

Helping Robots Imitate: Metrics and Technological Solutions Inspired by Human Behaviour

Aris Alissandrakis, Nuno Otero and Joe Saunders

Abstract—In this paper we describe three lines of research related to the issue of helping robots imitate people. These studies are based on observed human behaviour, technical metrics and implemented technical solutions. The three lines of research are: (a) a number of user studies that show how humans naturally tend to demonstrate a task for a robot to learn, (b) a formal approach to tackle the problem of what a robot should imitate, and (c) a technology-driven conceptual framework and technique, inspired by social learning theories, that addresses how a robot can be taught. In this merging exercise we will try to propose a way through this problem space, towards the design of a Human-Robot Interaction (HRI) system able to be taught by humans via demonstration.

I. INTRODUCTION

As robots increasingly integrate in human society, moving from their traditional use in the industry into assisting and collaborating in public and domestic spaces, new paradigms need to be developed towards more ‘natural’ forms of human-robot interaction avoiding the need, for example, for technical or scientific expertise from the human.

Social learning (in the sense of taking advantage of the presence of other agents), compared to individual learning (like trial-and-error), allows for task learning and skill transfer in more efficient and adaptive ways.

Imitation and social learning have been studied by psychologists and ethologists for over a century. Imitative learning of this type is defined and interpreted in many ways [1-4], however here we consider imitation in terms of the agent-based perspective (cf. [5]), and within this focus we address mainly the question of *what* [information is transmitted by teaching], and to a lesser degree the *how* [to reproduce the task]. The answers to the questions of *who* [to imitate], *when* [to imitate], and *how to evaluate the reproduction* are here assumed as given, or not directly addressed.

Concerning the *what* and the *how* issues within imitation, we consider three aspects: (a) the description and discussion of a number of user studies that show how humans naturally tend to demonstrate a task for a robot to learn, (b) a formal approach to tackle the problem of what a robot should imitate, and (c) a technology-driven conceptual framework and

technique, inspired by social learning theories that addresses how a robot can be taught.

The rest of this paper is structured as follows. Section II will give a summary of a number of user-studies conducted and discusses the lessons learned. This is followed by a formal approach briefly presented in section III which describes the search space of possible similarity metrics. These metrics can be used for evaluation and for planning an imitative response to a demonstration. This formal approach is complementary to the technical framework on how robots can be taught that is described in section IV. Section V offers some conclusions and future directions.

II. USER STUDIES

A. Overview of rationale and approaches followed

In relation to the overall conceptualisation of human-robot interactions and the demonstration/learning of everyday routine home tasks by robots, we have been considering a dialogical perspective where the system and the human need to establish common ground regarding their understanding of the tasks and on-going surrounding activities [6-11]. For example, the robot needs to ‘publicise’ its abilities, in order to be able to solicit further enhancements to the demonstration based on the ambiguous events at a timely and apposite manner. These responses from the robot should be tailored according to the ‘human perspective’, so that the human is able to interpret them, consider the robot’s difficulties and be, in turn, available to engage in further specification (or generation) of alternatives regarding the way an explanation unfolds. However, what are the specifics of this communication interface?

We have been conducting a series of user studies in order to understand how people naturally demonstrate some ‘basic’ home tasks to a robot. Up to the present, our overall strategy in addressing this problem has taken a step-by-step approach: we chose to introduce small changes (reaction/actions of the robot towards the human demonstration) from one study to the next whilst maintaining the overall experimental instruction requests to the human participants. More specifically, in the user studies described here, we were able to highlight the nature of human demonstrations of a specific routine home task: how to lay a table. This particular task was chosen because, while relatively simple and familiar to most people, the number of tableware involved, their order and their possible configurations on the table can vary depending on personal preferences.

Our interest in this line of research initiated within the

The first author is supported by a FY2007(P07369) Postdoctoral Fellowship for Foreign Researchers from the Japan Society for Promotion of Science (JSPS).

A. Alissandrakis is with Dept. of Computational Intelligence & Systems Science, Interdisciplinary Graduate School of Science & Engineering, Tokyo Institute of Technology, Japan alissandrakis@myk.dis.titech.ac.jp

N. Otero is with Dept. of Information Systems, University of Minho, Portugal nuno.otero@dsi.uminho.pt

J. Saunders is with the Adaptive Systems Research Group, University of Hertfordshire, UK j.l.saunders@herts.ac.uk

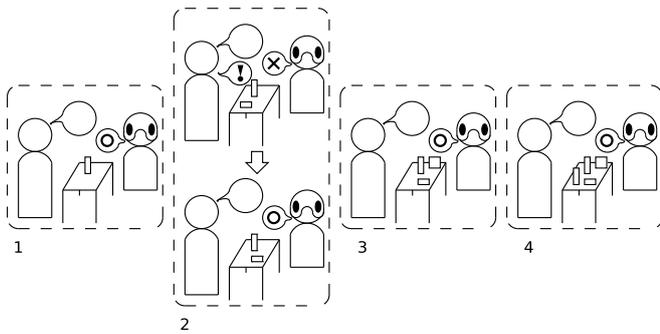


Fig. 1. **Investigating the participant's acknowledgment and response to the robot's feedback.** In the third and fourth studies, the robot would give positive feedback on the first, third and fourth instruction, but would initially state to not understand the second one (reverting back to positive if the participant clarified or repeated their demonstration). This also allowed to examine whether the participants remained consistent for the remaining instructions, in the case that they changed the way of demonstrating. (Figure originally from [10].)

COGNIRON project,¹ and two main goals set the course of our efforts: (a) we wanted to describe the frequency and duration of the different gestures people produce when performing routine home tasks, which in turn could be used to inform the development of appropriate algorithms allowing recognition of these gestures [12]; (b) to understand how to conceptualise and consequently design the robot in order to facilitate human-robot interactions, and as much as possible conform the robot to people's expectations and ideas of comfort and acceptance [13, 6].

Due to technological issues (current state-of-the-art robots are not yet able to detect and understand *unrestrained* behaviour by humans), plus the fact that for our purposes here the robot does not need to detect or respond to the actual participant's behaviour (as responses were predetermined), in all following studies the robots were not autonomous, but controlled using the Wizard-of-Oz methodology (cf. [14]).

In the first two studies, two experimental conditions were set up: the participants were asked to demonstrate how to lay the table using gestures only, or using gestures and speech (please see [6], for a detailed description). The rationale concerning the two conditions for these two exploratory studies was related to our additional interest in investigating to what extent the constraint of not using speech would affect the production of gestures. It could be the case that by restricting the speech people would make the gesturing more explicit, relative to the goals of the task. Both studies followed a within-subjects design, meaning that the same participants were used for the two conditions. The difference between these two studies lies in the role the robot had to perform: in the first study the robot was a passive observer of the participants' demonstration while in the second study the robot would acknowledge the actions/demonstration of the participants.

In the third study (please see [7, 8]), the general request

¹See <http://www.cogniron.org> for an overview of the project's goals and achievements.

to the participants was, again, similar to the previous studies – demonstrate to a robot how to lay a table. However, the robot's role changed. This time the robot not only acknowledged the actions/demonstration of the participant but, when the second object was placed on the table, the robot would state its misunderstanding (see Fig. 1). Furthermore, we also introduced an additional task: immediately after demonstrating to the robot, the participants were asked to view their own demonstrations and try to segment them (identify in the video of their demonstrations breakpoints of the activity) into meaningful chunks.

Two follow-up studies were carried out in Japan. The first one of these (please see [10]) was intended as a verification and extension of the third study. Here, the participants were asked to lay the table twice, using either Japanese or non-Japanese ('western') utensils. This was carried out in order to contrast two versions of the task, both known, but one more frequently practiced. In addition, a humanoid robot (Wakamaru, Mitsubishi Heavy Industries) was used rather than a robot with a mechanistic appearance (Peoplebot, MobileRobots Inc.) that was used in all previous studies.

In the second of those studies (please see [11]), the feedback from the robot was more elaborate; instead of the previously used "*I do (or sorry, I do not) understand [your demonstration]*", the robot would in turn expound what the participant just demonstrated, using speech and gestures, in either absolute or relative terms regarding the object localization. This explanation would be always correct regarding identifying the object, but either factually correct (in the sense of not necessarily using the same references that the participant used), or wrong, regarding the localization.

Summarising, the following set of questions gives a picture of the overall research issues being pursued in the user studies:

- Q1 Can we describe in a useful and efficient way the frequency and duration of different types of gestures when people are asked to explain to a robot how to perform a specific routine home task (e.g. lay a table)?
- Q2 Are there any differences regarding the frequency and/or duration of the different types of gestures produced when people are asked to use gestures only or are allowed to gesture and speak to explain this routine home task? Can the findings be used to engender new ways of considering human-robot interactions?
- Q3 Do participants clearly acknowledge the robot and change their explanations/demonstrations when the robot declares its inability to understand the participant's actions? What is the nature of this change (if any)? To what extent do people maintain the changes for the remainder of their explanation/demonstrations?
- Q4 What are the key moments that the participants consider to be dividing points of their own demonstrations? Does the feedback from the robot stating its misunderstanding at specific points of the demonstration alter the level of detail of the segmentation and corresponding demonstration effort?
- Q5 Do participants report difficulties regarding specific

#	Brief Summary	Research issues	Related publications
1	Participants instructed to use gestures only, or gestures and speech to demonstrate; robot was passive observer.	Q1, Q2	[6]
2	Participants instructed to use gestures only, or gestures and speech to demonstrate; robot would acknowledge the demonstration.	Q1, Q2	[6]
3	Participants relatively unrestrained on how to demonstrate; robot would state to understand or not the demonstration.	Q3, Q4, Q5	[7, 8]
4	Participants relatively unrestrained on how to demonstrate; robot would state to understand or not the demonstration; two versions of the task; humanoid instead of a mechanoid robot.	Q3, Q5, Q6	[10]
5	Participants relatively unrestrained on how to demonstrate; robot would give detailed feedback on understanding or not the demonstration, regarding object localisation; humanoid instead of a mechanoid robot.	Q3, Q5, Q7	[11]

TABLE I
USER-STUDIES QUICK REFERENCE TABLE

- events on their demonstrations? Do participants consider to have been influenced by the robot's feedback?
- Q6 Does the level of familiarity with the task influence decisively the nature of the demonstrations produced? If yes, in what ways?
- Q7 How and to what extent does more elaborate feedback from the robot regarding the object localization facilitate the interaction?

B. General Discussion of the Results from the User Studies

From the two initial studies [6], we were able to reflect on the need to produce a coding scheme for the analysis of gestures in human-robot interactions that could be readily and efficiently used to inform the development of appropriate gesture/activity recognition algorithms.

The conceptual framework presented in [15] was adapted to capture requirements for *contextual interpretation* of body postures and human activities for purposes of HRI. It defines five functional classes of gestures:

- **Manipulative gestures** These are gestures that involve the displacement of objects (e.g. picking a cup), or miming such displacements.
- **Symbolic gestures** These are gestures that follow a conventionalised signal. Their recognition is highly dependent on the context, both current task and cultural milieu (e.g. the thumbs up or thumb-index finger ring to convey "OK").
- **Interactional gestures** This category classifies gestures used to regulate interaction with a partner. These can be used to initiate, maintain, invite, synchronize, organize or terminate an interaction behaviour between agents (e.g. head nodding, hand gestures to encourage the communicator to continue).
- **Referencing/pointing gestures** (Deictics) The gestures that fall into this category are gestures used to indicate

objects or loci of interest.

- **Side effect of expressive behaviour** These are gestures that occur as side-effects of people's communicative behaviour. They can be motion with hands, arms, face, etc., but without specific interactive, communicative, symbolic or referential roles.
- **Irrelevant** These are gestures that do not have a primary communicative or interactive function, e.g. adjusting one's hair or rubbing the eye.

But it is important to note that certain gestures in particular situations might be multipurpose. Nehaniv et al. [15] stress the importance of knowing the context in which gestures are produced since it is crucial to disambiguate their meaning. In practice, data on the interaction history and context may help the classification process. This framework is intended to compliment existing and more detailed speech-focused classification systems (see for example [16–18]).

The concrete utilization within the COGNIRON project of the conceptual framework for the analysis of gestures referred to above and its practical instantiation as a coding scheme taught us that one important challenge is to adapt the level of granularity of the coding scheme to the sensory capability of the robot, without losing the meaning of the gestures produced. Further research is needed to produce some guidelines on how to do this adaptation from a practical point of view.

In relation to the actual gestural activities observed in the user studies, one relevant issue that clearly emerged from the analysis of the data collected in these two initial studies was the low frequency of any deictic (pointing) and symbolic gestures. In fact, the expectation was that the constraint of not being allowed to use speech would make people resort to pointing (e.g. to indicate the objects and corresponding locations) and use symbolic gestures to supplement their manipulative gestures (e.g. to mark the different steps of their explanations). The unexpected result may be due to the fact that people, when performing routine daily tasks, are not naturally likely to give detailed accounts of the way in which these tasks should be performed beyond the actual simple demonstration of how to accomplish it. Our studies also clearly suggest an interaction between the type of task and the type of gestures produced. This point stresses the importance of knowing the context in which gestures are produced and the interaction history [15]. Comparing the results from both studies, what seems more salient is the willingness of people to engage in interactions with the robot even if the trigger for such is just a 'small' acknowledgement of their actions. Moreover, people preferred the gestures and speech method of demonstrating. This choice is not surprising but it definitely supports the perspective that people might prefer to interact with robots in a 'natural' way [19]. The design challenge for current robotic systems is to be able to provide a fine balance between natural ways of interacting and the correct level of expectations generated in the final users concerning the system's capabilities.

In relation to the third study referred to previously, the actual demonstrations recorded suggest that people tried to

change their demonstrations when faced with the robot's misunderstanding but the alternatives were (sometimes) not any more informative than the original ones. People need more feedback about what to address specifically. In fact, we could observe some attempts to actively engage with the robot in a more complex interaction, i.e. participants seeming to probe the robot's abilities. Furthermore, people are not particularly consistent throughout their demonstration or modifications to it when faced with negative feedback.

Considering this specific issue on a related note, it is worth mentioning that Thomaz and Breazeal [20] make some observations about the way people tend to administer their own feedback when teaching a Reinforcement Learning agent:

- (a) they use the reward channel not only for feedback, but also for future-directed guidance;
- (b) they have a positive bias to their feedback, possibly using the signal as a motivational channel; and
- (c) they change their behaviour as they develop a mental model of the robotic learner.

These last points (about people's feedback) are outside the current scope of the work presented here (focusing more on the robot's feedback), but very much of interest in the broad context.

The results regarding the participants' segmentations of their own demonstrations of the home task suggest that people might differ regarding the level of detail they spontaneously consider. To what extent this result also actually implies different levels of detail in the explanation itself is still an on-going research question. However, it is worth noting that all the participants considered breakpoints corresponding to the end of an demonstration step (marked by the positive feedback from the robot). This suggests that these breakpoints mark the higher level of detail that the participants spontaneously considered. More surprisingly, the results indicate that participants did not segment the task differently when analysing the moment when the robot gave negative feedback, although in the post-session questionnaire the feedback was considered important for the segmentation.

The participants of the fourth study also acknowledged the robot's misunderstanding of their instructions, and responded by repeating their demonstration, modifying it to some extent. Together with the, as previously observed, lack of use of deictic gestures, this clarification did not vary significantly from the initial instruction (e.g. most simply used a louder voice as they thought the robot simply did not hear them). However, in the few cases where they actually modified their object localization (from absolute to relative reference and vice versa) to resolve the misunderstanding, they kept the same reference style for the remaining demonstrations; this consistency was not observed in the previous third study. Unfortunately, the majority of the participants remained inconsistent in a sense of mixing localization reference styles throughout their task demonstration. Any influence by the task familiarity (Japanese and non-Japanese style) was not observed in this study in terms of frequency of gestures, or object localization reference style; however, it was noted that

participants would comment on the particular function of the objects only for the Japanese utensils (e.g. "This is a soup bowl [...] it is used to eat *miso* soup.>").

At this time, only initial trial runs of the fifth study have been conducted, but they seem to indicate that the participants are favourable to the more detailed robot feedback, which in some cases made them adapt their own demonstrations. In this case, the name of the object (always correct) and its placement (sometimes wrong, both in terms of pointing to the table and using either absolute or relative referencing) are part of the feedback, effectively allowing the robot to 'publicise' its abilities of understanding the participant's demonstration.

III. A FORMAL APPROACH TO WHAT THE ROBOT SHOULD IMITATE

Alissandrakis' work on the correspondence problem (cf. [21]) and metrics for the robot's successful imitation of actions, states and effects [22, 23], establishes an initial formal stance (and partial solution) for the possibility of human robot interactions within the context of human teaching of robots. The correspondence problem effectively states that there can be multiple definitions and interpretations of what is 'similar' when trying to reproduce a behaviour, along the aspects of actions (e.g. body motions), states (e.g. body postures) and effects (e.g. positions, orientations and states of external objects, but also changes to the body-world relationship of the agent). There is essentially not a single 'goal', and the agent has to be able to decide, among the other four questions mentioned in the introduction section, which aspects to imitate. In an analogous fashion, an imitator/student has to be able to determine what aspect of the model/teacher's demonstration is important (and therefore constitutes the task knowledge).

In the table setting task used here, the most important aspect would be the effects, omitting in this case the teacher and student particular embodiments (which influence more the consideration of action and state metrics). Even so, there a number of possible effect metrics, taking into account (among other features) the displacement, orientation, and absolute or relative reference frame. The demonstrated task knowledge could be considered as any combination of these sub-aspects. However, if the initial configuration of the objects involved is dissimilar between the demonstration and the reproduction, different choices can lead to *qualitatively* dissimilar results.² In the context of task knowledge transfer, these possibly undesired results need to be avoided by having the teacher provide (possibly multiple) examples that satisfy only a small set of metrics, which in turn defines what is important.

Although useful for analysis from the 'system perspective', this research can not directly lead to a 'natural' HRI interface; we believe that the robot should be like a "black

²The issue here is that the chosen metrics may well be minimized, therefore resulting *quantitatively* in a successful imitation, however this choice of metrics by the system might not be the same as what the human intended while demonstrating her/his desired way to achieve the task.

box”, with no need for the interacting human to know details about the hardware and the inner data representations and structures. To that extent, we introduced a more detailed feedback on the fifth study, to examine whether the participants can adapt to the robot’s chosen style (dealing with either absolute or relative metrics), one that is implicitly expressed by its gestures and speech during the expounding of the just observed demonstration, rather than explicitly by identifying the metric in technical terms.

IV. INTEGRATING SOCIAL LEARNING ISSUES INTO A TECHNOLOGY-DRIVEN FRAMEWORK

In section II above it was suggested that, firstly, a key design challenge was to find a balance between natural ways of interacting with a robot and managing the expectations of human users. Secondly, that more feedback would be required from the robot to the human to indicate the robots success or otherwise in imitating a demonstrated task - simply giving negative feedback (e.g. ‘I don’t understand’) would be insufficient for the human to modify the task. Thirdly, segmentation of their own demonstrations to a robot suggested that individuals may differ as to the level of detail they spontaneously consider. In addressing these issues the robot faces many of the formalities described in section III above. Specifically these include finding a solution to the set of correspondence problems faced and deciding on the levels of granularity in the task set.

A partial technological solution to these issues has been addressed in [24] which describes a novel approach inspired by the notion of “zone of proximal development” initially proposed by Vygotsky [25, 26], and demonstrates how teaching a robot can be achieved through active manipulation of the robot’s actuators. Behavioural competencies are built step-by-step by the human, exploiting those competencies already taught. Such perspective is clearly in line with an embodied and situated perspective of cognition. The key issues that such an approach considers match those described earlier. In terms of the balance of natural interaction and expectations a mechanism that both amplifies the robots perceptions and provides a human-centred way of teaching is via the idea of ‘assisted’ or ‘self’-imitation [27]. Here the robot’s bodily configuration is physically manipulated by the human. This allows the robot to reconstruct motion paths having effectively been provided with a solution to the correspondence problem via the matching of human and robot actions. Amplification of the respective modalities can be achieved via information theoretic means whereby specific modalities which lead to the preferred outcome are favoured over those which do not. Similar mechanisms using ‘assisted imitation’ but employing a more model based approach are also used by Calinon and Billard [28].

Ideas of robot feedback are also explored by Saunders et al. [24]. In this instance the robot is effectively aware of what it already ‘knows’ in the form of a hierarchy of behavioural competencies. By exploiting these known competencies in a predictive mode (effectively acting as a forward model) the robot can assess when it is being taught

something that it already knows and can inform the human trainer appropriately. An extension to this idea would be when in a new and ill-described teaching event to respond with what it has understood so far in the process e.g. ‘I know how and where to place the knife but I have not understood the procedure after that’. Approaches to imitation have exploited combinations of inverse and forward models to model both observational and assisted imitation [29] and provide a computational mechanism which to some extent matches the mirror-neuron concepts [30] in neurophysiology.

In terms of spontaneous segmentation of tasks, we described in [7] how, in a robotic experiment (using the architecture described in [24]), that segmenting a behaviour not only allows task re-use and reduces human training time but can also serve to enhance the algorithmic effect in the environmental scaffolding process (in this case information gain). Thus the robot benefits with better performance, both on the task taught and from reduced computation. The human teacher also benefits as the number of training steps is significantly reduced as compared to a direct (i.e. teach everything at once) approach. However a further problematic issue (again explored via a robotic experiment using the system above [9]) considers situations where humans completely disregard steps which are crucial to the robot. Again a possible solution to this issue would be further feedback from the robot as suggested above.

Finally, Syrdal et al. [31] reports on interviews with non-expert end users, following the visualization of specific interaction episodes between humans and robots and suggests that people are receptive to this particular mode of interaction. In fact, the participants of this study reported not considering the physical manipulation of the robots actuators to be too far from their own experience of teaching children in particular situations. Of course, we are quite aware that commenting on a video might be very different to actually being engaged in the situation. However, the results reported are, at least, encouraging since there was no outright rejection. We do expect, though, that there will be some interaction between the levels of the actual acceptance of this mode of interaction, the interacting capabilities of the robot and also its physical appearance.

V. CONCLUSIONS AND FUTURE WORK

In this paper we have argued extensively that the robot will need to publicize its own abilities and understanding of ongoing tasks/activities in order to facilitate the interaction with a human partner. Considering this specific problem, however, one of the challenges related to teaching/learning/imitation tasks is to know what kind of initial building blocks need to be in place in order to make the robot’s actions understandable by the human partner. At the same time, though, from the system’s point of view, these building blocks have to be easily extended or adapted to different task needs. Furthermore, depending on the concrete social situations (for example, solo human teaching or multiple instructors, resilience to interruptions and breakdowns during teaching

episodes) and given the wide variability of human's behavioural repertoire, the robot's communication abilities will probably need to be tuned to the audience and/or recipient(s) (similar point made in [9]).

The problem space is vast and complex and we would like to highlight that the achievements in its exploration will definitely have to consider integrative frameworks, where user studies, formal/conceptual proposals and technical solutions need to be taken into careful consideration. In fact, some reflection needs to be made on how particular results impact on the perception of the technology to achieve concrete research goals, but also its usefulness, worth and values in society (see for example [32]). HRI is in its infancy and, like other digital technology advancements, easily catches the public eye. However, what we should avoid is what Fernaeus et al. [33] considered to be a 'robot cargo cult' phenomena, where research claims or suggestions raise expectations beyond reasonable heights, influencing the social representations people/laymen have of the field, maybe, in unexpected and undesirable ways.

REFERENCES

- [1] T. R. Zentall and B. G. Galef, Jr., Eds., *Social Learning: Psychological and Biological Perspectives*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1988.
- [2] T. R. Zentall, "An analysis of imitative learning in animals," in *Social Learning in Animals: The Roots of Culture*, B. G. Galef and C. M. Heyes, Eds. Academic Press, New York, 1996, pp. 221–243.
- [3] —, "Imitation in animals: Evidence, function and mechanisms," *Cybernetics and Systems*, vol. 32, no. 1–2, pp. 53–96, 2001.
- [4] B. G. Galef and C. M. Heyes, Eds., *Social Learning in Animals: The Roots of Culture*. Academic Press, 1996.
- [5] K. Dautenhahn and C. L. Nehaniv, "An agent-based perspective on imitation," in *Imitation in Animals and Artifacts*, K. Dautenhahn and C. L. Nehaniv, Eds. MIT Press, 2002, pp. 1–40.
- [6] N. Otero, C. L. Nehaniv, D. S. Syrdal, and K. Dautenhahn, "Naturally occurring gestures in a human-robot teaching scenario," *Interaction Studies*, vol. 9, no. 3, pp. 519–550, 2008.
- [7] N. Otero, J. Saunders, K. Dautenhahn, and C. L. Nehaniv, "Teaching robot companions: The role of scaffolding and event structuring," *Connection Science*, vol. 20, pp. 111–134, June 2008.
- [8] N. Otero, A. Alissandrakis, K. Dautenhahn, C. L. Nehaniv, D. S. Syrdal, and K. L. Koay, "Human to robot demonstrations of routine home tasks: Exploring the role of the robot's feedback," in *Proceedings of the 3rd ACM/IEEE International Conference On Human-Robot Interaction*. Amsterdam, The Netherlands: ACM, 2008, pp. 177–184.
- [9] J. Saunders, N. Otero, and C. Nehaniv, "Issues in human/robot task structuring and teaching," in *Robot and Human interactive Communication, 2007. RO-MAN 2007. The 16th IEEE International Symposium on*, Aug. 2007, pp. 708–713.
- [10] A. Alissandrakis and Y. Miyake, "Human to robot demonstrations of routine home tasks: Acknowledgment and response to the robot's feedback," in *New Frontiers in Human-Robot Interaction, (AISB 2009)*, April 2009, pp. 9–15.
- [11] —, "Human to robot demonstrations of routine home tasks: Adaptation to robot's preferred style of demonstration," in *Robot and Human interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*, Sept. 2009.
- [12] N. Otero, S. Knoop, C. L. Nehaniv, D. S. Syrdal, K. Dautenhahn, and R. Dillman, "Distribution and recognition of gestures in human-robot interaction," in *Robot and Human Interactive Communication, 2006. RO-MAN 2006. The 15th IEEE International Symposium on*, Sept. 2006, pp. 103–110.
- [13] N. Otero, C. L. Nehaniv, D. S. Syrdal, and K. Dautenhahn, "Naturally occurring gestures in a Human-Robot teaching scenario," in *Robot and Human Interactive Communication, 2006. RO-MAN 2006. The 15th IEEE International Symposium on*, Sept. 2006, pp. 533–540.
- [14] N. Dahlbäck, A. Jönsson, and L. Ahrenberg, "Wizard of Oz studies – why and how," in *Readings In Intelligent User Interfaces*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1998, pp. 610–619.
- [15] C. Nehaniv, K. Dautenhahn, J. Kubacki, M. Haegele, C. Parltitz, and R. Alami, "A methodological approach relating the classification of gesture to identification of human intent in the context of human-robot interaction," in *Robot and Human interactive Communication, 2005. RO-MAN 2005. The 14th IEEE International Symposium on*, Aug. 2005, pp. 371–377.
- [16] J. Cassell, "Nudge nudge wink wink: Elements of face-to-face conversation for embodied conversational agents," in *Embodied Conversational Agents*, J. Cassell et al., Ed. Cambridge, MA, USA: MIT Press, 2000, pp. 1–27.
- [17] J. Cassell, S. Kopp, P. Tepper, K. Ferrimanand, and K. Striegnitz, "Trading spaces: How humans and humanoids use speech and gesture to give directions," in *Conversational Informatics*, T. Nishida, Ed. New York: John Wiley & Sons, 2007.
- [18] D. McNeill, *Hand and Mind*. Chicago: University of Chicago Press, 1992.
- [19] K. Dautenhahn, "The art of designing socially intelligent agents: Science, fiction and the human in the loop," *Applied Artificial Intelligence Journal, Special Issue on Socially Intelligent Agents*, vol. 12, no. 7–8, pp. 573–617, 1998.
- [20] A. L. Thomaz and C. Breazeal, "Teachable robots: Understanding human teaching behavior to build more effective robot learners," *Artif. Intell.*, vol. 172, no. 6–7, pp. 716–737, 2008.
- [21] C. L. Nehaniv and K. Dautenhahn, "The correspondence problem," in *Imitation in Animals and Artifacts*, K. Dautenhahn and C. L. Nehaniv, Eds. MIT Press, 2002, pp. 41–61.
- [22] A. Alissandrakis, C. Nehaniv, and K. Dautenhahn, "Action, state and effect metrics for robot imitation," in *Robot and Human Interactive Communication, 2006. RO-MAN 2006. The 15th IEEE International Symposium on*, Sept. 2006, pp. 232–237.
- [23] —, "Correspondence mapping induced state and action metrics for robotic imitation," *Systems, Man, and Cybernetics, Part B, IEEE Transactions on*, vol. 37, no. 2, pp. 299–307, April 2007.
- [24] J. Saunders, C. L. Nehaniv, K. Dautenhahn, and A. Alissandrakis, "Self-imitation and environmental scaffolding for robot teaching," *International Journal of Advanced Robotics Systems, Special Issue on Human-Robot Interaction*, vol. 4, no. 1, pp. 109–124, March 2007.
- [25] L. S. Vygotsky, *Mind in Society: Development of Higher Psychological Processes*, new ed. Harvard University Press, 1978.
- [26] —, *Thought and Language*, 2nd ed. MIT Press, 1986.
- [27] P. Zukow-Goldring and M. A. Arbib, "Affordances, effectivities and assisted imitation: Caregivers and the directing of attention," *Neuro-computing*, vol. 70, no. 13–15, 2007.
- [28] S. Calinon, F. Guenter, and A. Billard, "On learning, representing and generalising a task in a humanoid robot," *IEEE Transactions on Systems, Man and Cybernetics, Part B. Special Issue on robot learning by observation, demonstration and imitation*, vol. 35:5, 2006.
- [29] J. Demiris and G. Hayes, "Imitation as a dual-route process featuring predictive and learning components: A biologically-plausible computational model," in *Imitation in Animals and Artifacts*, K. Dautenhahn and C. L. Nehaniv, Eds. MIT Press, 2002, pp. 327–361.
- [30] V. Gallese and A. Goldman, "Mirror neurons and the simulation theory of mind-reading," *Trends in Cognitive Sciences*, vol. 2, no. 12, pp. 493–501, 1998.
- [31] D. S. Syrdal, N. Otero, and K. Dautenhahn, "Video prototyping in human-robot interaction: Results from a qualitative study," in *Proceedings Of The 15th European Conference On Cognitive Ergonomics: The Ergonomics Of Cool Interaction*. Funchal, Portugal: ACM, 2008, pp. 1–8.
- [32] A. Sellen, Y. Rogers, R. Harper, and T. Rodden, "Reflecting human values in the digital age," *Commun. ACM*, vol. 52, no. 3, pp. 58–66, 2009.
- [33] Y. Fernaeus, M. Jacobsson, S. Ljungblad, and L. E. Holmquist, "Are we living in a robot cargo cult?" in *Proceedings of the 4th ACM/IEEE International Conference On Human-Robot Interaction*. La Jolla, California, USA: ACM, 2009, pp. 279–280.