Psychological Perspectives in Human–Robot Interaction

Mayiana Mitevska, Prof. D.Sc.
Plovdiv University "Paisii Hilendarski", Faculty of Pedagogy, Department of Psychology, Bulgaria, Plovdiv.

Emilia Tsvetanova Lazarova, PhD
Deputy Minister of Education and Science

Paulina Tsvetkova, PhD
Associate Professor, Institute of Robotics, Bulgarian Academy of Sciences

DOI: https://doi.org/10.56293/IJMSSSR.2025.5713

IJMSSSR 2025
VOLUME 7
ISSUE 4 JULY – AUGUST **ISSN: 2582 – 0265**

Abstract: Human–robot interaction (HRI) is a complex and dynamic process that engages a wide array of psychological mechanisms, including social communication, trust formation, emotional resonance, and cultural adaptation. This paper outlines five key theoretical models that offer in-depth perspectives on the psychological foundations of social robotics: Breazeal's (2003) model of socially motivated interaction, Hancock et al.'s (2011) trust model in HRI, Bartneck et al.'s (2007) model of anthropomorphism and user expectations, Li's (2015) cultural mediation model, and Mutlu et al.'s (2009) model of joint meaning-making. By analyzing these models, the study reveals the multifaceted nature of human–robot dynamics and their practical implications in fields like education and therapy. The comparative analysis shows how these frameworks complement each other in guiding the development and implementation of socially intelligent, emotionally engaging, and culturally adaptive robotic systems. Special attention is given to the integration of these models in academic contexts, highlighting their role in enhancing student engagement, trust in educational technologies, and intercultural inclusivity.

Keywords: Human–Robot Interaction (HRI), social robotics, trust, anthropomorphism, cultural mediation, joint meaning-making, user expectations, psychological models, social engagement, adaptive behavior.

Introduction

Interaction between Human and Social Robot

The interaction between a human and a social robot is a complex process based on various psychological mechanisms, including communication, trust formation, and individual differences (Fong et al., 2003; Breazeal, 2003). To better clarify this process, it is useful to examine several key theoretical models that provide a framework for understanding the psychological mechanisms in human–robot interaction.

Model of Socially Motivated Interaction (Breazeal, 2003)

Breazeal (2003) proposes a model emphasizing socially motivated interaction. This model examines how social signals—such as gaze, gestures, and emotional expression—facilitate interaction with robots, making it more natural and predictable. According to her, social engagement is a bidirectional process in which the robot must appropriately respond to human social signals and generate its own to maintain a "social bond" with the user.

In her foundational work Designing Sociable Robots, Cynthia Breazeal (2003) presents a socially motivated interaction model serving as a conceptual framework for developing robots capable of intuitive, contextually relevant, and emotionally sensitive communication with humans. The model underscores the importance of social signals—gaze, facial expressions, tone of voice, body posture, and gestures—as mediators of interaction, making it more natural, predictable, and sustainable over time.

Breazeal argues that social engagement is a bidirectional process, requiring robots not only to respond adequately to human behavioral signals but also to generate socially meaningful responses themselves to maintain mutual engagement and dialogic interaction. In this sense, social behavior is not a secondary element but an essential characteristic of their ability to collaborate effectively with humans:

"To maintain a meaningful interaction with a human, a robot must both perceive and act in a socially appropriate manner. This includes attending to a person, understanding their affective cues, and responding in ways that are perceived as contingent and socially competent."

(Breazeal, 2003, p. 167)

The socially-motivated model draws on developmental theories of social and emotional competence in children (e.g., Bruner, 1983; Vygotsky, 1978; Stern, 1985), where social interaction is a key mechanism for learning, trust-building, and affect regulation. By analogy, Breazeal argues that for social robots to be seen as "intentional agents," they must demonstrate not just cognitive but also affective sensitivity—adapting to human moods and social contexts.

This model is supported and extended by the work of Fong, Nourbakhsh, and Dautenhahn (2003), who view social interaction as a central component of robotic autonomy in human environments. They emphasize the need for "behavioral plasticity" and cultural relevance to have robots accepted as social partners:

"Social robots must be able to communicate with humans using natural cues and behaviors that promote trust, cooperation, and intuitiveness."

(Fong et al., 2003)

Similarly, Dautenhahn (2007) introduces the concept of social intelligence, defining not only functional capability but also the robot's social suitability—its capacity to integrate into everyday human activities in a non-intrusive yet present manner.

In conclusion, Breazeal's socially motivated interaction model offers a comprehensive framework for seeing robots not just as technological artifacts but as social agents whose effectiveness depends on their ability to establish and sustain a social connection with users. This is particularly critical in domains like education, therapy, and care for vulnerable populations, where emotional trust and engagement often outweigh purely instrumental efficiency.

Trust Model in HRI (Hancock et al., 2011)

Hancock et al. (2011) present a meta-analysis outlining factors that influence trust in human–robot interaction. Their model covers three main categories: robot characteristics (e.g., reliability, adaptability), user characteristics (e.g., prior experience, propensity to trust), and contextual factors (e.g., environment, task). Their work emphasizes trust as dynamic, adapting as the interaction evolves.

In their study, Hancock, Billings, and Schaefer (2011) conduct an extensive meta-analysis encompassing 143 empirical studies on trust in HRI. They identify and categorize key predictors of trust shaping how people perceive, develop, and maintain trust in robotic systems.

The proposed model groups factors into three categories:

Robot Characteristics

Reliability: ability to perform tasks accurately and consistently
Adaptability: flexibility in response to environmental and human behavior changes
Autonomy: capability to make decisions independently
Appearance and behavior: aspects like humanoid form, speech intonation, and movement that enhance robot

sociality

User Characteristics

Prior experience with technology/robots: influencing trust positively or negatively
Personality traits: such as general trust propensity

Current emotional state: e.g., anxiety or confidence during interaction
Contextual Factors

Type of task: routine, critical, or risky tasks
Urgency and uncertainty: level of situational pressure
Social and cultural context: surrounding milieu in which the robot operates

Hancock et al. stress that trust in HRI is dynamic, evolving over time based on user experience, robot behavior, and interaction context. Unlike static trust models in traditional technologies, emotional and social components carry extra weight in interactive, humanlike systems:

"Trust in automation is not static, but is rather a dynamic variable influenced by the performance of the system, the experience of the user, and the demands of the task."

(Hancock et al., 2011, p. 100)

This model is widely cited and applied in settings such as military operations, autonomous vehicles, healthcare, and social robotics—where trust in automated systems is critical for successful implementation and ethical use.

Model of Anthropomorphism and Expectations (Bartneck et al., 2007)

Bartneck et al. (2007) offer a model exploring how perceived anthropomorphism—or human likeness—in robots affects their social appeal and perceived intelligence. They show that users project human qualities onto robots displaying similar traits, facilitating trust and engagement.

In their empirical study, Bartneck, Kulic, Croft, and Zoghbi (2007) examine how perceived anthropomorphism (the extent to which robots resemble humans) shapes user expectations about robot behavior, intelligence, and social appropriateness. Their model draws from the premise that anthropomorphism acts as a cognitive trigger, prompting users to attribute human-like characteristics to robots.

They identify three key dimensions of anthropomorphism's impact:

Social Attractiveness

Robots with more anthropomorphic traits (e.g., facial features, eyes, expressive mimicry) are seen as more socially acceptable, "closer," and less threatening—boosting motivation to interact.

Expectations of Competence and Intelligence

The more humanlike a robot appears, the higher users' expectations for cognitive ability and autonomy. This "competence by appearance" can foster trust—but also disappointment if expectations are unmet.

Cognitive Projection and Attribution of Intentions

Anthropomorphism eases projection of emotions, intentions, and motivations onto the robot. Users interpret robot signals (e.g., a "smile") as human-like emotional cues—facilitating communication, but risking overestimation of the system.

"Anthropomorphism is not just an aesthetic property; it is a psychological trigger that affects how humans judge and interact with robots."

(Bartneck et al., 2007, p. 110)

The model highlights a paradox: high anthropomorphism can be a double-edged sword—promoting intuitive interaction and trust on one hand, while triggering unrealistic expectations that can lead to distrust and emotional withdrawal if unmet. Bartneck et al. also discuss the Uncanny Valley phenomenon (Mori, 1970): overly humanlike but imperfect robots may elicit discomfort and distrust. Therefore, social robot design must balance functional effectiveness with social acceptability of anthropomorphic features.

This model is key to understanding the social psychology of human–robot interaction, especially in contexts where emotional connection is critical—such as education, caregiving, therapy, and consumer electronics.

Cultural Mediation Model (Li, 2015)

Li (2015) proposes a cultural model emphasizing that cultural values and social norms shape how people perceive and interact with social robots. For instance, collectivist cultures tend to more readily accept social robots seen as part of the social group, while individualistic cultures focus more on personal autonomy and privacy.

As social and service robots become global, a deeper understanding of cultural factors in human–robot perception and interaction is essential. Li (2015) introduces a cultural mediation model exploring how cultural values, norms, and beliefs influence user attitudes toward robots, their behavioral engagement, and levels of closeness and trust.

The model asserts that human expectations toward robots are not universal but culturally conditioned. Based on cross-cultural research, Li identifies key cultural dimensions shaping different interaction approaches:

Individualism vs. Collectivism

In collectivist cultures (e.g. Japan, South Korea, China), robots are more readily integrated into social structures as "members" or helpers.

In individualistic cultures (e.g. US, Germany, UK), users prioritize personal autonomy, privacy, and control, preferring robots that maintain distance and respect boundaries.

High vs. Low Context Communication

High-context cultures (e.g. East Asia), which emphasize nonverbal cues and social subtlety, expect robots to behave intuitively, indirectly, and contextually.

Low-context cultures (e.g. Scandinavia, US), favor straightforward, rationally structured robot communication.

Power Distance

High power-distance cultures more easily accept asymmetric relationships, including human–machine ones—the robot's subservient or assisting role is not problematic.

In low power-distance cultures, equality and partnership are preferred, influencing robot role and personality design.
"Cultural values shape how users interpret the robot's role, personality, and place in the social hierarchy, which in turn affects acceptance, trust, and usability."

(Li, 2015, p. 123)
Li's model has practical relevance for designing culturally adaptive robots. Such systems must understand linguistic and behavioral differences and modulate their behavior based on the user's cultural context. This

includes adapting speech style, emotional expression, relational distance, and goal prioritization (e.g. group harmony vs. individual preference).

Thus, cultural mediation is fundamental to successful human–robot interaction in multicultural environments. The global nature of technology demands sensitivity to local cultural norms.

In comparison to the Bulgarian academic context, Li's model requires local adaptation. Mitevska & Lazarova's (2024) study confirms cultural indicators like high power distance and uncertainty avoidance, suggesting that social robot integration must be carefully culturally framed—with role alignment, gradual adaptation, and institutional mediation.

Table 1. Comparison & Forecast: Li's Cultural Mediation Model (2015) vs. Mitevska & Lazarova (2024)

Cultural Dimension	Li (2015) – Theoretical Position	Mitevska & Lazarova (2024) – Empirical Data	Comparison	Implementation Forecast
Individualism/Collectivism	Collectivist cultures more easily accept social agents as "their own"	Mixed orientation: institutional individualism, but with collectivist elements	Matches hybrid profile	Robots must be presented as helpers, not competitors
Power Distance	High distance suppresses expectations for equal communication with machines	High distance; hierarchical management; low faculty participation	Model confirmed	Robots will be better accepted if they support hierarchy rather than challenge it
Uncertainty Avoidance	High anxiety → slow adoption of novelty	Strong intolerance for uncertainty and resistance to change	Full alignment	Implementation through small, well-structured steps & proven pilots
Perception of Robot Role	Expectations culturally shaped	Robots seen more as assistants than social partners	Low social imagination	Predicted need for additional "humanization" via behavior design and emotional expressiveness
Communication Context	High-context cultures expect subtle nonverbal signals	Academic environment not explicitly high-context, but sensitive to formality and status	Partial match	Robots with clearly defined behavior relative to status and role will be more acceptable

This comparison shows that the Bulgarian academic context shares many of Li's model predictions but is culturally ambivalent: balancing collectivism and individualism, cautious about novelty, and valuing institutional order. Therefore, successful social robot introduction requires cultural adaptation of roles, behavior, communication scripts, and institutional mediation and trust.

Model of Joint Meaning-Making (Mutlu et al., 2009)

Mutlu et al. (2009) emphasize that human–robot interaction is not a one-way process but involves joint meaning-making. The robot not only conveys information but actively participates in dialog and co-constructing interpretations of the situation. This is especially important in educational contexts where the goal is not merely knowledge transfer but social student engagement.

Against traditional models viewing communication as linear or unidirectional, Mutlu et al. (2009) propose a conceptual model based on joint meaning-making. Human–robot interaction thus becomes a cooperative process in which the meaning of communicative acts is co-constructed by both participants within a particular social and contextual scenario.

Inspired by sociocultural theories of communication and learning (e.g., Vygotsky, 1978; Clark, 1996), this model emphasizes that understanding arises through dialogic participation, adaptation, and coordination between interlocutors.

Key components of this model:

Contextual Compatibility

Meaning is created in specific social, physical, and cultural contexts. The robot must interpret environmental signals (e.g., gaze, object layout, gestures) and integrate them into a shared situational interpretation. This enables flexible and context-appropriate behavior.

Joint Reference & Attention

Successful interaction requires coordinated attention achieved via gaze, pointing, cues, and synchronized actions. Through these, a shared focus emerges, underpinning meaning-making. The robot not only follows human attention but also initiates focus direction, partaking in an "action dialogue."

Interactive Adaptation & Learning

Robots must adapt communicative behavior in response to the user's reactions, creating interaction depth. This is vital in education, where mere content delivery is insufficient—social engagement and co-construction of understanding are required.

"Rather than simply delivering content, robots must participate in the co-construction of meaning, taking on an active role in establishing mutual understanding."

(Mutlu et al., 2009, p. 478)
Mutlu and colleagues' model has strong practical relevance in education, therapy, and teamwork, where interaction goals are both functional and meaning-laden. For instance, in the classroom, a robot must not just teach but foster social dynamics in which the student feels involved, belonging, and understood.

In burgeoning interactive robotics paradigms, this model lays the groundwork for designing systems that don't just communicate but actively participate in meaning creation—critical for building trust, sustained engagement, and effective collaboration.

Integrating the Models

Together, these models offer different perspectives on the complex dynamics of human–social robot interaction. They highlight that effective interaction depends not only on the robot's technical features but also on psychological and cultural factors shaping user expectations and behavior. Achieving sustainable, productive interaction in educational and therapeutic settings requires an integrated approach that considers these multiple layers.

The models reviewed—socially motivated interaction (Breazeal, 2003), trust in HRI (Hancock et al., 2011), anthropomorphism and expectations (Bartneck et al., 2007), cultural mediation (Li, 2015), and joint meaning-making (Mutlu et al., 2009)—offer complementary insights into human–robot interaction and have direct applications in academic education contexts.

Table 2. Comparative Analysis of HRI Models & Their Educational Benefits

Model (Author, Year)	Main Focus	Key Concepts	Educational Relevance
Breazeal (2003)	Social engagement	Bidirectional communication, social cues	Creating engaging, interaction-based learning environments
Hancock et al. (2011)	Trust in robots	Reliability, adaptability, user experience, context	Enhancing acceptance of educational robots via reliable behavior
Bartneck et al. (2007)	Anthropomorphism & expectations	Social attractiveness, cognitive projection	Stimulating emotional connection and student motivation with humanlike traits
Li (2015)	Cultural mediation	Individualism/collectivism, social norms	Adapting robots to a diverse academic audience
Mutlu et al. (2009)	Joint meaning-making	Dialogue, co-interpretation	Supporting active learning and co-construction of knowledge

Educational Benefits

Enhanced Student Engagement

Models by Breazeal (2003) and Bartneck et al. (2007) show how social and emotional presence of robots can positively affect motivation and engagement in learning; high engagement is directly linked to deeper learning and memory retention.

Building Trust in Educational Technology

Hancock et al. (2011) is especially important for introducing new technologies in higher education, as it identifies what should be considered when designing a trustworthy educational partner—reliability, predictability, and user adaptation.

Cultural Sensitivity and Inclusion

Li (2015) highlights the need for cultural adaptability, particularly in internationalized academic environments. Students from different cultures react differently to social robots, and successful integration demands aligning with their values and expectations.

Promoting Collaborative Learning

Mutlu et al. (2009) emphasize that education is a co-constructed knowledge process, not passive transmission. Robots actively participating through joint referencing, adaptation, and dialog can facilitate collaborative learning—even in complex subjects.

Interdisciplinary Preparation and Critical Thinking

Applying these models jointly in universities—especially in engineering, pedagogy, cognitive science, and design—supports the development of critical thinking about ethical, cultural, and social technology aspects, essential for modern higher education. Understanding these processes is crucial for successfully introducing social robots in educational and therapeutic environments (Kennedy et al., 2015).

The combined application of these models in the university environment—especially in fields such as engineering, pedagogy, cognitive science, and design—supports the development of critical thinking regarding the ethical, cultural, and social dimensions of technology. These are skills of crucial importance for contemporary higher education.

References

1. Bartneck, C., Kulic, D., Croft, E., & Zoghbi, S. (2007). Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics, 1*(1), 71–81. https://doi.org/10.1007/s12369-008-0001-3
2. Breazeal, C. (2003). Toward sociable robots. *Robotics and Autonomous Systems, 42*(3–4), 167–175. https://doi.org/10.1016/S0921-8890(02)00373-1
3. Clark, H. H. (1996). *Using language*. Cambridge University Press.
4. Dautenhahn, K. (2007). Socially intelligent robots: Dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1480), 679–704. https://doi.org/10.1098/rstb.2006.2004
5. Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems, 42*(3–4), 143–166. https://doi.org/10.1016/S0921-8890(02)00372-X
6. Hancock, P. A., Billings, D. R., & Schaefer, K. E. (2011). Can you trust your robot? A meta-analysis of human–robot interaction. *Human Factors, 53*(5), 517–527. https://doi.org/10.1177/0018720811417254
7. Li, D. (2015). Cultural influences on user acceptance of social robots. *International Journal of Social Robotics, 7*(3), 379–392. https://doi.org/10.1007/s12369-015-0283-7
8. Mori, M. (1970). The uncanny valley. *Energy, 7*(4), 33–35. (Translated by K. F. MacDorman & N. Kageki, 2012 in *IEEE Robotics & Automation Magazine*, 19(2), 98–100)
9. Mutlu, B., Shiwa, T., Kanda, T., Ishiguro, H., & Hagita, N. (2009). Footing in human–robot conversations: How robots might shape participant roles using gaze cues. *Proceedings of the 4th ACM/IEEE International Conference on Human–Robot Interaction*, 61–68. https://doi.org/10.1145/1514095.1514113
10. Stern, D. N. (1985). *The interpersonal world of the infant: A view from psychoanalysis and developmental psychology*. Basic Books.
11. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
12. Bruner, J. (1983). *Child's talk: Learning to use language*. Oxford University Press.
13. Mitevska, M., & Lazarova, E. (n.d.). *Integrativnost na lichnostni i situatsionni faktori na organizatsionnoto povedenie* [Integrativity of personality and situational factors of organizational behavior]. Universitetsko izdatelstvo "Paisiy Hilendarski". (Original work in Bulgarian)
14. Mitevska-Encheva, M., & Petkov, G. (2017). *Akademichna motivatsiya* [Academic motivation]. Za Bukvite.
15. Kennedy, J., Baxter, P., & Belpaeme, T. (2015). The robot who tried too hard: Social behavior of a robot tutor can negatively affect child learning. *Proceedings of the 10th ACM/IEEE International Conference on Human–Robot Interaction*, 67–74. https://doi.org/10.1145/2696454.2696457