Interaction as a bridge between cognition and robotics

Serge Thill
Interaction Lab, School of Informatics
University of Skövde, Sweden
serge.thill@his.se

Tom Ziemke
Interaction Lab, School of Informatics
University of Skövde, Sweden
and
Human-Centered Systems
Department of Computer & Information Science
Linköping University, Sweden
tom.ziemke@his.se

ABSTRACT

The triplet formed by studies of cognition, interaction, and robotics offers a number of opportunities for symbiotic relationships and mutual benefits. One such avenue is explored by the workshop's main theme in which cognition is seen as a bridge between interaction and robotics. Exploring ideas along that direction leads, as also discussed here, amongst others to the question of how theory-of-mind mechanisms might facilitate interaction between humans and robots.

A complementary view that we explore more fully here sees interaction as the bridge that connects robotics to relevant research on cognition. We follow recent trends in social cognition that go beyond studying social interaction as the outcome of the individuals' cognitive processes by seeing it as a constitutive and enabling element of social cognition. Here, we discuss this idea and show that it leads, amongst others, to the question of how interaction can be a constitutive element of a robot's cognitive architecture. It also leads to pointers towards research in the cognitive sciences that is beneficial to robotics but goes beyond cognitive architectures themselves. We show that considering the degree to which the robot is perceived by its end user as a tool and/or social partner points, for instance, to distributed and/or social cognition approaches for methodologies to evaluate human-robot interaction.

Categories and Subject Descriptors

I.2.9 [**Computing Methodologies**]: AI—*Robotics*

Keywords

Human-robot interaction evaluation, Distributed cognition, Social cognition, Social interaction

1. INTRODUCTION

Cognitive science and robotics research have long found common ground, whether this is to demonstrate the principles behind theories of cognition using, for example, robots or to improve artificial agents based on insights gained about human cognitive abilities. It is, however, noteworthy, that this common ground does not necessarily consider the *interaction* between robots and humans – most progress has indeed only necessitated one direction: human cognition inspiring robot cognitive architectures [45] or, alternatively, robotic models illustrating insights that could be relevant to the study of cognition [48].

Simultaneously, robots and other types of artificial agents are increasingly playing important roles in our society. Examples include research on robots for use in medical or therapeutic contexts [32, 43, 30, 3, 40], game-playing robots [1], but also the increasingly intelligent, adaptive and decision-making cars in use today [17, 39].

As such agents become more ubiquitous, there is thus a need to extend the common ground covered by robotics and the cognitive sciences into territories that concern such interactions. The 2015 HRI conference workshop tackling these issues head-on is titled "Cognition: a bridge between robotics and interaction". This cognition-centric view places an emphasis on cognitive mechanisms such as Theory of Mind (ToM): one agent's ability to create an internal model of another agent and use that to predict that agent's behaviour also improves the ability to interact with that agent [12]. Understanding these mechanisms – including, for instance, what social signals human pick up on, or how robot analogues of human ToM mechanisms might be constructed – leads to better interaction between humans and robots: it becomes the bridge between robots as such and interaction.

Here, as the paper title suggests, we explore the consequences of viewing this characterisation from a different angle in which interaction is the bridge between cognition and robotics. We do this to highlight aspects of cognitive science that are highly relevant to (social or sociable) robots that interact with humans and that complement the necessary focus on cognitive architectures. The purpose is therefore not to disagree with the view of cognition as a bridge between robotics and interaction, but rather to extend it with the complementary insights that interaction as a bridge between cognition and robotics leads to.

2. COGNITIVE SCIENCE AND ROBOTICS

The idea that robotics research might help to further research in the cognitive sciences (and vice versa) has been around for a while, fuelled in particular by the developing prominence of embodied theories of cognition [44, 6, 49, 52] on one side and increasingly well-engineered and cheap(er) robots on the other. In this section, we briefly discuss three traditional approaches to research that explore this symbiotic potential. This will serve as a background against which to discuss the added contributions of a focus on the interactive aspects.

2.1 Proof of concept

The first approach demonstrates principles of cognitive science (usually embodied or situated flavours thereof) using robots (see [48] for a brief review of relevant work). The general theme is that robots, through their design, display specific behaviours that are not themselves explicitly represented within the system, thus illustrating that "embodiment and embedding can therefore *replace* internal algorithms and lead to stable, functional behaviour" ([48], p. 4, emphasis in original). A similar example is that of morphological computation [28, 27]; the idea that computations can be offloaded into a suitably designed morphology. Well-designed legs on a quadruped robot can for example lead to an appropriate quadruped walking gait without the need for complicated control mechanisms [28].

As such, the purpose of these models is first and foremost to *illustrate by example* concepts that would otherwise be difficult to verify in a living organism. There is a benefit to robotics because these illustrations tend to be viable implementations of behaviours that might be useful for robotic applications too – such as pointers on how to simplify locomotory control as in the example above. The relevance to the study of cognition, on the other hand, is weaker: it is possible that predictions generated by such a mechanism turn out to match the biological counterpart (as in the case of Webb's cricket robots [47], see the discussion in [48]) but this is not a requirement since the original purpose is typically the demonstration of the concept (as in the case of robots that show tidying behaviour even though their underlying controllers do not explicitly implement any such behaviour [21], again discussed in [48]).

2.2 Embodying models of human cognition

The second approach attempts to more directly study human cognition using artificial agents. The motivation follows more or less as a consequence of accepting embodied or grounded theories of cognition according to which the body plays a fundamental, non-abstractable role in cognition. The Chinese Room argument [34], or the symbol grounding problem [16] are frequently cited in this context and the conclusion drawn tends to be that a cognitive model must be instantiated in a physical agent (how else could the role of the body otherwise be represented?). It is, however, worth noting that the mere provision of a robot instantiation does not by itself overcome the problems described, for instance, by the Chinese room argument: indeed, the "robot reply", in which a robot body is used to provide a sensorimotor apparatus in which to "embody" the computational model has already been considered and rejected by Searle in his original paper [34] (for a fuller discussion, see [55]).

Another challenge that these robot-reliant ways of studying human cognition face is simply that a robot body is not like a human body, even if it is described as "humanoid" [54]. Embodied accounts (irrespective of the particular theoretic flavour) ascribe a role to the body (and/or environment) that is fundamental in shaping cognition and cannot be abstracted away; yet robotic implementations often begin with a sensory apparatus that is radically different from the human senses and by necessity includes several simplifications and abstractions. Vision, for example, is often simplified, for instance by using brightly coloured and easily discriminable objects [22]. Although an advantage of robotic models is that they force integration from sensory perception to motor action [23, 26], this integration is not as forceful as it seems.

It is of course true of all models that they must contain abstractions and simplifications (otherwise they would not be a model). It has famously been said that "all models are wrong; the practical question is how wrong do they have to be to not be useful" [4]. When the model is not just of the cognitive process (because it is, in this view, meaningless to talk of "just" the cognitive process), but also of the body, and therefore all sensorimotor aspects, as well as the environment (whether this is because the agent is simulated or put into a purpose-engineered artificial situation), one has to exercise extra care when discussing the relevance of insights from such embodied models to human cognition [41].

However, this is not to say that such models have no utility beyond illustration of concepts (in which case, we should group them under the *proof-of-concept* approach discussed previously). For instance, any cognitive process that requires interaction with the environment needs to be modelled in a manner in which such interactions are possible. Even strongly abstracted sensorimotor mechanisms can provide insights into minimal requirements for the cognitive process of interest [26].

2.3 Cognitive science for the benefit of robots

The previous two approaches were examples of research whose aim is first and foremost a contribution to the study of cognitive mechanisms. By virtue of necessitating a robotic implementation, there is also a benefit to the field of robotics since, as previously argued, the algorithms and controllers that are developed may find new approaches or solutions to problems and challenges in robotics.

At the same time, there is an approach to research at the intersection between the study of cognition and robotics that aims first and foremost to benefit robotics research: knowledge and results from the cognitive sciences can be used to create "better" (defined, for instance, as an increased ability to cope with uncertainty or unpredicted events) robots. ToM mechanisms are an important example of cognitive mechanisms that have been used to this effect (see [12] for a discussion of the two main flavours of ToM – *theory theory* and *simulation theory* – in the context of social robots). Indeed, to interact proficiently with humans, such robots may simply require at least a rudimentary ToM; an internal model that can be used to estimate mental states of humans, in particular their intentions, expectations and predicted reactions to actions by the agent [31, 40, 41]. A second example is given in [18] (as cited in [50]) - here the insight that anticipation and perceptual simulation are important, for humans, in the perception of conspecifics and joint action are used to design a robot that can interact fluently with human partners. Finally, see [11] for an early review of a large number of socially interactive robots and the design principles and inspirations behind them.

3. INTERACTION AT THE CENTRE

The previous section has illustrated a number of active research areas that explore the symbiotic relationship between research in the cognitive sciences and robotics[45, 33]. It is readily apparent that interaction does not necessarily need to be considered in these areas – it is naturally not excluded: the ToM mechanisms discussed in section 2.3 are a prime example of a benefit that the study of cognition brings to robotics whereas research on human interpretation of robot movements leads to what aspects of robot motion may involve mechanisms thought to underlie, for instance, social interaction [13].

3.1 Interaction as a constitutive component of cognition

When interaction is considered in robotics research, however, it is often understood as two agents[1], one human, one artificial, each with their own cognitive apparatus, using that apparatus to engage in interaction with the other agent. This both reflects traditional views in social cognition (which are mainly interested in the individual's internal mechanisms underlying interaction) and features the same pitfall: not explicitly recognising that the interaction itself is fundamental, and *part of the overall cognitive process* as opposed to merely the result thereof [10]. In other words, the interactive setting does not merely play a contextual role for an individual's cognitive mechanisms but also takes on enabling and constitutive roles [10]. Just as a cognitive architecture in which the body does not play a fundamental, irremovable and irreplaceable part of the cognitive process is not an embodied architecture [55, 56], a cognitive architecture in which interaction is merely a contextual aspect lacks something.

This is the first core insight we gain from a focus on interaction: as the field of social cognition is moving away from an individualistic view of interaction, robotic cognitive architectures need to consider the implications of an enabling, constitutive role of interaction with other agents in their overall functionality (see also *e.g* [9] for a similar argument). For example, robots are often built for specific purposes – their desired behaviour is therefore given by that application. Yet, to deal with uncertainty and unforeseen events, it is not desirable to specify all behaviours axiomatically at design time – rather, the ability for appropriate behaviour to emerge from the robot's experience is needed. In this context, it can be shown that casting the objective function modulating such emergence in terms of *interaction* may lead to desirable, yet not unnecessarily constrained behaviour [41].

3.2 Evaluating human-robot interaction (HRI)

Robots (and other artificial agents), as discussed before, can in almost all cases be expected to interact with humans to some degree. There is therefore also a need to evaluate these robots *in terms of their interaction with humans*. There are no "simple" metrics to this end since successful performance, by definition, depends on the human/artificial agent system as a whole.

In other words, one cannot consider the robot's performance in isolation; its success is a function of how well the agent/human system functions (see [2, 10, 41] for related arguments). In some applications, for instance robot-enhanced therapy (RET) for children with autism spectrum disorder (ASD) [32, 40], the ideal measure (*e.g.* long-lasting benefits) is also simply unavailable since it can only be meaningfully be sampled after years if not decades after the artificial agent is deployed. Other scenarios might be entirely open-ended and without any direct task to be achieved, yet the need to evaluate the robot remains. Further, although it is possible to achieve some form of evaluation by asking persons who interacted with artificial agents to fill out questionnaires and similar (see also [39] for an example in which just that has been done), such options are typically not available if the persons interacting with the artificial agent are in fact children [2]. More generally speaking, these methods usually require the subjects to have a substantial degree of insight into their own cognitive processes.

[1]The present argument easily extends to multi-agent systems, but two are sufficient for illustrative purposes

How to characterise and evaluate interaction has long been a topic in HRI (see for instance the extensive survey and introduction to the topics in [14]). An immediate realisation in such efforts is that there is no "one size fits all" solution; robots can interact with humans in a number of ways that then define and shape what one expects from such interaction. This then leads to a number of proposals for dimensions along which to rate the precise nature of the interaction at hand. The ubiquitous example is that of autonomy: in 1978, Sheridan and Verplank proposed a 10-step scale describing degrees of automation, ranging from machines that are entirely remote-controlled to machines that ignore human beings altogether [36]. Since then, there have been numerous discussions of the scale in particular and the concept of autonomy in HRI in general (*e.g.* [14, 51, 38, 42]). It is for instance repeatedly argued that "human-robot interaction cannot be studied without consideration of a robot's degree of autonomy" [42] (p. 14).

It is therefore worth emphasising that autonomy is a particularly difficult term that can mean very different things to different people [45, 53]. In HRI, for instance, the take on autonomy is often task-oriented – referring, for example (as in Sheridan and Verplank's scale), to the degree to which the human has to assist the machine in accomplishing a given task [51], thus measuring the degree of automation. Cognitive scientists, on the other hand, might consider autonomy more in terms of self-sufficiency, or behaviour that is *not* determined entirely by external events but shaped by internal goals of an agent [35].

This highlights an important point pertinent to the possible benefits between the study of cognition and robotics: it needs to be kept in mind that autonomy is an overloaded term (as are others) when researchers from different disciplines meet. In [45], for instance, no less than 19 different takes on autonomy are discussed, a list that is by no means complete. Although we cannot possibly do the concept justice here (and instead point to [45], Ch. 4), the relevant insight is that, when the study of cognition and robotics meet, it is critical to be clear about the terms one uses; a symbiotic relationship depends on a common understanding of such concepts.

When autonomy refers to the degree of automation, it is a dimension in which social interactions occupy the middle of the range (since there is no meaningful interaction in the fully automated case and merely tele-operating a robot does not constitute social interaction with another cognitive agent). Likewise, other metrics that fundamentally seek to evaluate HRI performance in terms of task performance (*e.g.* robot efficiency and effectiveness in the task and human situation awareness [38]) do not assess the social interaction itself. Metrics that do would need to measure, it has been suggested [38], interaction characteristics, persuasiveness, trust, engagement and compliance, but the exact methodologies for that remain unclear.

3.3 Interaction-focussed HRI evaluation

It has been suggested [42] that we may not actually want to interact with robots in precisely the same way as we interact with other humans. Whether or not one reserves the term "social interaction" for human-human interaction or opens it up to human-robot interaction is a different debate and does not *per se* invalidate the idea of evaluating HRI as a type of interaction that can be usefully characterised by metrics similar to those used for human-human interaction.

It does, however, lead to the interesting question of how robots (and other artificial agents) are *perceived* – it is for instance known that, for some robots and actions at least, the human mirror system is activated when observing robot actions [13] that can then be interpreted as being goal-directed [33], which does point towards the likelihood that interacting with humans and robots – when they are perceived as having some agency at least – may not be entirely different.

The interesting question therefore is to what degree robots are actually perceived as agents by the people they interact with. With that characterisation, we can then return to the central theme of the paper and discuss methodologies in the cognitive sciences that may be useful for characterising human-robot interaction based on how the robot is perceived.

In the context of increasingly automated vehicles, it has been suggested that a useful way to characterise human-vehicle interaction is by establishing the degree to which the vehicle is perceived as a *tool*, used in navigation tasks, as opposed to an *intelligent agent*, with whom the driver collaborates in solving the task [39]. Here, we explore a similar characterisation for robots and artificial agents in general. In particular, we illustrate in the next two subsections that they can be understood *by their end users* as, to varying degrees, both tools and social partners.

Such characterisations have been used in the past: the "robot role" (ranging from tool/machine to companion/partner) is, for example, one of the suggested dimensions for determining the requirements on a robot's social skills [8]. Here, however, we use this dimension to identify theories of cognitive science useful in evaluating human-robot interaction. It is difficult to find such theories in the traditional overlap between cognitive science and robotics discussed in section 2: the first two approaches, proof-of-concept and embodying models, mainly use artificial agents for theoretical insights that could include interaction between agents (see *e.g.* robot language games [37]), but do not have to. When cognitive models are primarily used as an inspiration for better robots, validation is given by an adequate implementation of the targeted cognitive ability.

3.3.1 *Artificial agents are tools*

Artificial agents are usually created for a purpose - this can be academic (*e.g.* as demonstrators of cognitive theories or as tools for studying cognition as discussed above) or with a practical application in mind (*e.g.* for use in elderly care, therapy, navigation of dangerous or inhospitable terrain and so on). They exist, therefore, to assist humans in achieving certain goals (even if they are designed as autonomous agents). Artefacts used by people in addition to their own body to achieve a certain purpose are, by definition, tools.

3.3.2 *Artificial agents are social partners*

Although artificial agents are, as argued above, always created for a purpose, significant research efforts [45] are dedicated to creating agents with interesting cognitive abilities (whether it is to showcase models of these abilities or more directly to allow the agents to tackle more complicated and less trivial tasks). It is therefore clear that artificial agents can be seen as more than tools: indeed, they can be social partners with whom we interact, *collaborating* in solving the task for which they were created.

This highlights (again) that the artificial agent should not be seen by itself but rather as *interacting* with humans. [2] for instance argues that technical challenges in cHRI (HRI in which the humans are specifically children) may be overcome if we see the cognition

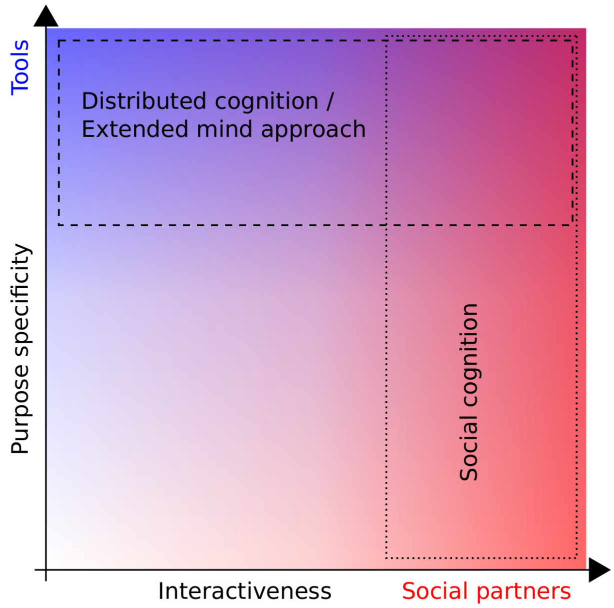

Figure 1: Diagram positioning artificial agents in function of how their interactiveness and purpose specificity are perceived by the end users. Boxes inside the diagram indicate the cognitive science research strands one should primarily consider when evaluating artificial agents in that area of the spectrum.

of a human/artificial agent ensemble as the product of their interaction. A critical point these authors make is that one agent (*e.g.* the human) can cover for potential failings of the other (*e.g.* the robot), which in itself illustrates that one cannot evaluate the robot by itself (see the credit assignment problem).

Viewing artificial agents as social partners also has consequences for how one expects humans to interact with them. For instance, humans tend to modulate their behaviour based on their beliefs about, amongst others, the cognitive abilities of the agent they interact with [5]. This has been shown to extend to robots [46, 20]. Furthermore, our recent research indicates that this extends even to cars [39].

3.3.3 *Artificial agents are tools and social partners*

It is worth emphasising that the two views of artificial agents, as sketched out above, are not mutually exclusive. In other words, they do not form two ends on a scale as in previous examples of similar scales [9]. Rather, artificial agents can be, to varying degrees, both:

- If a robot is built for very specific purposes, it is a tool created to achieve that purpose. But not all robots are created for such specific purposes: another scale used by Dautenhahn [8] considers "robot functionalities" which can range from clearly defined to open and adaptive. In a similar vein, we use *purpose specificity* as a dimension along which to measure whether or not an artificial agent may be perceived as a tool.
- Similarly, the degree to which a robot can be considered a social partner depends on the degree to which it is seen as interacting with its end user (by the end user). Again, in

Dautenhahn's set of scales used to determine social requirements, the analogue is the "contact with humans" dimension. Here, we refer to this dimension as *interactiveness* of a robot to more explicitly capture the fact the contact involves interaction.

These two dimensions - purpose specificity and interactiveness create a 2D map of artificial agents (as sketched in Fig. 1). We resist the temptation to populate the sketch with placements of example robots and other artificial agents. To give but two examples:

- It can be argued that cars would typically score high on purpose specificity (they are built purposefully for navigating from A to B) and low on interactiveness (since – until recently – they do not interact with the drive beyond providing information about their internal state). There are, however, significant technological developments [17, 39] that will increase the interactiveness of cars. In the near future, cars might thus move further to the right in Fig. 1.
- Therapeutic robots, for instance as used for ASD therapy [32] are naturally highly interactive but their purpose is more open-ended, [8], reducing their purpose specificity (especially as perceived by the child). One could conceivably expect to place them around the middle of the right side of the graph.

3.3.4 Cognitive theories for HRI

With this in mind, we can now consider cognitive theories that have traditionally dealt with human interaction with tools and social partners. First, robots that score highly on the purpose specificity scale more or less directly speak to *extended and distributed views of cognition* [19, 7].

From the extended mind view [7], we can take the position that the artificial agent becomes just such an extension of the mind. The cognitive process according to which the human uses the artificial agent to achieve a certain purpose cannot be defined within the human alone; the artefact at a minimum becomes a resource (of what type depends on the agent).

From distributed cognition [19], we similarly get the perspective that cognition should be understood in terms of the interaction with the material and social world. The paradigm additionally comes with a large set of tools for analysing such interactions, most dominantly ethnography (see [24] for an extensive review of these aspects of distributed cognition, including criticisms and rebuttals). Distributed cognition has also already found applications in HCI, for instance as a method "with which to understand the underlying mechanisms of the relationships between humans and computer" [24] (p. 63). For instance, a distributed cognition-inspired methodology for studying the interaction between humans and machines in a maritime control room has been developed [25]. While it may of course be too bold to refer to such a control room as an artificial agent, the example illustrates that it is possible to take the basic ideas from distributed cognition into a more formalised approach to studying the interaction between man and machine.

When robots score high on the interactiveness scale, meanwhile, it is possible (and necessary) to go beyond distributed cognition and explicitly treat the interaction as social. Consequently, this points to insights from social cognition. Here, social interaction can, for instance, be defined as "two or more autonomous agents co-regulating their coupling with the effect that their autonomy is not destroyed and their relational dynamics *acquire an autonomy of their own*" [10] (p. 441, emphasis added). A highly interactive robot would necessarily possess some autonomy in the same sense (and notably not necessarily the sense usually given to autonomy in HRI, see the previous discussion in section 3.2); it is therefore clear that any take on this agent that ignores the interactive aspect will fail to adequately take into account this coupling.

A comprehensive review of methods that are useful in studying social interaction can also be found in [10]. These include conversation and gesture analysis, with the particular insight that Motion Energy Analysis [15] could predict subjective assessments of a therapy session's quality based on bodily coordination between patient and therapist [29] (as cited in [10]). Work such as this provides a clear entry point by which one could possibly evaluate therapeutic robots, addressing for instance the concerns of [2] that one cannot easily make children fill out a questionnaire. Even though putting the focus of social cognition on embodied social interaction is, as noted at the beginning, a relatively recent trend [10], it is clear that the field is developing a range of techniques that are useful for evaluating the quality of this interaction. These techniques may well find further applications in the study of the interaction between humans and robots.

4. CONCLUSIONS

We have highlighted the importance of interaction in the ⟨*cognition, robotics, interaction*⟩ triplet. This perspective has enabled us to illustrate that interaction is not just contextual, but rather an enabling and constitutive component of social cognition [10]. Although cognition can, as also illustrated here, rightfully be seen as a bridge between robotics and interaction, the latter also functions as a bridge between robotics and cognition; in particular enabling robotics research to develop cognitive architectures in which the interaction likewise plays a constitutive, enabling component (as opposed to being the outcome).

The perspective has also enabled us to consider the roles that robots play when interacting with humans. We have argued that the degree to which the robot is perceived as fulfilling a specific purpose as well as the degree to which it is perceived as interacting with humans – in both cases as seen from the end user – are useful dimensions to consider in this respect. In particular, the relative degree to which robots score on these dimensions form a guide to theories in cognitive science that can be useful to understand and evaluate the interaction between the human and the robot.

Given that robots and other artificial agents (we have mentioned cars in particular) are increasingly entering into our daily lives, such evaluations become increasingly important. It of course remains to be seen to what extent exactly one can translate the methodologies, explanatory tools and techniques from distributed and social cognition onto the study of artificial agents. Here, our purpose has been to highlight that the relevance of cognitive research for robotics goes beyond inspiration for better cognitive architectures as such to include the study of how the human-robot system as a whole functions. Such a perspective has relevance in many application areas. In robots used for therapy, for instance RET aimed at children with ASD, the child-robot system is more than just the sum of a child and a robot - a relationship between the two exists that cannot be abstracted away [9] and that has implications both for the design of the cognitive architecture of the robot [41] and, as argued here, for the evaluation of the robot.

5. ACKNOWLEDGEMENTS

This work has been supported by DREAM (www.dream2020.eu), funded by the European Commission (FP7-ICT, project number 611391), and TINA/AIR, funded by KK-SIDUS, Sweden.

6. REFERENCES

[1] P. Baxter, J. de Greeff, and T. Belpaeme. Do children behave differently with a social robot if with peers? In *International Conference on Social Robotics (ICSR 2013)*, October 2013.

[2] T. Belpaeme, P. Baxter, J. de Greeff, J. Kennedy, R. Read, R. Looije, M. Neerincx, I. Baroni, and M. Zelati. Child-robot interaction: Perspectives and challenges. In G. Herrmann, M. Pearson, A. Lenz, P. Bremner, A. Spiers, and U. Leonards, editors, *Social Robotics, volume 8239 of Lecture Notes in Computer Science*, volume 8239, pages 452–459. Springer International Publishing, 2013.

[3] O. A. Blanson Henkemans, B. P. Bierman, J. Janssen, M. A. Neerincx, R. Looije, H. van der Bosch, and J. A. van der Giessen. Using a robot to personalise health education for children with diabetes type 1: A pilot study. *Patient Education and Counseling*, (PEC-4519):8, 2013.

[4] G. E. P. Box and N. R. Draper. *Empirical Model Building and Response Surfaces*. John Wiley & Sons, New York, NY, 1987.

[5] H. P. Branigan, M. J. Pickering, J. Pearson, J. F. McLean, and A. Brown. The role of beliefs in lexical alignment: Evidence from dialogs with humans and computers. *Cognition*, 121(1):41 – 57, 2011.

[6] A. Clark. *Being there: Putting brain, body, and world together again*. MIT press, Cambridge, MA, 1997.

[7] A. Clark and D. J. Chalmers. The extended mind. *Analysis*, 58:7–19, 1998.

[8] K. Dautenhahn. Roles and functions of robots in human society: implications from research in autism therapy. *Robotica*, 21:443–452, 8 2003.

[9] K. Dautenhahn. Socially intelligent robots: dimensions of human–robot interaction. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362(1480):679–704, 2007.

[10] H. De Jaegher, E. Di Paolo, and S. Gallagher. Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, 14(10):441 – 447, 2010.

[11] T. Fong, I. Nourbakhsh, and K. Dautenhahn. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3âĂŞ4):143 – 166, 2003. Socially Interactive Robots.

[12] S. Gallagher. Social cognition and social robots. *Pragmatics & Cognition*, 15(3):435–453, 2007.

[13] V. Gazzola, G. Rizzolatti, B. Wicker, and C. Keysers. The anthropomorphic brain: The mirror neuron system responds to human and robotic actions. *NeuroImage*, 35(4):1674 – 1684, 2007.

[14] M. A. Goodrich and A. C. Schultz. Human-robot interaction: A survey. *Found. Trends Hum.-Comput. Interact.*, 1(3):203–275, Jan. 2007.

[15] K. Grammer, M. Honda, A. Juette, and A. Schmitt. Fuzziness of nonverbal courtship communication unblurred by motion energy detection. *Journal of Personality and Social Psychology*, 77(3):487–508, 1999.

[16] S. Harnad. The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1-3):335–346, 1990.

[17] A. Heide and K. Henning. The “cognitive car”: A roadmap for research issues in the automotive sector. *Annual Reviews in Control*, 30(2):197 – 203, 2006.

[18] G. Hoffman and C. Breazeal. Effects of anticipatory perceptual simulation on practiced human-robot tasks. *Autonomous Robots*, 28(4):403–423, 2010.

[19] E. Hutchins. *Cognition in the Wild*. MIT Press, Cambridge, MA, 1995.

[20] S. Kopp. Social resonance and embodied coordination in face-to-face conversation with artificial interlocutors. *Speech Communication*, 52(6):587–597, 2010.

[21] M. Maris and R. Boeckhorst. Exploiting physical constraints: heap formation through behavioral error in a group of robots. In *Intelligent Robots and Systems ’96, IROS 96, Proceedings of the 1996 IEEE/RSJ International Conference on*, volume 3, pages 1655–1660 vol.3, 1996.

[22] A. F. Morse, T. Belpaeme, A. Cangelosi, and L. B. Smith. Thinking with your body: Modelling spatial biases in categorization using a real humanoid robot. In S. Ohlsson and R. Catrambone, editors, *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, pages 1362–1367, Austin, TX, 2010. Cognitive Science Society.

[23] A. F. Morse, C. Herrera, R. Clowes, A. Montebelli, and T. Ziemke. The role of robotic modelling in cognitive science. *New Ideas in Psychology*, 29(3):312–324, 2011.

[24] M. Nilsson. *Capturing semi-automated decision making: the methodology of CASADEMA*. PhD thesis, Örebro University, 2010.

[25] M. Nilsson, J. van Laere, T. Susi, and T. Ziemke. Information fusion in practice: A distributed cognition perspective on the active role of users. *Information Fusion*, 13(1):60 – 78, 2012.

[26] G. Pezzulo, L. W. Barsalou, A. Cangelosi, M. H. Fischer, K. McRae, and M. J. Spivey. The mechanics of embodiment: a dialog on embodiment and computational modeling. *Frontiers in Psychology*, 2(5), 2011.

[27] R. Pfeifer, J. Bongard, and S. Grand. *How the body shapes the way we think: a new view of intelligence*. MIT press, Cambridge, MA, 2007.

[28] R. Pfeifer and F. Iida. Morphological computation: Connecting body, brain and environment. *Japanese Scientific Monthly*, 2005.

[29] F. Ramseyer and W. Tschacher. Synchrony: A core concept for a constructivist approach to psychotherapy. *Constructivism in the Human Sciences*, 11(1):150–171, 2006.

[30] B. Robins, K. Dautenhahn, R. Boekhorst, and A. Billard. Robotic assistants in therapy and education of children with autism: Can a small humanoid robot help encourage social interaction skills? *Universal Access in the Information Society*, 4(2):105–120, 2005.

[31] B. Scassellati. Theory of mind for a humanoid robot. *Autonomous Robots*, 12(1):13–24, 2002.

[32] B. Scassellati, H. Admoni, and M. Matarić. Robots for use in autism research. *Annual Review of Biomedical Engineering*, 14:275–294, 2012.

[33] A. Sciutti, A. Bisio, F. Nori, G. Metta, L. Fadiga, and G. Sandini. Robots can be perceived as goal-oriented agents. *Interaction Studies*, 14(3):329–350, 2013.

[34] J. R. Searle. Minds, brains, and programs. *Behavioral and Brain Sciences*, 3:417–424, 9 1980.

[35] A. Seth. Measuring autonomy and emergence via Granger

causality. *Artificial Life*, 16(2):179–196, April 2010.

[36] T. B. Sheridan and W. L. Verplank. Human and computer control of undersea teleoperators. Technical report, MIT Man-Machine Systems Laboratory, 1978.

[37] L. Steels. Evolving grounded communication for robots. *Trends in cognitive sciences*, 7(7):308 – 312, 2003.

[38] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich. Common metrics for human-robot interaction. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-robot Interaction*, HRI '06, pages 33–40, New York, NY, USA, 2006. ACM.

[39] S. Thill, P. E. Hemeren, and M. Nilsson. The apparent intelligence of a system as a factor in situation awareness. In *Proceedings of the 4th IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, pages 52 – 58, 2014.

[40] S. Thill, C. Pop, T. Belpaeme, T. Ziemke, and B. Vanderborght. Robot-assisted therapy for autism spectrum disorders with (partially) autonomous control: Challenges and outlook. *Paladyn*, 3(4):209–217, 2012.

[41] S. Thill and D. Vernon. How to design emergent models of cognition for application-driven artificial agents. In *Proceedings of the 14th Neural Computation and Psychology Workshop (NCPW14)*, submitted.

[42] S. Thrun. Toward a framework for human-robot interaction. *Human–Computer Interaction*, 19(1-2):9–24, 2004.

[43] B. Vanderborght, R. E. Simut, J. Saldien, C. A. Pop, A. S. Rusu, S. Pintea, D. Lefeber, and D. David. Social stories for autistic children told by the huggable robot Probo. In *Cognitive Neuroscience Robotics workshop IROS*, pages 1–6, 2011.

[44] F. J. Varela, E. Rosch, and E. Thompson. *The embodied mind: Cognitive science and human experience*. MIT press, Cambridge, Ma, 1992.

[45] D. Vernon. *Artificial Cognitive Systems: A primer*. MIT Press, Cambridge, MA, 2014.

[46] A.-L. Vollmer, B. Wrede, K. J. Rohlfing, and A. Cangelosi. Do beliefs about a robot's capabilities influence alignment to its actions? In *Development and Learning and Epigenetic Robotics (ICDL), 2013 IEEE Third Joint International Conference on*, pages 1–6, 2013.

[47] B. Webb. Using robots to model animals: a cricket test. *Robotics and Autonomous Systems*, 16(117–134), 1995.

[48] A. D. Wilson and S. Golonka. Embodied cognition is not what you think it is. *Frontiers in Psychology*, 4(58), 2013.

[49] M. Wilson. Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4):625–636, 2002.

[50] M. Wilson and G. Knoblich. The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131(3):460–473, 2005.

[51] H. Yanco and J. Drury. Classifying human-robot interaction: an updated taxonomy. In *Systems, Man and Cybernetics, 2004 IEEE International Conference on*, volume 3, pages 2841–2846 vol.3, Oct 2004.

[52] T. Ziemke. WhatâĂŹs that thing called embodiment. In *Proceedings of the 25th Annual meeting of the Cognitive Science Society*, pages 1305–1310, 2003.

[53] T. Ziemke. On the role of emotion in biological and robotic autonomy. *Biosystems*, 91(2):401 – 408, 2008. Modelling Autonomy Modelling Autonomy.

[54] T. Ziemke and J. Lindbolm. Some methodological issues in android science. *Interaction Studies*, 7(4):339–342, 2006.

[55] T. Ziemke and S. Thill. Robots are not embodied! conceptions of embodiment and their implications for social human-robot interaction. In *Proceedings of Robo-Philosophy 2014: Sociable robots and the future of social relations*, pages 49– 53. IOS Press BV, 2014.

[56] T. Ziemke, S. Thill, and D. Vernon. Embodiment is a double-edged sword in human-robot interaction: Ascribed vs. intrinsic intentionality. In *Cognition: a bridge between robotics and interaction. Workshop at HRI2015*, 2015.