when unannounced, often confuse people. To resolve this, we propose legible stopping behaviors, including gradual deceleration, body orientation, and gaze cues that signal the robot’s behavior change. Through experimental comparisons, we show that these legible stopping cues improve predictability, reduce human-robot conflict, and are evaluated positively by observers. This study underscores the importance of robot self-presentation and transparency in avoiding unintentional low-moral signaling.

By focusing on socially appropriate interventions and mutual accountability, this thesis presents a framework for designing robots that not only coexist with the public but also participate in shaping socially normative environments. These insights advance the development of robots capable of engaging with human morality, improving acceptance, safety, and harmony in shared public spaces.

Keywords: Social Robotics, Human-robot Interaction, Behavior Detection, Social Navigation, Legible motion, Cross-cultural Study