Personalised Self-Explanation by Robots: The Role of Goals versus Beliefs in Robot-Action Explanation for Children and Adults

Frank Kaptein[1], Joost Broekens[1], Koen Hindriks[1], and Mark Neerincx[1,2]

***Abstract*— A good explanation takes the user who is receiving the explanation into account. We aim to get a better understanding of user preferences and the differences between children and adults who receive explanations from a robot. We implemented a Nao-robot as a belief-desire-intention (BDI)-based agent and explained its actions using two different explanation styles. Both are based on how humans explain and justify their actions to each other. One explanation style communicates the *beliefs* that give context information on why the agent performed the action. The other explanation style communicates the *goals* that inform the user of the agent's desired state when performing the action. We conducted a user study (19 children, 19 adults) in which a Nao-robot performed actions to support type 1 diabetes mellitus management. We investigated the preference of children and adults for goal- versus belief-based action explanations. From this, we learned that adults have a significantly higher tendency to prefer goal-based action explanations. This work is a necessary step in addressing the challenge of providing *personalised explanations* in human-robot and human-agent interaction.**

I. Introduction

Explainable Artificial Intelligence (XAI) is the capability of a system to explain its own behaviour. XAI is known to have a positive influence on user trust in and understanding of the intelligent systems [1–3]. Intelligent systems are becoming increasingly complex, which makes it difficult for the users to understand the system's actions [4]. XAI is important in areas such as medical support [1], fire-fighting [5, 6], and education [3].

A theoretical approach towards explaining actions is the *intentional stance*. When adopting the intentional stance, one assumes that actions result from intentions of the *actor* (i.e., the human or agent performing the action) [7]. In everyday human communication, two common explanation styles for intentional actions are goal-based and belief-based explanations [8]. A goal-based explanation communicates the actor's desired outcome of the action. A belief-based explanation provides information about the context and the circumstances that caused the actor to choose one action over another. How humans explain an actor's actions in everyday communication is referred to as *folk psychology* [9, 10], and has been used in developing XAI for BDI-based (belief, desire intention) agents [6, 11–13].

A good explanation is *personalised*, i.e., it takes the user that is receiving the explanation into account. As we mature, we develop our capabilities to create and understand explanations for someone else's actions [8, 14, 15]. Furthermore, different educational strategies are required for adults and children [16, 17]. Therefore, self-explaining robots and agents that educate their users are likely to need different explanation strategies for children and adults.

To address this challenge of personalising explanations, we compared two different explanation styles on two different user groups. We constructed goal-based and belief-based robot-action explanations. We then asked both children and adults what explanation best helped them to understand the different actions. We tested whether these different explanation styles significantly differ on this metric, taking user group into account as a factor.

The context of this paper is the PAL (a Personal Assistant for a Healthy Lifestyle) project. This project helps children (aged 7-14) to cope with type 1 diabetes mellitus. In this project, we develop an agent controlling a Nao-robot or its virtual avatar. The system autonomously interacts with the children and their parents for prolonged periods of time. It helps the children to cope with their medical health issues. Therefore, it is important that the different users trust and understand the actions of the PAL-agent. To facilitate this, we develop the capability to explain these actions to the different users. Consider the following robot action: '*the PAL-robot tells the child that one has a hypo when one's blood glucose level is below 4.0 millimoles per litre*'. We may explain this by saying the robot wants to teach the child how to detect and treat a hypo (goal); or, that it thinks the child does not know what blood measurements indicate that one has a hypo (belief).

We will first review related work in the field of XAI, and, in particular, work that focuses on self-explaining agents in Section II. Then, in Section III, we describe a generally applicable representation of a BDI-based agent's decision making and how we derive belief-based and goal-based explanations from this. We explain the set-up of our experiment in Section IV. Finally, we present the results and discuss them.

II. Motivation for Research Conducted

When developing intelligent agents, one should consider enhancing them with self-explaining capabilities. Previous studies have shown that XAI enhances user trust in and understanding of intelligent systems [1–3, 18, 19]. This is especially important for intelligent agents because they are often designed to operate semi-autonomously, and they often operate in consequential domains like medicine or military [20].

*The PAL project is supported by Horizon2020 grant nr. 643783-RIA.

[1]Department of Intelligent Systems, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands F.C.A.Kaptein@tudelft.nl

[2]TNO, Postbus 23, 3769 ZG Soesterberg, The Netherlands

Recent work on XAI for intelligent agents used automatically generated folk psychology based explanations [6, 11, 12]. A folk psychology based explanation communicates the beliefs and goals that led to the agent's behaviour. One adopts the *intentional stance*, meaning one explains the agent's action by explaining the *reasons* (beliefs and goals) for the agent's intention [7]. Example folk psychology based explanations are, 'I proposed to play a sorting game together because I thought that you liked that game' (belief); or, 'I proposed to play a sorting game together because I wanted to play a game with you' (goal). Folk psychology based explanations provide concise, human-like explanations of an agent's actions. They are mainly directed at the end-users of the intelligent system (see, e.g., [13, 21] for explanations more directed towards agent developers).

Previous work on XAI took user knowledge into account [22, 23]. These studies classified a user as a beginner or expert and used this to provide explanations that better fit the individual user's preferences. However, we belief more elaborate user models are required for good personalised explanations. In this paper, we compare two common explanation styles for two distinct user groups: children and adults.

A. Goal-based and Belief-based Explanations

Folk psychology is how humans in everyday communication explain and predict intentional actions [8, 9]. Two common explanations styles in folk psychology are goal-based and belief-based explanations [7–10]. A goal-based explanation communicates the actor's desired outcome of the action. It provides an answer to the questions, 'To what end?' or 'For what purpose?' A belief-based explanation provides information on why the actor chose a certain action over another. It provides information about the context and the circumstances. Goals are easier to infer from the action itself, whereas beliefs provide information specific to the particular actor that performed the action and context in which the action was performed. Malle [8] writes that to infer an actor's belief, one needs to take the *perspective* of this particular actor.

B. Hypothesis

In this paper, we define two explanation algorithms: one that always provides the triggering condition (belief) that caused the agent to perform the action and one that always provides the parent goal that the agent is trying to achieve. We test which explanation algorithm is preferred by adults and children by presenting them example actions explained by these algorithms. Our hypothesis is:

Hypothesis. *Adults have a stronger preference than children for goal-based over belief-based explanations.*

There is psychological support for this hypothesis. First, a difference in itself is likely because explanations based on folk psychology change as humans mature [8, 15]. For example, young children (4 years old) have trouble realising someone may have a belief that is false [14]. Second, children and adults alike are inclined to belief that others have similar beliefs and knowledge as they do [15]. However, adults have accumulated a vast amount of knowledge to which they can link new information [17]. Third, adults strongly desire (more than children) to know the goals you are pursuing when educating them [16, 17].

III. Goal Hierarchy Trees

Previous work on XAI showed how one can use an agent's beliefs and goals for generating action explanations [12, 24]. However, this means that one should take special care in designing and formulating the beliefs and goals of the agent [12, 25]. Previous work proposed the use of a goal hierarchy tree (GHT) to develop a high level design of the agent's reasoning [6, 11] and provides guidelines for their development [12]. GHTs are based on hierarchical task analysis, a technique from cognitive psychology used to specify complex human tasks [26]. In this section, we describe the structure of a GHT, and how we can construct explanations from it. We adopt GHTs as agent design and use these to test how one can personalise explanations for different users.

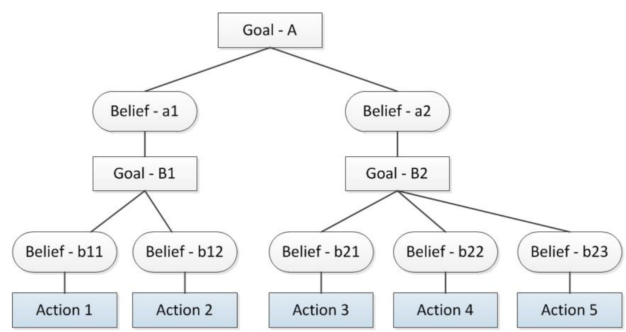

Fig. 1. The square nodes are goals the agent adopts. The top node is the agent's main goal. When following the edges, sub-goals are represented that the agent adopts to achieve the main goal. Actions to achieve (sub-)goals are shown as shaded square nodes (the leaves) of the tree. Conditions (beliefs) that cause an agent to adopt a sub-goal or perform a particular action are shown as rounded nodes.

A. The Structure of a Goal Hierarchy Tree

Figure 1 shows the structure of a GHT. A BDI-based agent that runs in accordance with a GHT chooses actions as follows. Based on the agent's current goals and beliefs, it chooses an action to perform. If multiple actions are applicable, then the agent randomly chooses one. When no actions are applicable then the agent remains idle until its beliefs change, which can cause it to adopt new goals and can make new actions become applicable.

A GHT does not model what external events can occur and how events and agent-actions cause the agent to update its beliefs. Rather, the GHT shows a high-level design of the agent's *reasoning*, i.e., what action it should perform given a current state of beliefs and goals. This is sufficient for our purpose of generating explanations based on an agent's reasoning. However, if one wants to run a BDI-based agent that acts in accordance with the GHT, then this additional modelling is required.

B. *Goal-based and Belief-based Agent-action Explanations*

One can explain an action by means of the goal one aims to achieve, or by explaining why it was possible to perform the action (belief) [8, 10]. Explanations should not be too long [12, 15]. We need to be selective in what beliefs and goals we communicate. In this paper we use the following explanation algorithms.

The **belief-based explanation algorithm** selects the belief directly above the action (triggering condition). For example, action-2 is explained by Belief-b12 and action-3 by Belief-b21 (Figure 1).

The **goal-based explanation algorithm** selects the goal directly above the action (parent goal). For example, action-2 is explained by Goal-B1 and action-3 by Goal-B2 (Figure 1).

We use these algorithms for the sake of testing the hypothesis of Section II-B. Both explanation algorithms consist of one element (one belief or one goal) in the goal hierarchy tree. The goal-based explanation provides the most direct response to the question, ‘What is your purpose?’ The belief-based explanation provides the most direct answer to why the agent chose a particular action over another. Thus, they closely resemble how humans explain their actions as discussed in Section II-A. Furthermore, they are short and thus unlikely to flood the receiver of the explanation with too much information [12, 15]. Therefore, we can use these algorithms to determine a preference for belief-based or goal-based agent-action explanations in a general way.

IV. User Study

We developed a GHT within the context of the PAL-project and set up an experiment using the explanation algorithms from Section III-B. We tested whether goal-based or belief-based action explanations are better received by the participants (children and adults). We tested for a significant difference in preference within and between these user groups.

We belief the PAL-project is a good domain for testing XAI. Firstly, it provides us with both children and adults that interact with the agent; and secondly, it is an exemplary, consequential domain were the agent interacts with its users for prolonged periods of time (the type of domains where XAI is especially important [20]).

A. *Participants*

The participants were recruited from a diabetes camp for children. Children diagnosed with type 1 diabetes mellitus were recruited by the Dutch Diabetes Association DVN. These children and their parents were invited to participate in our experiment. Rejecting this had no influence on participation in other activities during the camp.

In total, there were 21 children and 20 adults (parents of the children) present at the camp. One child did not participate in the experiment. One child participated, but was looking over his friend’s shoulder while filling in answers. One adult did not fill in the initial sheet asking for data like age, gender, and education. This left us with 19 children (12 male, aged 8-11) and 19 adults (8 male, aged 35-48).

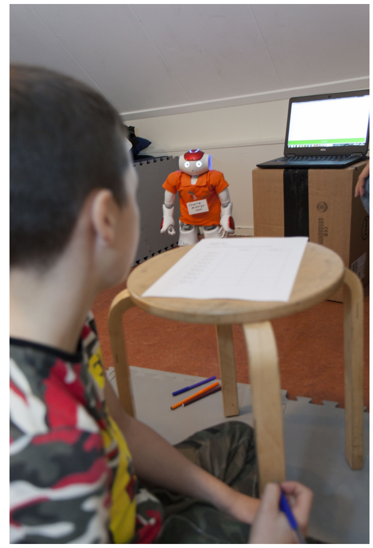

Fig. 2. Set-up of the experiment. The Nao-robot verbally presents example scenarios and provides two explanations for each of these. The screen textually shows what the Nao-robot is saying, so the child can always read-back on the screen what happened. The child then puts a mark at the most preferred explanation.

B. *Designing a Goal Hierarchy Tree*

Figure 3 presents our design of a GHT for our agent. This GHT is a translation from its Dutch counterpart, since the experiment was performed in the Netherlands. Based on this GHT, the agent chose different actions to support a child in diabetes management.

The GHT specifies two styles of support. The agents aims to educate the child when the child is in a good mood to learn new things. The agent aims to cheer up the child when the child is sad. We call this *cognitive support*, and *affective support*, respectively. The way that this agent provides these types of support is defined by ontologies that were developed in cooperation with health-care professionals [27]. The here developed GHT resembles the treatment plan provided by these experts.

C. *Set-up & Materials*

The GHT shown in Section IV-B has nine different robot actions. These actions can all be explained by using the belief-based explanation algorithm or by using the goal-based explanation algorithm. A Nao-robot presented all the actions to the participants. For each action, it proposed two explanations obtained from the algorithms. The participants had a forced choice to prefer either one of the proposed explanations.

The robot was located in front of the participants and next to a laptop screen (Figure 2). For every action and corresponding explanations presented by the Nao-robot, the

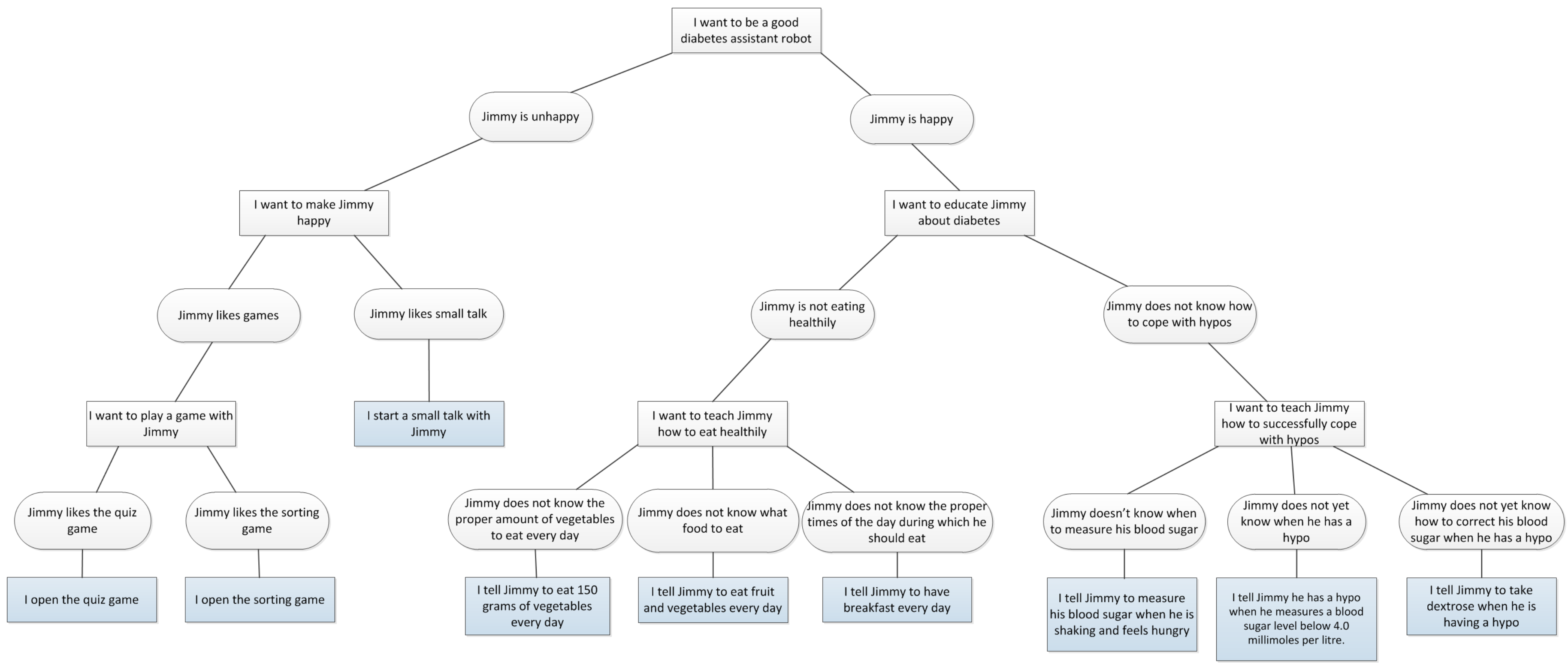

Fig. 3. In this figure, the goal hierarchy tree of the PAL agent is shown. The rectangular nodes are goals the agent adopts. The top node is the agent's main goal. When following the edges, sub-goals are represented that the agent adopts in order to achieve the main goal. The triggering conditions (beliefs) that determine whether the agent should adopt a sub-goal are represented in rounded nodes. The agent's actual actions are the shaded nodes (the leaves) of the tree.

TABLE I

DISTRIBUTION OF CHILDREN AND ADULTS OVER THE 4 CONDITIONS

	Random Seed of Scenarios			
	Normal Order of Explanations		Reversed Order of Explanations	
	Normal Order of Scenarios	Reversed Order of Scenarios	Normal Order of Scenarios	Reversed Order of Scenarios
Children	6	5	4	4
Adults	5	4	5	5

laptop screen showed the action performed and explanations provided. In this way, the participants could read back what the robot said. For example, the screen could look like:

Action: 'I tell Jimmy to take dextrose when he is having a hypo.'
Explanation 1: 'Jimmy does not yet know how to correct his blood-sugar when he has a hypo.'
Explanation 2: 'I want to teach Jimmy how to successfully cope with hypos.'

By using the robot, the children experience the experiment as a fun activity rather than a chore. Robots have been shown to have a positive impact on motivation and learning [28]. We chose to also use a screen, since Nao-robots do not always pronounce words very well. By using a screen, the participants can always read back what the robot said.

D. Variables & Design

The presentation of an action including the two explanations is henceforth referred to as a *scenario*. A scenario starts with the robot saying: 'I performed action a'. With a being one of the actions in the GHT and phrased exactly as shown in the GHT (Figure 3). Then, the robot says: 'How should I explain this? One: $explanation$-1; or two, $explanation$-2'. Where explanation 1 and 2 are the belief-based explanation and the goal-based explanation.

A participant is shown nine scenarios, one for each action, in random order. Whether the participant was shown the belief-based explanation first or the goal-based explanation first was also chosen randomly for every scenario. After the experiment, we counted the *percentage of scenarios where the participant preferred a goal-based explanation*. So, if the participant preferred the belief-based explanation in six scenarios and the goal-based explanation in the other three scenarios, then this variable is 33% for the participant.

Due to the camp setting, we were not able to do the experiment with every user separately. We were forced to have the participants do the experiment in small groups (group size of 2-3 for the children, 4-5 for the adults). The individuals in the groups were not allowed to discuss amongst each other nor look at each other's answers before the experiment was over. However, a consequence of having groups was that participants in the same group also saw the same order of actions and the same order of explanations. Thus, we counterbalanced the conditions. We produced a single random seed of scenarios. The actions were first put in random order. For every action separately, the system then randomly chose explanation 1 to be belief-based and explanation 2 goal-based, or vice versa. We counterbalanced this among the participants. I.e., the participants saw the actions in this randomly chosen order, or they saw the actions in reversed order. Furthermore, they saw the order of the explanations in this randomly chosen order, or they saw the explanations for these actions in reversed order. The participants were evenly distributed over these 4 conditions (see Table I).

E. Procedure

In small groups, the participants were asked to enter the room and were seated in front of the robot and laptop. The researcher informed the participants that he would remain present during the experiment but that the robot would guide the experiment. Additional questions could be directed to the researcher.

The Nao-robot started the experiment with a small presentation. Here, it told the participants that it wants to learn how to explain its behaviour to them and that it needs their input. It explained that it will provide example scenarios where it helped a fictional child 'Jimmy' to deal with diabetes. In this starting presentation, the robot said it sometimes plays a game with Jimmy and sometimes tries to educate Jimmy concerning diabetes management. The robot said that in all the example scenarios it wants to explain its action and always considers two possible explanations. The participants were then asked to select the explanation that best helped them to understand why the PAL-robot performed that action.

After the presentation the robot verbally presented the nine scenarios, one for each action, and the screen showed the scenarios in text. With the children, the researcher paid special attention to prevent them from looking over each other's shoulder while choosing the best explanation. Once the experiment was finished the robot and researcher thanked the participants for their help.

V. RESULTS

To test the preference towards the different explanation styles, we counted the percentage of scenarios where the participants preferred a goal-based explanation. A one-sample Wilcoxon signed rank test shows that the median of preferring goal-based explanations, rather than belief-based explanations, is significantly above 50% for children ($med = 0.667, 95\%\ CI = [0.667, 0.778],\ p = .007$) and adults ($med = 0.778, 95\%\ CI = [0.667, 1.0],\ p < .000$). So, both user groups significantly prefer goal-based explanations over belief-based explanations. Figure 4 shows the distributions of preferring goal-based explanations for children and adults.

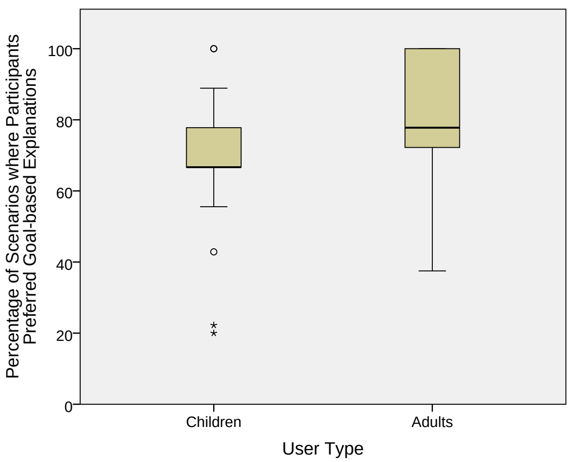

Fig. 4. A box plot showing the distribution of preferring goal-based explanations over belief-based explanations. On the x-axis the two user types (children and adults) are depicted. The y-axis shows the percentage of scenarios where the subject preferred the goal-based explanation over the belief-based explanation. Adults have a significantly higher preference for goal-based action explanations (Median = 0.778) than children (Median = 0.667).

Furthermore, a Mann-Whitney test indicated that the preference towards goal-based explanations, rather than belief-based explanations, was greater for adults ($med = 0.778$) than for children ($med = 0.667$), U = 112.5, p = .042, r = .33. Adults prefer goal-based explanations significantly more than children.

VI. Discussion

The results in the previous section show that there is a significant preference for goal-based explanations in both user groups. However, it would be premature to state that goal-based explanations are always preferable. Previous studies have shown contradicting findings on this subject. The work in [12] analyses three studies that provide explanations based on a goal hierarchy tree. Two studies in a firefighting domain [5, 6], (one with experts and one with laymen) and one in a cooking domain [11] (where users were perceived as experts, since they all knew how to cook). In the non-expert domain, the participants showed a preference for belief-based explanations. In the expert domains, goal-based explanations were preferred. It is hypothesized in [12] that this difference may be due to the expert level. Since both the children (who have diabetes mellitus themselves) and the adults (their parents) are familiar with the domain, it can be expected that these users would prefer goal-based explanations. Future work should further explore this by testing this domain on layman (adult & child) users and comparing the results with the here presented findings.

Another finding in the results is that adults significantly prefer goal-based explanations more than their children. This has two possible explanations. First, according to adult learning psychology, adults need to know the objectives of the instruction [16, 17]. They are goal-oriented learners that rely on their vast personal experience. They prefer to know how instructions help them to enhance their existing abilities, rather than children who learn under the assumption that all instructions will help them sometime in the future [16]. Adults thus prefer knowing the objectives (goals) that the robot is pursuing when performing actions to educate its user.

The second explanation is that children are more motivated to understand a robot character. To infer an agent's belief, one needs to take the *perspective* of this particular agent [8]. Children and adults alike are better at perspective taking when their motivation to do this is high [29]. A higher motivation then correlates with a better understanding of belief-based explanations. On the other hand, adults are better adapted to perspective taking than children. Adults are faster at adjusting when they learn their initial perspective is incorrect [29]. Being an adult then also suggests a better understanding of belief-based explanations, since the adult is more flexible at adjusting her perspective to match that of the agent. In conclusion, if perspective taking is the explanation for these results, then motivation of the children must have had a stronger influence than the flexibility of the adults.

There are three sources that potentially limit the generalisability of our results. First, the chosen scenarios depicted in the GHT (Figure 3) potentially have traits that we are unaware of but that influence the preference for an explanation style. Second, the PAL-agent aims to educate its users on type 1 diabetes mellitus and maintain a positive mood in the user. A domain that resembles this type of system behaviour may be more likely to find similar preferences for explanation styles. Third, the children and adults may also have traits that are not representative for the entire population (e.g., culture) and that influence the preference for an explanation style. There are, however, many similarities between our user groups (i.e., they both face the problem of managing diabetes, they work with the same caregivers at the same hospitals, they even share the same genes). Within our sample space, we therefore belief that child/ adult is the only factor responsible for this difference in preferred explanation style. However, to address this issue of generalisability, future work includes replicating our study with more divers scenarios, contexts, and users.

The presented experiment was a start. In future work, we will systematically expand the design space. For example, individual user preferences and differences in the context in which the agent performed the action can also have an influence on how one should construct the explanation. Furthermore, one can combine goal-based and belief-based explanations providing the user with more information. However, explanations should not become too long [15]. The explainer should thus be careful with when and how to add further information to an explanation. Finally, we tested the subjective preference towards the explanation styles. A next step is to test how this influences user behaviour and trust in the system.

VII. Conclusion

In this paper we compared the preference for goal-based versus belief-based action explanations between two user groups. We presented children and adults with a set of example robot actions and provided two possible explanations for these. Belief-based explanations communicated the context (a belief) preceding the decision to perform an action. Goal-based explanations provided the agent's purpose (a goal) of the action. The users were asked to choose the explanation that *best helped them to understand why the Nao-robot performed this action*.

We found that adults have a significantly *higher* preference for goal-based explanations than children. This is the first evidence that self-explanations of intelligent agents are perceived differently by children and adults. This work is a necessary step towards providing *personalised explanations* in human-robot and human-agent interaction.

REFERENCES

[1] E. H. Shortliffe, R. Davis, S. G. Axline, B. G. Buchanan, C. C. Green, and S. N. Cohen, "Computer-based consultations in clinical therapeutics: explanation and rule acquisition capabilities of the mycin system," *Computers and biomedical research*, vol. 8, no. 4, pp. 303–320, 1975.

[2] D. N. Lam and K. S. Barber, "Comprehending agent software," in *Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems*. ACM, 2005, pp. 586–593.

[3] C. Conati and K. VanLehn, "Providing adaptive support to the understanding of instructional material," in *Proceedings of the 6th international conference on Intelligent user interfaces*. ACM, 2001, pp. 41–47.

[4] M. G. Core, H. C. Lane, M. Van Lent, D. Gomboc, S. Solomon, and M. Rosenberg, "Building explainable artificial intelligence systems," in *Proceedings of the national conference on artificial intelligence*, vol. 21, no. 2. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2006, p. 1766.

[5] M. Harbers, K. van den Bosch, and J.-J. C. Meyer, "A study into preferred explanations of virtual agent behavior," in *International Workshop on Intelligent Virtual Agents*. Springer, 2009, pp. 132–145.

[6] M. Harbers, K. Van den Bosch, and J.-J. Meyer, "Design and evaluation of explainable bdi agents," in *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on*, vol. 2. IEEE, 2010, pp. 125–132.

[7] D. C. Dennett, "Three kinds of intentional psychology," *Perspectives in the Philosophy of Language: A Concise Anthology*, pp. 163–186, 1978.

[8] B. F. Malle, "How the mind explains behavior," *Folk Explanation, Meaning and Social Interaction. Massachusetts: MIT-Press*, 2004.

[9] P. M. Churchland, "Folk psychology and the explanation of human behavior," *Philosophical Perspectives*, vol. 3, pp. 225–241, 1989.

[10] B. F. Malle, "How people explain behavior: A new theoretical framework," *Personality and social psychology review*, vol. 3, no. 1, pp. 23–48, 1999.

[11] J. Broekens, M. Harbers, K. Hindriks, K. Van Den Bosch, C. Jonker, and J. J. Meyer, "Do you get it? user-evaluated explainable bdi agents," in *Multiagent System Technologies MATES*, vol. 6251 LNAI. Springer Berlin Heidelberg, 2010, pp. 28–39. [Online]. Available: http://link.springer.com/10.1007/978-3-642-16178-0_5

[12] M. Harbers, J. Broekens, K. Van Den Bosch, and J.-J. Meyer, "Guidelines for developing explainable cognitive models," in *Proceedings of ICCM*. Citeseer, 2010, pp. 85–90.

[13] K. V. Hindriks, "Debugging is explaining," in *International Conference on Principles and Practice of Multi-Agent Systems*. Springer, 2012, pp. 31–45.

[14] H. W. H. Mayringer, "False belief understanding in young children: Explanations do not develop before predictions," *International Journal of Behavioral Development*, vol. 22, no. 2, pp. 403–422, 1998.

[15] F. C. Keil, "Explanation and understanding," *Annual review of psychology*, vol. 57, p. 227, 2006.

[16] M. S. Knowles *et al.*, *The modern practice of adult education*. New York Association Press New York, 1970, vol. 41.

[17] S. Lieb and J. Goodlad, "Principles of adult learning," 2005.

[18] L. R. Ye and P. E. Johnson, "The impact of explanation facilities on user acceptance of expert systems advice," *Mis Quarterly*, pp. 157–172, 1995.

[19] B. Y. Lim, A. K. Dey, and D. Avrahami, "Why and why not explanations improve the intelligibility of context-aware intelligent systems," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2009, pp. 2119–2128.

[20] S. R. Haynes, M. A. Cohen, and F. E. Ritter, "Designs for explaining intelligent agents," *International Journal of Human-Computer Studies*, vol. 67, no. 1, pp. 90–110, 2009.

[21] A. Hedhili, W. L. Chaari, and K. Ghédira, "Causal maps for explanation in multi-agent system," in *Intelligent Informatics*. Springer, 2013, pp. 183–191.

[22] S. Gregor and I. Benbasat, "Explanations from intelligent systems: Theoretical foundations and implications for practice," *MIS quarterly*, pp. 497–530, 1999.

[23] G. Milliez, R. Lallement, M. Fiore, and R. Alami, "Using human knowledge awareness to adapt collaborative plan generation, explanation and monitoring," in *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*. IEEE Press, 2016, pp. 43–50.

[24] J. Broekens, D. Degroot, and W. A. Kosters, "Formal models of appraisal: Theory, specification, and computational model," *Cognitive Systems Research*, vol. 9, no. 3, pp. 173–197, 2008.

[25] G. Taylor, K. Knudsen, and L. S. Holt, "Explaining agent behavior," *Ann Arbor*, vol. 1001, p. 48105, 2006.

[26] J. M. Schraagen, S. F. Chipman, and V. L. Shalin, *Cognitive task analysis*. Psychology Press, 2000.

[27] M. A. Neerincx, F. Kaptein, M. A. van Bekkum, H.-U. Krieger, B. Kiefer, R. Peters, J. Broekens, Y. Demiris, and M. Sapelli, "Ontologies for social, cognitive and affective agent-based support of childs diabetes self-management," *Artificial Intelligence for Diabetes*, p. 35.

[28] R. Looije, M. A. Neerincx, and V. d. Lange, "Childrens responses and opinion on three bots that motivate, educate and play," 2008.

[29] N. Epley, C. K. Morewedge, and B. Keysar, "Perspective taking in children and adults: Equivalent egocentrism but differential correction," *Journal of Experimental Social Psychology*, vol. 40, no. 6, pp. 760–768, 2004.