What do we think we are doing: principles of coupled self-regulation in human-robot interaction (HRI)

Idit Shalev[1]* and Tal Oron-Gilad[2]

[1] *Department of Education and the Zlotowski Center of Neuroscience, Ben-Gurion University of the Negev, Beer-Sheva, Israel,* [2] *Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Beer-Sheva, Israel*

The use of domestic service robots is becoming widespread. While in industrial settings robots are often used for specified tasks, the challenge in the case of robots put to domestic use is to afford human-robot collaboration in a variety of non-predefined and different daily tasks. Herein, we aim at identifying and understanding the conditions that will facilitate flexible collaboration between humans and robots. Past research of social and personality psychology was mainly focused on individual's self-regulation, defined as the ability to govern, or direct attention, resources, or action toward the realization of a particular goal (Higgins, 1989; Kruglanski et al., 2002). There is evidence that pursuing goals with the presence of others influences self-control (Fishbach and Trope, 2005), however only little is known on dyadic processes of self-regulation. Additionally, whereas research of goal pursuit in social psychology has mainly been associated with general processes of the structure and function of goals (Gollwitzer and Bargh, 1996; Carver and Scheier, 1998; Kruglanski et al., 2002; Fishbach and Ferguson, 2007; Custers and Aarts, 2010), human-robot interaction involves pragmatic interpersonal dilemmas such as how to coordinate human-robot activity and what knowledge should be shared between humans and robots over the course of action. To fill this gap, in what follows, we will define the unique characteristics of what we term as *human-robot coupled self-regulation*, which has the unique features of a dyadic asymmetric team aimed to increase the affordances of an individual in different activities. We will describe the unique characteristics of human-robot interaction and its special challenges toward goal pursuit.

Human and Robot are a Dyadic Instrumental Asymmetric Team

Our first assumption is that self-regulation of a human-robot couple could be conceptualized as a unique team configuration. A team is "a distinguishable set of two or more people who interact, dynamically, interdependently, and adaptively toward a common and valued goal/objective/mission, who have each been assigned specific roles or functions to perform, and who have a limited life-span of membership" (Salas et al., 1992, p. 4; Salas et al., 2010). Team members have differentiated responsibilities and roles (Cannon-Bowers et al., 1993). Therefore, essential for a team's successful performance is the understanding of the abilities and behaviors of its members that fit their experience and unique expertise for the task at hand.

Because humans and robots differ in their level of agency (the capacity to act and do) and their level of experience (the capacity to feel and sense), (Gray and Wegner, 2012), we argue that their contribution to the team is not symmetric. Based on the reasoning that genuine authorship of an action or situation may not always be clear (Dijksterhuis et al., 2008), we suggest that defined requirements of person, robot, and situation are essential to reduce the expectation gap.

Our perspective is that human-robot collaboration should be viewed in terms of functionality, to extend possibilities for the kinds of goals that humans want to pursue. These instrumental relations

between a person and her tool, used to increase the fit between person and environment, are termed affordances (Gibson, 1979). Following this view, we argue that robots can be perceived as self-regulatory tools to increase affordances across different situations (Koole and Veenstra, 2015). Our instrumental relational approach enables flexibility in tuning the robot's level of responsiveness and dominance in human-robot social contexts. For example, whereas the human member of the team holds a fixed ownership position, the robot's level of dominance could vary by user demands, or depending on the situation. To understand the usefulness of this principle, let us take for example 80 year old Mrs. Brown. She is physically fragile, but it is important for her to maintain an independent life style. This is why she has "Rupert," a multi-functional platform robot that serves as her aid. When she leaves the house she may want "Rupert" to lead and find the safest walking path to the store, thus she may set it to high dominance and responsiveness, in case she startles. At home, she may not desire high level of proactive care-taking and leave "Rupert" to be on call.

Concrete Level of Human-robot Negotiation

Our second assumption is that human-robot coupled self-regulation is based on concrete rather than abstract level of agreement. Carrying out human-robot joint actions demands continuous coordination on at least five elements: (1) who takes part; (2) what is the role of each member; (3) what is the joint goal; (4) how does each team member contribute to the timing and synchronization; and (5) where the actions take place (Clark, 2005). To address this, the robot should identify where the focus of attention of the human is, to what degree the attention of the human is focused on team actions, and how to convey feedback. Similarly, the human needs to calibrate expectations from the robot, i.e., be invested in the robot's immediate action or approval of action, and how to respond to the robot's requests (Alami et al., 2005).

Coupled self-regulation of goals requires agreement on goal setting and goal striving as two basic phases in goal pursuit (Gollwitzer and Oettingen, 2011). Whereas, robots may act automatically from initiation to completion of the task, humans' possible reflection on their performance may involve conscious awareness and create new representations of behavior, thus leading to communication gaps (Baumeister and Bargh, 2014). According to the action identification theory, a specific action can be verbally identified and interpreted from different levels of abstraction, ranging from low-level identities that specify how the action is performed, to high-level identities that signify why the action is performed. For instance, a person who "drinks water" can identify it as "holding a glass" (low level), or as "relieving thirst" (high level) (Vallacher and Wegner, 1987, 1989). This helps explain why different action identifications by human and robot may lead to dissimilar systems of goals and means of attainment (Kruglanski et al., 2002; Shah et al., 2002).

To address these challenges, we suggest the use of multiple human-robot forms of communication to pursue the joint goal. Lohan et al. (2014) proposed a distinction between two kinds of actions: path-oriented and manner-oriented, that can be communicated via two different linguistic utterance styles. Whereas, in path-oriented utterances the goal is stressed, in manner-oriented utterances, the means of motion are emphasized (e.g., Talmy, 1991). In our example, Mrs. Brown and "Rupert" carry a recliner to the porch (Path-"let's move the chair to the porch" or Manner-"I want to read my book on the porch"). Suddenly the phone rings and Mrs. Brown wants to go and answer ((Path-"let me go get the phone" or Manner-"I need to answer this call"). "Rupert" must understand that the goal has changed and pause.

Continuous and Various Communication Forms Over Goal Pursuit

Research indicates that professional and social interactions between team members can develop the team's social cognition (Klimoski and Mohammed, 1994). There is evidence that a team's fluent on-going communication regarding goal pursuit reduces the need for preexisting knowledge (Kozlowski and Bell, 2003). In social HRI, it is critical to generate many levels of interaction with the automation. Hence, the robot should always be present and aim to facilitate the goal, even if only to provide recommendations. In civil aviation, for example, communication is key especially if things turn out unexpectedly. In the Northwest 2009 incident in Minneapolis the automation had the capability, but was not designed to point out that the task was not performed as planned and that the pilots missed their destination. To borrow from our previous example, let us suppose Mrs. Brown wants to grab a pillow from the upper cabinet. The robot may not be able to reach so high, but it should continue to collaborate by providing feedback and advice; I cannot reach the uppermost cupboard (failure to complete task) but it is too dangerous for you to try to reach it on your own, if not urgent, perhaps we should call your son, or is there another pillow on a lower shelf?

Much of human communication over goal pursuit is based on social cues (e.g., gestures, and mimicry) that automatically generate social judgment and behavior (Chartrand and Bargh, 1999; van Baaren et al., 2003; Leander et al., 2010). Similarly, translation of social cues to social signals leads to inference of human intentions by robotic agents (Fiore et al., 2013). The relevance of automatic embodied cues for joint goal pursuit was demonstrated in human-human and human-robot synchronicity, suggesting that physical synchronicity is associated with experience of responsiveness and empathy (Sebanz and Knoblich, 2009; Cohen et al., 2010; Paladino et al., 2010; Boucher et al., 2012; Hoffman et al., 2014). Embodied communication is not only "used" by robots, but integrated in them to support both the recognition of the human's behavior and the generation of their behavior. Research of social signal processing and modeling multimodal communication, suggests that social and behavioral cues may be detectable from a machine, hence perceivable. Likewise, models of behavior are integrated in a way that a robot exhibits a more natural behavior, aiming at

a more successful interaction with the human (Pentland, 2007; Vinciarelli et al., 2012).

However, despite emerging findings from the field of embodied cognition on the potential of physical and social cues as an alternative route for communication, it was also claimed that embodied cognition cues can lead to different patterns of activation across different contexts (Loersch and Payne, 2011), thus prediction of behavior may be difficult (Shalev, 2015). A possible way to address this limitation is to use robots in fixed context, where interpretation to human's embodied signals is less ambiguous. For example Loth et al. (2013), have demonstrated that bar staff responded to a set of two non-verbal signals. Foster (2014), indicated that robotic sensors can similarly detect and respond to these signals.

Addressing the Human-robot Communication Gap over Goal Pursuit

Individuals frequently use embodied cues for functional self-regulatory purposes (Balcetis and Cole, 2009; Schnall et al., 2010; Bargh and Shalev, 2012; Shalev, 2014). However, using embodied cues as diagnostic inputs (Williams et al., 2009; Ackerman et al., 2010; Meier et al., 2012; Robinson and Fetterman, 2015; Winkielman et al., 2015) may lead to human-robot miscommunications. For example, human speakers expect co-located listeners to link visually perceivable objects and the verbally described references to them. Thus, humans may expect a co-located robot to have the same visual-verbal linking abilities (e.g., look at the green object on the right), thus developers must integrate the robot's visual system with natural language components to enable this flow of communication (Kopp, 2010; Cantrell et al., 2012; Vollmer et al., 2013).

Furthermore, there is also anecdotal evidence of human-human communication misunderstandings in complex scenes. For example orientation can be relative to egocentric, or exocentric (absolute or relative) locations. Soldiers for example, are taught to communicate via the exocentric coordinates of the compass rose. However, most humans tend to naturally orient relative to their egocentric perspective, which may be difficult for robots to depict. Interestingly, Cassenti et al. (2012) found that instructors used exocentric references to direct the robot and that it improved their performance relative to egocentric-only commands.

To address this communication gap, we argue that shared database, sensors and multiple types of displays and interaction means (e.g., physiological measures, eye tracking, voice, touch, text, button presses etc.) can enrich the robot's capacity of perception and expression. Similarly, to reduce expectation issues, technology can shape the way the user acts on the robot, how individuals understand what to expect from it, and how they can interact with a robot to refine their mutual understanding of the task at hand. Providing the relevant information about the current state of the robot, the progress of the task, and of the surrounding environment, can facilitate successful performance. Similarly, education efforts need to convey the ambiguity of ongoing human-robot communication, particularly the robot's physical and data-driven limitations, and to encourage problem solving and novelty seeking.

Acknowledgments

References

Ackerman, J. M., Nocera, C. C., and Bargh, J. A. (2010). Incidental haptic sensations influence social judgments and decisions. *Science* 328, 1712–1715. doi: 10.1126/science.1189993

Alami, R., Clodic, A., Montreuil, V., Sisbot, E. A., and Chatila, R. (2005). *Task Planning for Human-Robot Interaction.* Grenobel: JointsOc-EUSAI Conference.

Balcetis, E., and Cole, S. (2009). Body in mind: the role of embodied cognition in self-regulation. *Soc. Pers. Psychol. Commun.* 759–774. doi: 10.1111/j.1751-9004.2009.00197.x

Bargh, J. A., and Shalev, I. (2012). The substitutability of physical and social warmth in daily life. *Emotion* 12, 154–162. doi: 10.1037/a0023527

Baumeister, R. F., and Bargh, J. A. (2014). "Conscious and unconscious: toward an integrative understanding of human life and action," in *Dual Process Theories of the Social Mind*, eds J. Sherman, B. Gawronski, and Y. Trope (New York, NY: Guilford Press), 33–49.

Boucher, J. D., Pattacini, U., Lelong, A., Bailly, G., Elisei, F., Fagel, S. P. F., et al. (2012). I reach faster when i see you look: gaze effects in human-human and human-robot face-to-face cooperation. *Front. Neurorob.* 6:3. doi: 10.3389/fnbot.2012.00003

Cannon-Bowers, J. A., Salas, E., and Converse, S. A. (1993). "Shared mental models in expert team decision making," in *Current Issues in Individual and Group Decision Making*, ed N. J. Castellan (Mahwah, NJ: Erlbaum), 221–246.

Cantrell, R., Krause, E., Scheutz, M., Zillich, M., and Potapova, E. (2012). "Incremental referent grounding with NLP-biased visual search," in *Proceedings of AAAI 2012 Workshop on Grounding Language for Physical Systems.*

Carver, C. S., and Scheier, M. F. (1998). *On the Self-regulation of Behavior.* New York, NY: Cambridge University Press.

Cassenti, D. N., Kelley, T. D., Yagoda, R., and Avery, E. (2012). Improvements in robot navigation through operator speech preference. *Paladyn* 3, 102–111. doi: 10.2478/s13230-012-0100-6

Chartrand, T. L., and Bargh, J. A. (1999). The chameleon effect: the perception–behavior link and social interaction. *J. Pers. Soc. Psychol.* 76:893. doi: 10.1037/0022-3514.76.6.893

Clark, H. H. (2005). Coordinating with each other in a material world. *Dis. Stud.* 7, 507–525. doi: 10.1177/1461445605054404

Cohen, E., Ejsmond-Frey, R., Knight, N., and and Dunbar, R. I. M. (2010). Rowers' high: behavioural synchrony is correlated with elevated pain thresholds. *Biol. Lett.* 6, 106–108. doi: 10.1098/rsbl.2009.0670

Custers, R., and Aarts, H. (2010). The unconscious will: how the pursuit of goals operates outside of conscious awareness. *Science* 329, 47–50. doi: 10.1126/science.1188595

Dijksterhuis, A., Preston, J., Wegner, D. M., and Aarts, H. (2008). Effects of subliminal priming of self and God on self-attribution of authorship for events. *J. Exp. Soc. Psychol.* 44, 2–9. doi: 10.1016/j.jesp.2007.01.003

Fiore, S. M., Wiltshire, T. J., Lobato, E. J., Jentsch, F. G., Huang, W. H., and Axelrod, B. (2013). Toward understanding social cues and signals in human–robot interaction: effects of robot gaze and

proxemic behavior. *Front. Psychol.* 4:859. doi: 10.3389/fpsyg.2013.00859

Fishbach, A., and Ferguson, M. F. (2007). "The goal construct in social psychology," in *Social Psychology: Handbook of Basic Principles*, eds A. W. Kruglanski and T. E. Higgins (New York, NY: Guilford Press), 490–515.

Fishbach, A., and Trope, Y. (2005). The substitutability of external control and self-control in overcoming temptation. *J. Exp. Soc. Psychol.* 41, 256–270. doi: 10.1016/j.jesp.2004.07.002

Foster, M. E. (2014). "Validating attention classifiers for multi-party human-robot interaction," in *Proceedings of the 2014 ACM/IEEE International Conference on Human-robot Interaction: Workshop on Attention Models in Robotics.* (Bielefeld: ACM Press).

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.

Gollwitzer, P. M., and Bargh, J. A. (eds.). (1996). *The Psychology of Action: Linking Cognition and Motivation to Behavior*. New York, NY: Guilford Press.

Gollwitzer, P. M., and Oettingen, G. (2011). "Planning promotes goal striving," in *Handbook of self-regulation: Research, Theory, and Applications, 2nd edn.,* eds K. D. Vohs and R. F. Baumeister (New York, NY: Guilford), 162–185.

Gray, K., and Wegner, D. M. (2012). Feeling robots and human zombies: mind perception and the uncanny valley. *Cognition* 125, 125–130. doi: 10.1016/j.cognition.2012.06.007

Higgins, E. T. (1989). Continuities and discontinuities in self-regulatory and self-evaluative processes: a developmental theory relating self and affect. *J. Pers.* 57, 407–444. doi: 10.1111/j.1467-6494.1989.tb00488.x

Hoffman, G., Birnbaum, G. E., Reis, H. T., Vanunu, K., and Sass, O. (2014). "Robot responsiveness to human disclosure affects social impression and appeal," in *Proceedings of the 9th ACM/IEEE International Conference on Human-Robot Interaction.*

Klimoski, R., and Mohammed, S. (1994). Team mental model: construct or metaphor? *J. Man.* 20, 403–437.

Koole, S. L., and Veenstra, L. (2015). Does emotion regulation occur only inside people's heads? Towards a situated cognition analysis of emotion-regulatory dynamics. *Psychol. Inq.* 26, 61–68. doi: 10.1080/1047840X.2015.964657

Kopp, S. (2010). Social resonance and embodied coordination in face-to-face conversation with artificial interlocutors *Speech Commun.* 52, 587–597. doi: 10.1016/j.specom.2010.02.007

Kozlowski, S. W. J., and Bell, B. S. (2003). "Work groups and teams in organizations," in *Handbook of Psychology: Industrial and Organizational Psychology* eds W. C. Borman, D. R. Ilgen, R. J. Klimoski (London: Wiley), 333–375.

Kruglanski, A. W., Shah, J. Y., Fishbach, A., Friedman, R., Chun, W. Y., and Sleeth-Keppler, D. (2002). A theory of goal systems. *Adv. Exp. Soc. Psychol.* 34, 331–378. doi: 10.1016/s0065-2601(02)80008-9

Leander, N. P., Chartrand, T. L., and Wood, W. (2010). Mind your mannerisms: behavioral mimicry elicits stereotype conformity. *J. Exp. Soc. Psychol.* 47, 195–201. doi: 10.1016/j.jesp.2010.09.002

Loersch, C., and Payne, B. K. (2011). The situated inference model an integrative account of the effects of primes on perception, behavior, and motivation. *Pers. Psychol. Sci.* 6, 234–252. doi: 10.1177/1745691611406921

Lohan, K. S., Griffiths, S. S., Sciutti, A., Partmann, T. C., and Rohlfing, K. J. (2014). Co−development of manner and path concepts in language, action, and eye−gaze behavior. *Top. Cogn. Sci.* 6, 492–512. doi: 10.1111/tops.12098

Loth, S., Huth, K., and de Ruiter, J. P. (2013). Automatic detection of service initiation signals used in bars. *Front. Psychol.* 4:557. doi: 10.3389/fpsyg.2013.00557

Meier, B. P., Schnall, S., Schwarz, N., and Bargh, J. A. (2012). Embodiment in social psychology. *Top. Cogn. Sci.* 4, 705–716. doi: 10.1111/j.1756-8765.2012.01212.x

Paladino, M. P., Mazzurega, M., Pavani, F., and Schubert, T. W. (2010). Synchronous multisensory stimulation blurs self-other boundaries. *Psychol. Sci.* 21, 1202–1207. doi: 10.1177/0956797610379234

Pentland, A. A. (2007). Social signal processing. *IEEE Signal Process. Mag.* 24, 108–111. doi: 10.1109/MSP.2007.4286569

Robinson, M. D., and Fetterman, A. K. (2015). The embodiment of success and failure as forward versus backward movements. *PLoS ONE* 10:e0117285. doi: 10.1371/journal.pone.0117285

Salas, E., Cooke, N. J., and Gorman, J. C. (2010). The science of team performance: progress and the need for more. *Int. J. Hum. Comput. Stud.* 45, 75–104. doi: 10.1177/0018720810374614

Salas, E., Dickinson, T. L., Converse, S. A., and Tannenbaum, S. I. (1992). "Toward an understanding of team performance and training," in *Teams: Their Training and Performance*, eds R. J. Swezey and E. Salas (Norwood, NJ: Ablex), 3–29.

Schnall, S., Zadra, J. R., and Proffitt, D. R. (2010). Direct evidence for the economy of action: glucose and the perception of geographical slant. *Perception* 39, 464. doi: 10.1068/p6445

Sebanz, N., and Knoblich, G. (2009). Prediction in joint action: what, when, and where. *Top. Cogn. Sci.* 1, 353–367. doi: 10.1111/j.1756-8765.2009.01024.x

Shah, J. Y., Friedman, R., and Kruglanski, A. W. (2002). Forgetting all else: on the antecedents and consequences of goal shielding. *J. Pers. Soc. Psychol.* 83:1261. doi: 10.1037/0022-3514.83.6.1261

Shalev, I. (2014). Implicit energy loss: embodied dryness cues influence vitality and depletion. *J. Cons. Psychol.* 24, 260–270. doi: 10.1016/j.jcps.2013.09.011

Shalev, I. (2015). The architecture of embodied cue integration: insight from the "motivation as cognition" perspective. *Front. Psychol.* 6:658. doi: 10.3389/fpsyg.2015.00658

Talmy, L. (1991). "Path to realization: a typology of event conflation," in *Proceedings of the 17th Annual Meeting of the Berkeley Linguistics Society*, eds L. A. Sutton, C. Johnson and R. Shields (Berkeley, CA: Berkeley Linguistic Society), 480–519.

Vallacher, R. R., and Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychol. Rev.* 94:3. doi: 10.1037/0033-295X.94.1.3

Vallacher, R. R., and Wegner, D. M. (1989). Levels of personal agency: individual variation in action identification. *J. Per. Soc. Psychol.* 57:660. doi: 10.1037/0022-3514.57.4.660

van Baaren, R. B., Maddux, W. W., Chartrand, T. L., DeBouter, C., and van Knippenberg, A. (2003). It takes two to mimic: behavioral consequences of self-construals. *J. Pers. Soc. Psychol.* 84, 1093–1102. doi: 10.1037/0022-3514.84.5.1093

Vinciarelli, M., Pantic, D., Heylen, C., Pelachaud, I., Poggi, F., D'ericco, M., et al. (2012). Bridging the gap between social animal and unsocial machine: a survey of social signal processing. *IEEE Trans. Affect. Comput.* 3, 69–87. doi: 10.1109/T-AFFC.2011.27

Vollmer, A.-L., Wrede, B., Rohlfing, K. J., and Cangelosi, A. (2013). Do beliefs about a robot's capabilities influence alignment to its actions? *Proc. EpiRob/ICDL* 2013, 1–6. doi: 10.1109/DevLrn.2013.6652521

Williams, L. E., Huang, J. Y., and Bargh, J. A. (2009). The scaffolded mind: higher mental processes are grounded in early experience of the physical world. *Eur. J. Soc. Psychol.* 39, 1257–1267. doi: 10.1002/ejsp.665

Winkielman, P., Niedenthal, P., Wielgosz, J., Eelen, J., and Kavanagh, L. C. (2015). "Embodiment of cognition and emotion," in *APA handbooks in psychology. APA handbook of personality and social psychology, Vol. 1. Attitudes and social cognition* eds M. Mikulincer, P. R. Shaver, E. Borgida and J. A. Bargh (Washington, DC: American Psychological Association), 151–175. doi: 10.1037/14341-004

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.