Personalising Behaviour of and Content for Socially Interactive Agents in eHealth Training for Children

Rijgersberg, R.M.

DOI
10.4233/uuid:5edddc16-21ec-4752-8643-b7851565a083

Publication date
2025

Document Version
Final published version

Citation (APA)
Rijgersberg, R. M. (2025). *Personalising Behaviour of and Content for Socially Interactive Agents: in eHealth Training for Children*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:5edddc16-21ec-4752-8643-b7851565a083

Personalising Behaviour of and Content for Socially Interactive Agents

in eHealth Training for Children

Personalising Behaviour of and Content for Socially Interactive Agents

in eHealth Training for Children

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. dr. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op dinsdag 4 februari 2025 om 15:00 uur

door

Rifca Marcella RIJGERSBERG-PETERS

Master of Science in Human Media Interaction
Master of Science in Educatie in de Bètawetenschappen
Universiteit Twente, Enschede, Nederland
geboren te Zevenaar, Nederland.

Dit proefschrift is goedgekeurd door de promotoren.

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. M.A . Neerincx,	Technische Universiteit Delft, promotor
Dr. ir. D.J. Broekens,	Universiteit Leiden, promotor
Onafhankelijke leden:	
Prof. dr. G. Castellano	Uppsala University
Prof. dr. K.V. Hindriks	Vrije Universiteit Amsterdam
Prof. dr. M.M. Specht	Technische Universiteit Delft
Prof. dr. ir. J.C.F. de Winter	Technische Universiteit Delft
Prof. dr. C.M. Jonker	Technische Universiteit Delft, reservelid

Keywords: Intelligent Tutoring System, Social Robot, Education, Ontology Engineering, Diabetes, Child-Robot Interaction

Printed by: Proefschiftmaken.nl

Front & Back: Rifca Rijgersberg-Peters

ISBN 978-94-6510-414-0

An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

Contents

Summary

In this thesis, we focus on developing behaviours for socially interactive agents (SIAs). The context in which the agent is used is a self-regulated learning system for children. We focus on personalising learning objectives and interaction content within an intelligent tutoring system (ITS). We envision a system where children can train diabetes self-management knowledge and skills independent of space and time and in collaboration with the health care professional, legal caretakers, and a SIA. To facilitate long-term interaction with such a system, relevant learning content and appropriate 'intelligent' social behaviour of the SIA are necessary. The envisioned system was developed within the Horizon 2020 PAL-project and evaluated in an iterative design process. The main contributions of the research described in this thesis are: insights into the behaviour design for a NAO robot and its virtual avatar, and the formalisation of learning objectives facilitating personalised learning content.

Most studies on SIA behaviour focus on the design of emotional expressions or implement roles (e.g., peer or tutor) that were not validated for perception. We argue that strategic pedagogical interaction style (i.e., style purposefully selected based on knowledge about the user, task and context such as done by teachers in traditional classroom settings) is necessary but not yet sufficiently studied to design meaningful interactions that surpass the initial novelty and fun. Further, we argue that learning content must be relevant to the child's needs and developmental stage. These two challenges are the subjects of study in the two parts of this thesis.

The main research question addressed in part I is: *How to design SIA behaviours that express different pedagogical styles and what is the effect on learning outcomes?* We answer this question in the four included chapters.

In a systematic review we focus on non-verbal expressions by parameter-based manipulations of bodily shape and motion of humanoid robots and virtual agents, and how these manipulations are perceived by humans. We present a comprehensive review of peer-reviewed published articles and analyse and summarise the available work. Research in this field is multidisciplinary and shows a large variety in concept definitions, behavioural manipulations and evaluation methodologies. We developed the TAXMOD taxonomy as a starting point to develop a shared understanding and interpretation of research objectives and outcomes, and to formulate a road-map. We applied TAXMOD to position and compare research, and to structure the progress made in this area. We found structural support for the fact that some social signals can be displayed by behaviour manipulation in the form of posture- or motion modulation or designed key expressions (fixed behaviours with a specific target expression). Key findings include: 1) the expression of personality traits using virtual or robot bodies is limited to the trait extraversion; 2) the expression of social dimensions such as warmth, competence and dominance is possible, but only when using the whole body, and more research is needed

to disentangle individual effects on warmth, competence and dominance; 3) the expression of emotion is restricted to generic positive versus negative signals; and 4) context seems important for users for the correct interpretation of the expressive behaviour.

In a first perception study we evaluate an educational robot displaying non-verbal behaviours expressing high or low warmth and competence with children at primary schools and a diabetes camp. We show that style expression by a humanoid robot is possible. Bodily posture, hand gestures and paralinguistic cues were manipulated to evoke an expression of a specific level of warmth and competence. The competence dimension in our model was successful. Warmth manipulations were perceived as intended only in combination with high-competence. Moreover, context influenced children's perceptions: at school the robot was perceived warmer and more competent than at camp.

In a second perception study we evaluate an educational robot displaying non-verbal behaviours expressing high or low dominance. We modulate bodily posture and movement, specifically by manipulating body expansiveness. We show the validity of body expansiveness modulation for dominance expression in both postures and gestures and show that with a limited set of parameters we can moderate dominance expression. Specific postures and gestures have a natural tendency towards being perceived as more or less dominant. Further, the manipulation effect is consistent for a variety of behaviours except a sitting pose. This study provides evidence that body expansiveness is an important factor for dominance expression and that this effect is independent of specific behaviours and view angle.

We study the effect of stylised behaviours on children's learning approach and learning gain by having a NAO robot guide children while performing an inquiry-based science learning task where children roll rollers down a slope to discover laws of movement, friction and gravity. Robot style is implemented as variations in verbal strategy and non-verbal style expression, resulting in an expert or facilitator interaction style. No effect of robot interaction style on children's learning approach or gain is reported. Based on only verbal behaviour variations children perceive the explaining robot (either the expert style or explaining verbal strategy with neutral non-verbal behaviour) as more competent than the robot giving evidence descriptions (either the facilitator style or evidence descriptions verbal strategy with neutral non-verbal behaviour). These perception differences did not impact the learning approach or gain in the present study. We did not find perception differences based on the variation of non-verbal behaviour. We did find that the presence of a robot giving feedback on children during rolling trials did cause children to play longer and do more informative experiments compared to no feedback. However, this difference in learning approach did not impact learning gain.

The main research question addressed in part II is: *How to personalise learning content based on personal learning objectives?*

First, we look into how learning goals are formulated in pedagogy and ontologies for education: effective learning goals must attune to the appropriate and desired difficulty level. A way to structure this is Bloom's taxonomy [1]. A learning goal must also have attributes presenting relations between and descriptions of goals. Then, we model educational objectives (i.e., achievements, learning goals and accompanying tasks) in an ontology. The upper ontology structures the classes and relations and defines domain independent constructs (i.e., level and topic). The domain model specifies diabetes self-management training objectives for young children based on current checklists and expert input. The resulting knowledge base was considered relevant to, and covering, the diabetes domain to a considerable extent. From this we conclude that our upper model adequately supports the formalisation of implicit knowledge of health care professionals on diabetes self-management training. A field study with children with type 1 diabetes in the Netherlands and Italy showed that an SIA-ITS offering tasks based on our model to support basic needs for autonomy, competence, and relatedness of children with diabetes [2]. For the formalisation of domain specific learning goals, achievements, tasks and materials in the knowledge structure we recommend the following design guidelines: work in a multidisciplinary team (to define an inventory of important learning goals and define learning activities include domain- and pedagogic experts next to knowledge engineers); formulate achievements from logical learning units (e.g., daily challenges) that require a subset of the knowledge and skills encapsulated in the goals to improve relevance; formulate achievements and goals from the perspective of the child to facilitate ownership and increase experienced relevance; and, define user characteristics relevant to goal and/or task selection. For the integration of the knowledge structure in a multi-modal intelligent tutoring system we recommend the following design guidelines: provide instruction and explanation to the child on how achievements, goals and tasks are selected and can be attained (i.e., that progress on a goal is gained by task completion, and benefits earned by this); embed the objectives in the ITS application to make them easily accessible to the (child) user and integrate them in other system functionality such as feedback on progress provided by a SIA; and offer sufficient learning content such as games and quizzes to maintain interest and engagement.

We developed an authoring tool with a tree-based interface adapted from game design for collaborative personal goal setting and monitoring that implements the ontology of diabetes self-management education, and we co-evaluated this interface with health care professionals. We propose the following design guidelines for an authoring tool: provide clear, visual feedback on goal structure, and active state and progress; consistently use a different representation (e.g., shape) for different concepts of the model (e.g., goal and achievement); cover the full domain and different skill levels with the finite set of goals; and, support assessment of current abilities next to goal setting, progress monitoring and goal attainment registration.

We developed an mHealth dashboard as interface for personal goal and task selection and monitoring, and co-evaluated this interface with children with diabetes. The interface implements our ontology of diabetes self-management education. The following design elements were understandable for all children: colouring indicating status, and navigation between layers of information. Children experienced difficulties interpreting the meaning conveyed in iconic presentations, understanding of the layered information, and navigation. Based on reported usability issues, we present guidelines for the design of a dashboard for children: provide descriptive labels next to visual elements because children lack experience using apps and thus understanding of icons and such; connect elements accordingly by placing them in close proximity and in boxes with appropriate labels; ease navigation between layers for hidden detailed information to avoid cognitive overload; and avoid cluttering elements such as navigation bars.

The work in this thesis shows that robots can express different pedagogical styles perceivable by young children. Dominance expression is mainly dependent on body expansiveness. Warmth and competence expression rely on a complex set of behaviour modulations. However, current style variations are too subtle to impact learning approach and gain. With respect to content personalisation, we show that a structure for and selection of learning objectives provide both a personalised learning path as well as personalised content.

Overall we conclude that to impact learning approach and gain not only SIA behaviour must be modulated, it must be noticed by the learner as well. Learning objectives and content should be formalised within a structure and a user-friendly interface is needed to select objectives and tasks with accompanying content, and monitor progress. The success of an SIA-ITS depends on the amount of available content and social interaction capabilities of the SIA.

Samenvatting

In dit proefschrift richten we ons op het ontwikkelen van gedrag voor 'socially interactive agents' (SIA's). De context waarin de SIA wordt gebruikt is een zelfregulerend leersysteem voor kinderen. We richten ons op het personaliseren van leerdoelen en leerinhoud binnen een 'intelligent tutoring system' (ITS). We stellen ons een systeem voor waarin kinderen kennis en vaardigheden op het gebied van diabeteszelfmanagement kunnen oefenen, onafhankelijk van ruimte en tijd en in samenwerking met de zorgprofessional, wettelijke verzorgers en de SIA. Om langdurige interactie met een dergelijk systeem mogelijk te maken zijn relevante leerinhouden en passend 'intelligent' sociaal gedrag van de SIA noodzakelijk. Het beoogde systeem is ontwikkeld binnen het Horizon 2020 PAL-project en geëvalueerd in een iteratief ontwerpproces. De belangrijkste bijdragen van het onderzoek dat in dit proefschrift wordt beschreven zijn inzichten in het gedragsontwerp van een NAO-robot en zijn virtuele avatar, en de formalisering van leerdoelen die gepersonaliseerde leerinhoud mogelijk maken.

De meerderheid van eerdere onderzoeken naar SIA-gedrag richt zich op het ontwerp van emotionele expressies of gebruiken voorgedefinieerde rollen (bijvoorbeeld peer of tutor) die niet gevalideerd zijn op perceptie. We veronderstellen dat een strategisch gekozen pedagogische interactiestijl (dat wil zeggen, een stijl die doelbewust wordt geselecteerd op basis van kennis over de gebruiker, taak en context, zoals gedaan door leraren in traditionele klaslokalen) noodzakelijk is maar nog niet voldoende onderzocht is om betekenisvolle interacties te ontwerpen die de aanvankelijke nieuwigheid en het plezier overstijgen. Verder veronderstellen we dat de leerinhoud relevant moet zijn voor de behoeften en ontwikkelingsfase van het lerende kind. Deze twee uitdagingen zijn onderwerp van studie in de twee delen van dit proefschrift.

De belangrijkste onderzoeksvraag die in deel I aan bod komt is: *Hoe kunnen we SIA-gedrag ontwerpen dat verschillende pedagogische stijlen tot uitdrukking brengt en wat is het effect op de leerresultaten?* We beantwoorden deze vraag in de vier opgenomen hoofdstukken.

In een systematische review richten we ons op non-verbale expressies door parametergebaseerde manipulaties van de lichaamsvorm en -beweging van humanoïde robots en 'virtual agents'. We inventariseren hoe deze manipulaties door mensen worden waargenomen. We presenteren een uitgebreid overzicht van gepubliceerde artikelen en analyseren het beschikbare werk. Onderzoek op dit gebied is multidisciplinair en vertoont een grote verscheidenheid aan conceptdefinities, gedragsmanipulaties en evaluatiemethodieken. We hebben de TAXMOD-taxonomie ontwikkeld als uitgangspunt om een gedeeld begrip en interpretatie van onderzoeksdoelstellingen en -resultaten te ontwikkelen, en om een routekaart te maken. We hebben TAXMOD toegepast om bestaand onderzoek te positioneren, te vergelijken en om de voortgang op dit gebied inzichtelijk te maken. We tonen dat sommige sociale signalen kunnen worden weergegeven door

gedragsmanipulatie in de vorm van houdings- of bewegingsmodulatie of voorgedefinieerde gedragingen (vaststaand gedrag met een specifieke doeluitdrukking). De belangrijkste bevindingen zijn: 1) de expressie van persoonlijkheidskenmerken met behulp van een virtuele agent of robot is beperkt tot de extraversie; 2) de expressie van sociale dimensies zoals warmte, competentie en dominantie is mogelijk, maar alleen wanneer het hele lichaam wordt gebruikt, en er is meer onderzoek nodig om individuele effecten op de dimensies te onderscheiden; 3) de uiting van emotie beperkt zich tot generieke positieve versus negatieve signalen; en 4) context lijkt belangrijk voor gebruikers voor de gewenste interpretatie van het gedrag.

In een eerste perceptiestudie evalueren we met kinderen op een basisschool en een herfstkamp een educatieve robot die non-verbaal gedrag vertoont en hiermee hoge of lage warmte en competentie uitdrukt. We laten zien dat stijlexpressie door een mensachtige robot mogelijk is. Lichaamshouding, handgebaren en paralinguïstische signalen werden gemanipuleerd om een specifiek niveau van warmte en competentie op te roepen. De competentiedimensie in ons model was succesvol. Warmtemanipulaties werden alleen als bedoeld ervaren als ze samen met hoge competentie getoond werden. Bovendien beïnvloedde de context de perceptie van kinderen: op school werd de robot warmer en competenter ervaren dan op het kamp.

In een tweede perceptiestudie evalueren we een educatieve robot die non-verbaal gedrag vertoont en hiermee hoge of lage dominantie uitdrukt. We moduleren de lichaamshouding en -beweging, met name door manipulatie van expansie. We laten zien dat de modulatie van lichaamsexpansie succesvol is voor dominantie-expressie in zowel houdingen als gebaren en we tonen aan dat een beperkte set parameters afdoende is om dominantie-expressie te kunnen modereren. Specifieke houdingen en gebaren hebben een natuurlijke neiging om als min of meer dominant te worden ervaren. Het manipulatie-effect is consistent voor een verscheidenheid aan gedragingen, met uitzondering van een zittende houding. Deze studie suggereert dat lichaamsexpansie een belangrijke factor is voor de expressie van dominantie en dat dit effect onafhankelijk is van specifiek gedrag en kijkhoek.

We bestuderen het effect van robotgedrag op de leeraanpak en leeropbrengst van kinderen door een NAO-robot kinderen te laten begeleiden terwijl ze een onderzoekende, wetenschappelijke leertaak uitvoeren. De taak houdt in dat kinderen rollers van een helling af laten rollen om wetten van beweging, wrijving en zwaartekracht te ontdekken. Robotstijl is geïmplementeerd als variaties in verbale strategie en non-verbale expressie, resulterend in een expert- of facilitator-interactiestijl. Er is geen effect gezien van de robotinteractiestijl op de leeraanpak of -opbrengst van kinderen. Op basis van alleen verbale gedragsvariaties ervaren kinderen een verklarende robot (expert-stijl of verklarende verbale strategie met neutraal non-verbaal gedrag) als competenter dan de robot die observaties deelt (facilitator-stijl of beschrijvende verbale strategie met neutraal non-verbaal gedrag). Deze perceptieverschillen hadden in het huidige onderzoek geen invloed op de leeraanpak of -opbrengst. We hebben geen perceptieverschillen gevonden op basis van de variatie in non-verbaal gedrag. We zagen wel dat de aanwezigheid van een robot die feedback geeft tijdens rolproeven ervoor zorgt dat kinderen langer spelen en meer informatieve experimenten doen in vergelijking met afwezigheid van feedback gegeven door de robot. Dit verschil in leeraanpak had geen invloed op de leeropbrengst.

De belangrijkste onderzoeksvraag die in deel II aan bod komt is: *Hoe kunnen we leerinhoud personaliseren op basis van gepersonaliseerde leerdoelen?*

Eerst bestuderen we hoe leerdoelen worden geformuleerd in de onderwijskunde en in ontologieën voor het onderwijs. Effectieve leerdoelen moeten in elk geval aansluiten op de juiste en gewenste moeilijkheidsgraad. Een manier om dit te structureren is de taxonomie van Bloom [1]. Een leerdoel moet ook attributen bevatten die relaties tussen en beschrijvingen van de doelen weergeeft. Vervolgens modelleren we educatieve doelstellingen (d.w.z. prestaties, leerdoelen en bijbehorende taken) in een ontologie. De 'upper' ontologie structureert de klassen en relaties en definieert domeinonafhankelijke constructies (bijvoorbeeld niveau en onderwerp). Het domeinmodel specificeert trainingsdoelstellingen voor diabeteszelfmanagement voor jonge kinderen op basis van actuele checklists en input van deskundigen. De resulterende kennisbank werd als relevant beschouwd en omvat het diabetesdomein in aanzienlijke mate. Hieruit concluderen we dat ons domeinonafhankelijke model de formalisering van impliciete kennis van gezondheidszorgprofessionals over diabeteszelfmanagementtraining adequaat ondersteunt. Een veldstudie met kinderen met diabetes type 1 in Nederland en Italië toonde aan dat een leersysteem op basis van ons model de basisbehoeften aan autonomie, competentie en verbondenheid van kinderen met diabetes ondersteunt [2]. Voor het formaliseren van domeinspecifieke leerdoelen, prestaties, en taken presenteren we de volgende ontwerprichtlijnen: werk in een multidisciplinair team (voor het inventariseren van belangrijke leerdoelen en het definiëren van leeractiviteiten worden domein- en pedagogische experts ingeschakeld naast technische uitwerking); formuleer om relevantie te verhogen prestaties op basis van logische leereenheden (bijvoorbeeld dagelijkse uitdagingen) die een subset vereisen van de kennis en vaardigheden die in de doelen zijn vastgelegd; formuleer prestaties en doelen vanuit het perspectief van het kind om eigenaarschap en de ervaren relevantie te vergroten; en neem gebruikerseigenschappen mee die relevant zijn voor doel- en/of taakselectie. Voor de integratie van de kennisstructuur in een multimodale ITS presenteren we de volgende ontwerprichtlijnen: geef instructie en uitleg aan het kind over hoe prestaties, doelen en taken worden geselecteerd en kunnen worden bereikt; bed doelstellingen in de ITS-applicatie in om deze eenvoudig toegankelijk te maken voor de (kind)gebruiker en deze te integreren in andere systeemfunctionaliteiten zoals feedback over de voortgang; en bied voldoende leerinhoud (games, quizzen) aan om de interesse en betrokkenheid te behouden.

We hebben een ondersteunende applicatie voor beheer van leerdoelen ontwikkeld met een 'tree-graph'-gebaseerde interface. Het ontwerp is gebaseerd op game-ontwerp en aangepast op het gezamenlijk stellen van persoonlijke doelen en monitoring van deze doelen. De interface maakt gebruik van de ontologie van diabeteszelfmanagementeducatie. We hebben deze interface samen met professionals in de gezondheidszorg geëvalueerd. We presenteren de volgende ontwerprichtlijnen voor: geef duidelijke, visuele feedback over de doelstructuur en de actieve status en voortgang; gebruik consequent een andere representatie (bijvoorbeeld vorm) voor verschillende concepten van het model (bijvoorbeeld doel en prestatie); bestrijk het volledige domein en verschillende vaardigheidsniveaus met een eindige reeks doelen; en biedt ondersteuning bij de beoordeling van de huidige vaardigheden (naast het stellen van doelen, het monitoren van de voortgang en het registreren van het bereiken van doelen).

We hebben een in-app-dashboard voor kinderen ontwikkeld als interface voor persoonlijke doel- en taakselectie en monitoring. We evalueerden deze interface samen met kinderen met diabetes. De interface maakt gebruik van de ontologie van diabeteszelfmanagementeducatie. De volgende ontwerpelementen waren voor alle kinderen begrijpelijk: kleuren die de status aangeven en navigatie tussen informatielagen. Kinderen ondervonden moeilijkheden bij het interpreteren van de betekenis van iconen, bij het begrijpen van de gelaagde informatie en vertoonden navigatieproblemen. Op basis van gerapporteerde bruikbaarheidsproblemen presenteren we richtlijnen voor het ontwerp van een dashboard voor kinderen: zorg voor beschrijvende labels naast visuele elementen; verbind bij elkaar behorende elementen door ze dicht bij elkaar en in kaders met labels te plaatsen; en vereenvoudig de navigatie tussen lagen bij het verbergen van gedetailleerde informatie om rommeligheid en cognitieve overbelasting te voorkomen, en voorkom drukke elementen zoals navigatiebalken.

Het werk in dit proefschrift laat zien dat robots verschillende pedagogische stijlen kunnen uitdrukken die jonge kinderen kunnen waarnemen. Dominantie-expressie is voornamelijk afhankelijk van de lichaamsexpansie. Warmte- en competentie-expressie zijn afhankelijk van een complexe reeks gedragsmodulaties. De huidige stijlvariaties zijn echter te subtiel binnen een heel leersysteem om de leeraanpak en -opbrengst te beïnvloeden. Wat betreft de personalisatie van leerinhoud laten we zien dat een structuur voor en selectie van leerdoelen zowel een gepersonaliseerd leertraject als gepersonaliseerde inhoud oplevert.

Over het geheel genomen concluderen we dat om de leeraanpak te beïnvloeden en leeropbrengst te behalen, niet alleen het SIA-gedrag moet worden gemoduleerd, maar dat dit ook door de leerling moet worden opgemerkt. Leerdoelen en inhoud moeten binnen een structuur worden geformaliseerd en er is een gebruiksvriendelijke interface nodig om doelen en taken met bijbehorende inhoud te selecteren en de voortgang te monitoren. Het succes van een SIA-ITS hangt af van de hoeveelheid beschikbare inhoud en sociale interactiemogelijkheden van de SIA.

REFERENCES

[1] D. R. Krathwohl, *A revision of bloom's taxonomy: An overview,* Theory into practice **41**, 212 (2002).

[2] M. Neerincx, W. Vught, O. Blanson Henkemans, E. Oleari, J. Broekens, R. Peters, F. Kaptein, Y. Demiris, B. Kiefer, D. Fumagalli, *et al.*, *Socio-cognitive engineering of a robotic partner for child's diabetes self-management,* Frontiers in Robotics and AI **6**, 118 (2019).

1

INTRODUCTION

Social robots are entering the educational domain where their aim is to support children in their learning process. The goal is to facilitate collaborative learning (either in the role of peer or educator) and long-term learning. This requires the robot to display behaviour supportive to the learning context and to offer appropriate learning materials. In recent years, the field of Child-Robot Interaction (CRI) has seen significant advancements, particularly in understanding the various roles a robot can take (i.e., tutor, peer or tool [1]) and how to express robot emotion, mood and personality. Further, advancements of technology have led to the emergence of Intelligent Tutoring Systems (ITSs) as powerful tools for enhancing learning experiences. ITSs offer opportunities to advance educational processes and guidance by direct and personalised instructions and feedback. Specifically, the integration of robots and virtual agents into these ITSs has demonstrated potential in addressing the educational needs of children in traditional subjects such as STEM and language learning (e.g. [2, 3]) as well as self-regulated extracurricular learning such as health training (e.g. [4]). While previous studies have touched upon CRI and ITSs separately, an area that remains relatively under-explored is the domain of long-term child-robot interaction in comprehensive educational environments (i.e., robot or agent-included ITSs).

This dissertation delves into a distinctive aspect of this gap, developing robots that complement the roles of professionals and parents in supporting children's self-regulated learning in daily life. The primary scope of this study is to investigate and facilitate collaborative, long-term learning interactions between children and robots. To facilitate these long-term educational interactions, they must remain interesting and engaging. Personalisation of both robot interaction and presented learning content is key. Thus, the robot must be knowledgeable about the user's state (of knowledge) and context and have the ability to respond appropriately. In this dissertation we research two challenges important for a robot integrated in an ITS: I) the design of the behaviour style of a robot as collaborative learning partner (either tutor or peer); and II) the development of a model and method to deliver personalised content.

In recent literature people refer to both robots and virtual agents (also mentioned as embodied conversational agents) as Socially Interactive Agents (SIAs). ”SIAs, whether virtually or physically embodied, are autonomous agents that are able to perceive an environment including people or other agents, reason, and decide how to interact, and express attitudes such as emotions, engagement, or empathy. They are capable of interacting with people and each other in a socially intelligent manner using multimodal communicative behaviors with the goal to support humans in various domains.”[5] The use of SIAs in education and specifically e-health training introduces a novel and engaging approach that can capture children's attention and foster their active participation in the learning process. By leveraging the capabilities of SIAs it becomes possible to create a more immersive and personalised learning environment. However, the challenge is to overcome the so-called *novelty effect*: humans initially exhibit high motivation and interest in the new interactive technology, but this interest can wear off quickly after repeated interactions [6]. To achieve long-term and effective interactions with untrained humans, it is crucial to offer an interesting experience that is tailored to the child's needs and the task at hand, and for the SIA to have some *social intelligence*. Based on observable behaviours such as bodily motion (i.e., gestures) and shape (i.e., postures) people form impressions of the others and develop expectations [7]. For SIAs, to engage in meaningful interactions with humans, the importance of social intelligence is widely acknowledged (e.g., [8–10]). The SIA should be able to express social behaviour appropriately perceived by the human, meaning that the human understands the communicative intent of the agent. Progress has been made in emotion modelling (e.g., [11–13]). However, research into behaviour style is less thorough: no comprehensive, validated model exists yet of non-verbal behaviours to express style in human-robot interactions. In inter-human educational interactions the importance of strategic behaviour and expression thereof (i.e., teaching style) is acknowledged. Theoretical models exist, as well as studies of the effects. In human-robot interaction robot role (the *what* e.g., [14, 15]) has been studied. However, these studies rarely provide a validated behaviour model for style expression (the *how*). An exception is Johal [16] who studied parenting style.

In Part I of this dissertation we address the first challenge: how to design the interaction style of a SIA as a collaborative learning partner (either tutor or peer) that is appropriate given the learning context. First, we investigate whether style is perceivable and if there is an effect on learning by studying existing work on validated expressive behaviour. Then we model pedagogical style behaviour and validate the expression thereof. Finally, we study the effect of these validated behaviours on learning outcomes.

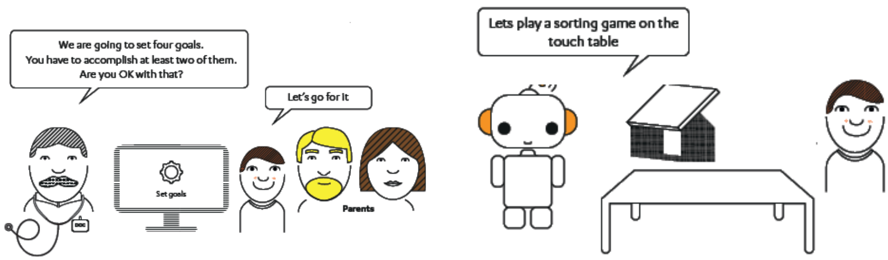
Another challenge in developing a SIA-ITS for e-health training lies in the need for personalisation of educational content, tailoring of the learning path, progress monitoring, and adaptive feedback. ITSs provide tailored instructions or feedback to a learner to facilitate effective learning while lessening the students dependency on a teacher. Success rates of ITSs depend on the user's (intrinsic) motivation for system usage and adherence to the educational program [17]. Motivation (for learning) was suggested to benefit from goal setting and feedback on goal attainment (e.g., [18, 19]), as well as tailoring of the educational path to the learner's prior knowledge. Self-regulated learning systems can exploit this by focusing on goal-oriented learning. A personalised guidance approach in ITSs attempts to provide learners with the most appropriate learning content accounting for individual learner characteristics such as current knowledge, interests and motivation [20]. Incorporating personal learning goals in educational technologies requires: knowledge of learning goals relevant to the domain; a mechanism to set personal learning goals, and share this information with the system and between different types of users (e.g., doctor and patient, teacher and student); and means to monitor learning progress.

In Part II of this dissertation we address the second challenge: delivering relevant, personalised content by proposing the development of a domain independent ontology to formalise both learning content and related objectives (goals and tasks). We further provide a domain specific instantiation based on domain knowledge of diabetes self-management training for children aged 7-14. This ontology serves as a knowledge base that captures the relationships among different learning goals, tasks and content, allowing for adaptive and personalised instruction based on individual needs. Further, interfaces to communicate the content and objectives to various user groups have been designed and evaluated with the users.

1.1. VISION

The research presented in this dissertation is centred around the development of a personal assistant for a healthy lifestyle (PAL) [21], encompassing long-term child-robot interactions. A (NAO) robot and its virtual avatar are intended to support children with type 1 diabetes mellitus (T1DM) in extracurricular self-management education.

Imagine a child, Robin, aged 8 who has been diagnosed with T1DM since he was 6 years old. Robin sits together with his health care professional (HCP) and mentions he soon has an important soccer game. Robin definitely wants to be on the field, but is afraid he can't because he can't maintain a good blood sugar level during sports. The HCP then shows a program where the HCP and Robin in collaboration can set learning objectives for Robin to work on at home (Figure 1.1a). There are even objectives specifically tailored to sports and diabetes. By setting this (set of) learning goals Robin gains access to educational materials such as games and videos. In the hospital Robin can play one of the games with a robot Charlie (Figure 1.1b).

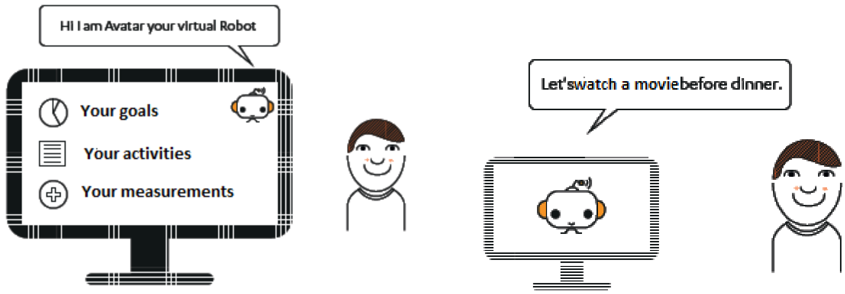


(a) setting objectives

(b) playing a game with the robot

Figure 1.1: PAL usage in the hospital.

The HCP hands Robin a tablet to take home. On the tablet is an avatar of Robot Charlie and it gives Robin access to the educational materials as well as insight in the activated goals and progress on these goals (Figure 1.2a). From playing with Charlie in the hospital Robin already learned foods high on carbohydrates, he can eat this food to quickly heighten his blood sugar level during soccer! At home Robin starts the app. Charlie suggests watching a video that explains the effect of sports on his blood sugar level (figure 1.2b). Robin does so. Charlie provides feedback complimenting Robin on his motivation and he receives an update of another attained goal! Robin regains trust that he can actually play the soccer game.



(a) options in the home app

(b) Charlie proposing an activity

Figure 1.2: PAL usage at home.

We envision 'Charlie' to adapt its behaviour to learner preferences and to be sensitive to context. Based on the user model and contextual factors an interaction style with a high likelihood of appropriateness is selected and expressed, the active choice is continuously evaluated and if necessary adjusted.

In this thesis, we define non-verbal behavioural styles based on the social dimensions *Friendliness* (i.e., warmth and affiliation), *Competence*, and *Dominance* allowing us to construct behaviours with non-verbal styles such as *Direct*, *Friendly* or *Neutral*. We adhere two verbal communication strategies *Evidence Descriptions* and *Explanations* as presented in [22].

Pedagogical interaction style should be attuned to the co-learner role of the robot (e.g., *Expert*, *Facilitator* [23] or *Peer*) and is defined by the combined selected verbal and non-verbal style and expressed in robot Charlie's integrated behaviour. The user model may consist of knowledge of user preferences, objectives and dynamic perception of the user state (e.g. affective state, interaction intensity and motivation). Moreover, contextual factors such as the current location, users and activity can be taken into account. For each interaction style output behaviour is defined; this includes task selection and expression of style by, for example, bodily shape, motion and dialog acts.

Imagine again Robin who wants to play his soccer game. Taking no risks of not being allowed to join, Robin decides to write down a fake measurement of his blood sugar level. Charlie, Robin's personal support SIA, notices that Robin changed the value. Charlie knows that Robin has a learning goal to manage diabetes during sports, what should Charlie do? Having the knowledge that Robin lacks insight in consequences of bad blood values during sports, Charlie should support Robin in his education on this and provide adequate tasks and information. *Explanation* is the a-priori preferred verbal strategy thus the *Expert* interaction style is selected and Charlie proposes to play an educational quiz. However, Charlie notices that the interaction intensity is very low, and believes that Robin is sad. Knowing that Robin prefers empathic interaction conveyed in the peer style when feeling sad, Charlie changes to the affective talk strategy and starts a conversation about sports asking Robin about his likes and dislikes regarding his illness and soccer. Not only the verbal behaviour is changed, also the non-verbal behaviours are adapted to be congruent with the verbal strategy. For example, the peer style (i.e., affective talk strategy) is accompanied with *Friendly* (i.e., high friendliness, high competence, low dominance) behaviours.

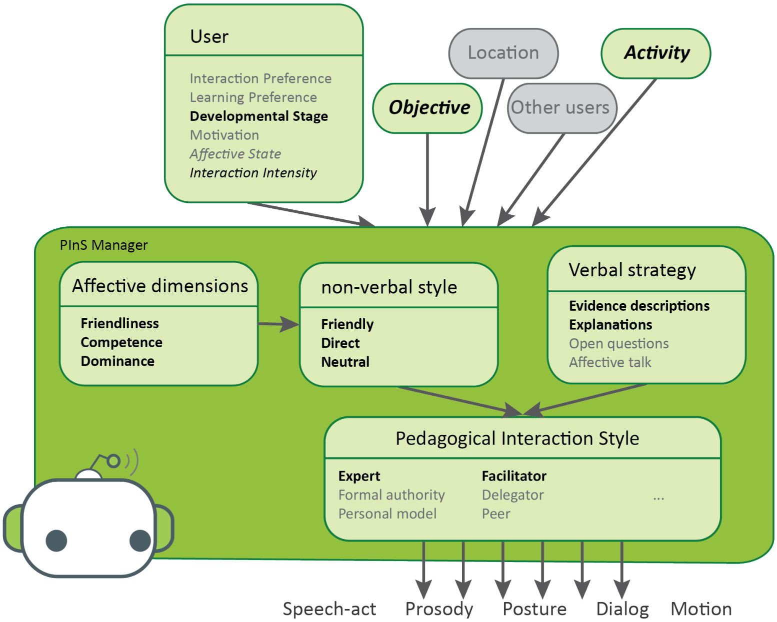


Figure 1.3: Model for management of pedagogical interaction style depicting factors (user, goal, location, other users and activity) influencing selection of verbal strategy and non-verbal style that together determine which pedagogical interaction style is chosen, which is then expressed in robot behaviour using, for example, motion, speech and posture.

1.2. Research Questions and Thesis Structure

We introduce the main problem of current research in ITSs that use SIAs as a part of their system. Agent behaviour is neither designed to align with educational needs nor validated for human perception. Further, there is a need to personalise content in SIA-ITSs. This thesis provides an effort to contribute to this. The main research questions of this thesis are:

Research Question 1 *How to design SIA behaviours that express different pedagogical styles and what is the effect on learning outcomes?*

Research Question 2 *How to personalise learning content based on personal learning objectives?*

The answers to these questions contribute to the advancement of SIAs and ITSs. By designing a validated pedagogical interaction style aimed to optimise learning outcomes, and by developing an ontology to effectively define and assign learning objectives to personalise learning content, this work aims to contribute to an adaptive learning environment that supports the need of younger learners in the healthcare domain within the PAL project [24].

The two main research questions are investigated in the two parts of this thesis. Part I focuses on designing behaviours for expression of pedagogical interaction style in child-robot interaction. Part II has a focus on personalisation of learning content via learning objectives. Each part is introduced with a background section and the research questions addressed in the included chapters.

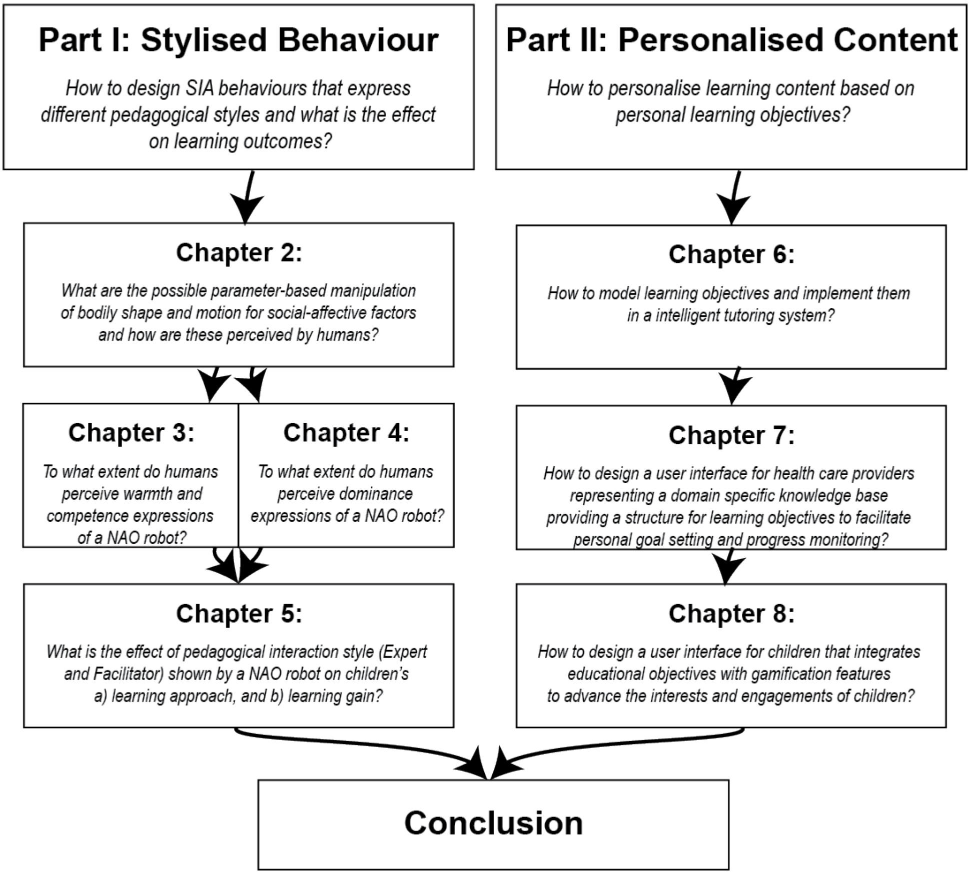


Figure 1.4: Visual outline of the thesis chapters and research questions addressed therein.

1

REFERENCES

[1] O. Mubin, C. J. Stevens, S. Shahid, A. Al Mahmud, and J.-J. Dong, *A review of the applicability of robots in education,* Journal of Technology in Education and Learning **1**, 13 (2013).

[2] S. Y. Okita, V. Ng-Thow-Hing, and R. Sarvadevabhatla, *Learning together: ASIMO developing an interactive learning partnership with children,* in *RO-MAN* (IEEE, 2009) pp. 1125–1130.

[3] M. Saerbeck, T. Schut, C. Bartneck, and M. D. Janse, *Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor,* in *CHI* (ACM, 2010) pp. 1613–1622.

[4] O. a. Blanson Henkemans, V. Hoondert, F. Schrama-Groot, R. Looije, L. L. Alpay, and M. a. Neerincx, *"I just have diabetes": children's need for diabetes self-management support and how a social robot can accommodate their needs,* Patient Intelligence **4**, 51 (2012).

[5] B. Lugrin, C. Pelachaud, and D. Traum, *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application* (ACM, 2022).

[6] C. Breazeal, K. Dautenhahn, and T. Kanda, *Social robotics,* Springer handbook of robotics , 1935 (2016).

[7] G. B. Moskowitz, *Social cognition: Understanding self and others* (Guilford Press, New York, 2005).

[8] C. Breazeal, *Role of expressive behaviour for robots that learn from people,* Philosophical Transactions of the Royal Society of London B: Biological Sciences **364**, 3527 (2009).

[9] A. Vinciarelli, M. Pantic, D. Heylen, C. Pelachaud, I. Poggi, F. D'Errico, and M. Schroeder, *Bridging the Gap Between Social Animal and Unsocial Machine: A Survey of Social Signal Processing,* Affective Computing **3**, 69 (2011).

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

[11] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx, *Effects of bodily mood expression of a robotic teacher on students,* in *Proc. Int. Conf. on Intell. Robots and Syst.* (IEEE, Chicago, Illinois, 2014) pp. 2614–2620.

[12] M. L'hommet and S. Marsella, *Expressing Emotion Through Posture and Gesture,* in *The Oxford Handbook of affective computing* (Oxford University Press, 2015) pp. 273–285.

1

[13] G. Ball and J. Breese, *Emotion and Personality in a Conversational Agent,* Embodied Conversational Agents , 189 (2001).

[14] R. Looije, M. a. Neerincx, and V. D. Lange, *Children's responses and opinion on three bots that motivate , educate and play,* Physical Agents **2**, 13 (2008).

[15] C. Zaga, M. Lohse, K. P. Truong, and V. Evers, *The Effect of a Robot's Social Character on Children's Task Engagement: Peer Versus Tutor,* in *ICSR*, Vol. 9388 LNAI (Springer, Paris, 2015) pp. 704–713.

[16] W. Johal, S. Pesty, and G. Calvary, *Towards companion robots behaving with style,* in *Proc. 23rd Int. Symp. on Robot and Human Interactive Commun.* (IEEE, Edinburgh, Scotland, 2014) pp. 1063–1068.

[17] D. H. Schunk and B. J. Zimmerman, *Motivation and Self-Regulated Learning: Theory, Research, and Applications* (Routledge, 2008).

[18] J. F. Bryan and E. A. Locke, *Goal setting as a means of increasing motivation.* Journal of applied psychology **51**, 274 (1967).

[19] P. Black and D. Wiliam, *Assessment*, Vol. 5 (2006) pp. 37–41, arXiv:0402594v3 [arXiv:cond-mat] .

[20] Brusilovsky, *Domain modeling for personalized guidance,* in *Design Recommendations for Intelligent Tutoring Systems: Volume 4 Domain Modeling,* edited by R. A. Sottilare, A. C. Graesser, X. Hu, A. Olney, B. Nye, and A. M. Sinatra (US Army Research Laboratory, Orlando, Florida, 2016) Chap. 15, pp. 165–2184.

[21] M. Neerincx, W. Vught, O. Blanson Henkemans, E. Oleari, J. Broekens, R. Peters, F. Kaptein, Y. Demiris, B. Kiefer, D. Fumagalli, *et al.*, *Socio-cognitive engineering of a robotic partner for child's diabetes self-management,* Frontiers in Robotics and AI **6**, 118 (2019).

[22] T. J. Van Schijndel and M. E. Raijmakers, *Parent explanation and preschoolers' exploratory behavior and learning in a shadow exhibition,* Science Education **100**, 153 (2016).

[23] A. F. Grasha, *A matter of style: The teacher as expert, formal authority, personal model, facilitator, and delegator,* College teaching **42**, 142 (1994).

[24] F. Kaptein, B. Kiefer, A. Cully, O. Celiktutan, B. Bierman, R. Rijgersberg-Peters, J. Broekens, W. Van Vught, M. Van Bekkum, Y. Demiris, *et al.*, *A cloud-based robot system for long-term interaction: Principles, implementation, lessons learned,* ACM Transactions on Human-Robot Interaction (THRI) **11**, 1 (2021).

Part I

Stylised Behaviour

In this part we present four chapters that investigate the effect of SIA bodily motion and shape on human perception of the agent. The ultimate goal is to express different pedagogical interaction styles to facilitate long-term educational interactions. The behaviour design patterns discussed all relate to the expression of social-affective factors. Within our use case, i.e. the PAL project, we focus on the role of a SIA as an educator. Therefore we ground our work in theories from human-human educational interactions and social signals which will be discussed below.

Background

Teaching style

Teaching styles are behaviour patterns that affect information presentation and interaction. Style is considered a technique addressing differences in learning- and cognitive styles [3]. Grasha [4] proposed five teaching styles: Expert (transmitting information), Formal authority (providing feedback and establishing boundaries), Personal model (showing an example), Facilitator (encouraging critical thinking), and Delegator (available in the background during project work). An experienced human educator selects the style strategically based on three factors: 1) students' capability; 2) teacher's willingness to give up control / desire for control (resistance to change); and 3) teacher's willingness to build and maintain relationships.

These factors are related to student competence and teacher dominance and warmth / affiliation. Competence, dominance and warmth are three fundamental interpersonal social dimensions.

Social Dimensions

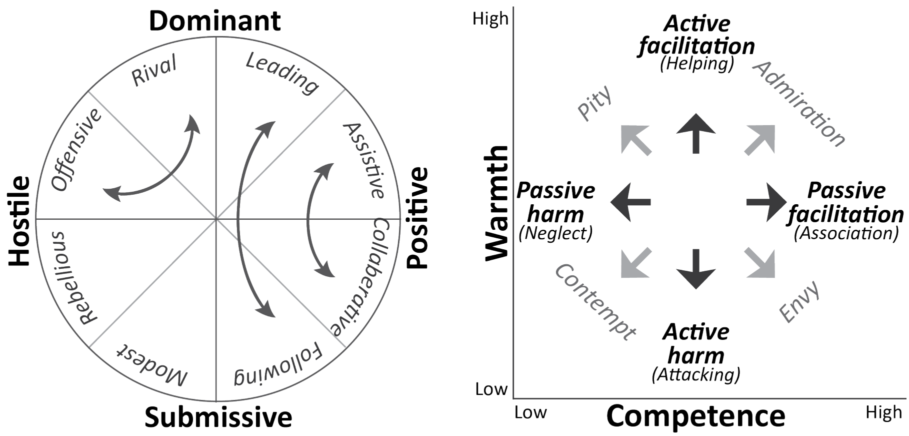
Social dimensions are universal, interpersonal concepts that can only exist for a person with a reference to another person. The social dimensions originate from the interpersonal circumplex (Leary's Rose) and the Stereotype Content Model (SCM) (see below). Although "researchers tend to agree on two primary dimensions that underlie relational communication —affiliation (or intimacy) and dominance (status/competence)" [5], the two models named the axes differently. The different terms of Leary's Rose and SCM for warmth and affiliation entail similar key concepts: judging respectively good-ill, and affection-hostility (i.e., 'horizontality' [6]).

Interpersonal Circumplex This model—also known as Leary's Rose—defines interaction stance by two dimensions: dominance and affiliation [7, 8] (Figure I.5a). The horizontal affective-axis depicts willingness to cooperate. The vertical dominance-axis depicts the degree of power. Commonly, the circumplex is partitioned into eight octants, resulting in eight types of patterns of interpersonal relationships: steering, friendly, understanding, accommodating, uncertain, dissatisfied, reprimanding, enforcing [9]. Leary's theory states two interaction rules: dominance is complementary and affiliation symmetric, meaning that, an opposed stance evokes opposed stance by the other interlocutor while a dominant stance evokes submissiveness by the other. A cooperative style has been linked to maintaining contact, a competitive style to aversion. For example, head

Parts of this chapter have been published in [1, 2].

nods [10], gaze, and open posture [11] are associated with high affiliation. A dominant style is commonly stereotyped as loud and obtrusive, while a submissive style is believed to show discrete, unnoticed behaviour. A meta-analysis of the relation between dominance and non-verbal behaviours argued that previous work has been inconclusive or based on limited data, and concluded that the relation exists, but to different degrees and directions, depending on the person and situation [12]. Nonetheless, the model is widely used in social skill training and teacher training.

Stereotype Content Model The Stereotype Content Model defines warmth and competence as two fundamental dimensions of social perception [13]. Warmth (kindness, empathy, friendliness and trustworthiness) evaluates the valence of any intention, competence (intelligence, power, efficacy and skill) assesses the ability to act on these intentions. Perceived warmth and competence generate emotions of admiration, envy, pity, and disgust towards someone, and predict active / passive and facilitative / harmful behaviour patterns (Figure I.5b). Perceived warmth is believed to be evoked by sincere smiles, head tilt, nodding, leaning forward, and open gestures [14, 15]. Coldness is expressed by closed hands, cutting motion, chin down, and the body pivoting away [15]. Upright posture and open gestures are predictors of perceived power, and associated with competence [14, 15]. Fiddling was suggested to signal low control and confidence, therefore resulting in low-competence judgement [15]. Additionally, warmth judgements are believed to be made before, and influence, competence evaluations —persons evaluated as warm, are likely to be judged more competent [14]. The dimensions of warmth and competence are related to the teaching styles presented above.



(a) Interpersonal Circumplex

(b) Stereotype Content Model

Figure I.5: Models of social interaction evaluation and prediction.

In summary, style can be defined as a behaviour patterns signalling our attitudes towards a person or situation, affecting how others evaluate us and subsequently respond. Four social constructs are important to the notion of style: *warmth, competence, dominance*, and *affiliation*. However, affiliation, warmth and friendliness can all be defined as horizontality [6]. In the remainder of this thesis we use the following three behavioural style factors: *warmth/affiliation* as *friendliness*, *competence*, and *dominance*.

Research Questions

The main research question for part I is: *How to design SIA behaviours that express different pedagogical styles and what is the effect on learning outcomes?*

We are interested in designing behaviours expressing pedagogical interaction style focusing on manipulation of the bodily shape and motion of a NAO robot. Pedagogical style is an interaction style of the SIA tuned to educational activities and composed of the social dimensions presented above: warmth/affiliation as *friendliness*, *competence*, and *dominance*. We derived four research questions, addressed in the chapters of part I.

Research question 1; Chapter 2
What are the possible parameter-based manipulations of bodily shape and motion for social-affective factors and how are these perceived by humans?

In this chapter we review existing research in parameter-based manipulation of bodily shape and motion and the influence on the human perception of social and affective factors. We include studies that present: 1) a detailed description of displayed non-verbal behaviour and applied (parameter-based) manipulations in bodily shape and/or motion displayed by an artificial agent (either virtual or robotic); and, 2) the result of a user study evaluating how humans perceive the agent displaying these behaviours. We structure both the possible outcomes of perceived social-affective factors and the possible parameter-based manipulations in TAXMOD, a TAXonomy for the MODulation of non-verbal behaviour and perception thereof by humans.

Submitted as: **R. Rijgersberg-Peters, J. Broekens, M.A. Neerincx**, *Parameter-based Modulation of Bodily Shape and Motion in Socially Interactive Agents: A Taxonomy and Review*, Submitted to ACM Transactions on Human-Robot Interaction.

Research question 2; Chapter 3
To what extent do humans perceive warmth and competence expressions of a NAO robot?

Warmth and competence are two fundamental dimensions of social perception [13]. These dimensions contribute to teaching style [4]. Based on existing work, in particular the work of [16], we created a model of non-verbal behaviour to express high/low warmth and competence. We expect children to evaluate robots as expressing varying levels of warmth and competence based on the displayed behaviours. In a perception study, we evaluate this model applied to a NAO robot giving a lecture at primary schools and a diabetes camp in the Netherlands.

Published as: **R. Peters, J. Broekens, and M.A. Neerincx**, *Robots Educate in Style: The Effect of Context and Non-verbal Behaviour on Children's Perceptions of Warmth and Competence*, 26th IEEE international symposium on robot and human interactive communication, (2017).

Research question 3; Chapter 4
To what extent do humans perceive dominance expressions of a NAO robot?

Dominance (or vertically) is a dimension in the interpersonal circumplex —or Leary's Rose [7], and an important factor in social interaction [17]. Leary's Rose has been used in teacher education to train strategical behaviour in student interaction. We developed a parameter-based model for head tilt and body expansiveness. This model was applied to a variety of gestures and postures. These behaviours are evaluated by human observers in two different studies with respectively static pictures of key postures and real-time gestures. We expect participants to evaluate robots as expressing more dominance when bodily shape or motion was expanded.

Published as: **R. Peters, J. Broekens, K. Li, M.A. Neerincx**, *Robots expressing dominance: Effects of behaviours and modulation*, 8th International conference on affective computing and intelligent interaction (2019).

Research question 4; Chapter 5
What is the effect of pedagogical interaction style (Expert and Facilitator) shown by a NAO robot on children's a) learning approach, and b) learning gain?

We study the effect of robot stylised behaviour on children's learning approach and learning gain by having a robot guide children while performing an inquiry-based learning task. Robot style is implemented as variations in verbal strategy and non-verbal style expression. We hypothesise that children demonstrate greater active involvement, exploratory behaviour, and learning when interacting with a Facilitator robot (giving Evidence Descriptions and showing 'Friendly' motions) compared to an Expert robot (giving Explanations and showing 'Direct' motions) because the Facilitator style is suitable for self-discovery learning [4]. Also we study the individual effect of verbal or non-verbal behaviour. We expect children to show a higher level of exploratory behaviour when guided by a coach using Evidence Descriptions [18], and non-verbal behaviour congruent to verbal strategy to yield better performances compared to nonstylised neutral motions.

Submitted as: **R. Rijgersberg-Peters, T.J.P. van Schijndel, J. Broekens, and M.A. Neerincx**, *Effect of Robot Interaction Style on Children's Learning Approach and Cognitive Outcomes*, Submitted to International Journal of Social Robotics.

I

REFERENCES

[1] R. Peters, J. Broekens, and M. A. Neerincx, *Robots educate in style: The effect of context and non-verbal behaviour on children's perceptions of warmth and competence,* in *Proc. 26th Int. Symp. on Robot and Human Interactive Commun.* (IEEE, Lisbon, Portugal, 2017) pp. 449–455.

[2] R. Rijgersberg-Peters, J. Broekens, and M. A. Neerincx, *Parameter-based modulation of bodily shape and motion in robots andvirtual agents: A taxonomy and review,* (submitted).

[3] R. M. Felder, L. K. Silverman, *et al.*, *Learning and teaching styles in engineering education,* Engineering education **78**, 674 (1988).

[4] A. F. Grasha, *A matter of style: The teacher as expert, formal authority, personal model, facilitator, and delegator,* College teaching **42**, 142 (1994).

[5] C. H. Dupree and S. T. Fiske, *Universal dimensions of social signals: Warmth and competence,* in *Social Signal Processing* (Cambridge University Press, 2017) pp. 23–33.

[6] M. S. Mast and J. A. Hall, *The vertical dimension of social signaling,* in *Social Signal Processing* (Cambridge University Press, 2017) pp. 34–45.

[7] T. Leary, *Interpersonal diagnosis of personality,* Amer. J. of Physical Med. & Rehabilitation **37**, 331 (1958).

[8] J. S. Wiggins, *An informal history of the interpersonal circumplex tradition,* J. of personality assessment **66**, 217 (1996).

[9] T. Wubbels, M. Brekelmans, P. den Brok, J. Levy, T. Mainhard, and J. van Tartwijk, *Let's make things better: Developments in research on interpersonal relationships in education,* in *Interpersonal Relationships in Education*, edited by T. Wubbels, P. den Brok, J. van Tartwijk, and J. Levy (Sense Publishers, Rotterdam, 2012) Chap. 14, pp. 225–249.

[10] R. Gifford, *The Role of Nonverbal Communication in Interpersonal Relations,* in *Handbook of interpersonal psychology: Theory, research, assessment, and therapeutic interventions* (Wiley, NY, 2011) Chap. 11, pp. 171–190.

[11] D. S. Berry and J. S. Hansen, *Personality, Nonverbal Behaviour, and Interaction Quality in Female Dyads,* PSPB **26**, 278 (2000).

[12] J. a. Hall, E. J. Coats, and L. S. LeBeau, *Nonverbal behavior and the vertical dimension of social relations: a meta-analysis,* Psychological bulletin **131**, 898 (2005).

[13] S. T. Fiske, A. J. Cuddy, and P. Glick, *Universal dimensions of social cognition: Warmth and competence,* Trends in cogn. sci. **11**, 77 (2007).

[14] L. L. Carli, S. LaFleur, and C. Loeber, *Nonverbal behavior, gender, and influence,* Personality and Social Psychology **68**, 1030 (1995).

I

[15] A. J. Cuddy, M. Kohut, and J. Neffinger, *Connect, Then Lead,* Harvard Business Review **91**, 54 (2013).

[16] T.-H. D. Nguyen, E. Carstensdottir, N. Ngo, M. S. El-Nasr, M. Gray, D. Isaacowitz, and D. Desteno, *Modeling warmth and competence in virtual characters,* in *Proc. Int. Conf. on Intell. Virtual Agents* (Springer, Delft, The Netherlands, 2015) pp. 167–180.

[17] M. Schmid Mast and J. A. Hall, *Social signal processing,* (Cambridge University Press, Cambridge, 2017) Chap. The vertical dimension of social signaling, pp. 34–45.

[18] T. J. Van Schijndel, R. K. Franse, and M. E. Raijmakers, *The exploratory behavior scale: Assessing young visitors' hands-on behavior in science museums,* Science Education **94**, 794 (2010).

2

Parameter-based Modulation of Bodily Shape and Motion in Socially Interactive Agents: A Taxonomy and Literature Review

Socially Interactive Agents express social behaviours that need to be appropriately perceived by humans. Appropriate means that the perception reflects the underlying communicative intention accurately. This review focuses on non-verbal expressions by parameter-based manipulations of bodily shape and motion, and how these are perceived by humans. A systematic search on Google Scholar for articles published until January 2023 returned 63 studies that report parameter-based manipulations in non-verbal behaviour and social perception thereof by humans. To structure these studies we propose TAXOMOD, a TAXOnomy for the MODulation of non-verbal behaviour and perception thereof by humans. With this taxonomy, we characterise and relate all of the included studies. We found strong indications that some social expressions can be reliably displayed by behaviour manipulation in the form of posture- or motion modulation. However, for most manipulation-perception pairs the number of studies is limited (and variation within studies too large) to find considerable support for correlations. Expansiveness manipulations seem the most effective, specifically for expression of extraversion and positiveness (i.e., valence and happy/sad). We discuss recommendations and directions for future research.

This paper has been submitted to ACM Transactions on Human-Robot Interaction

2.1. Introduction

As robots and virtual agents emerge in the public and private domain, they engage in social interactions and need to develop *social intelligence*. For example, whether as a tool, co-learner or mentor [1], these 'Socially Interactive Agents' (SIAs) [2] should have the ability to express appropriate affective social signals by display of non-verbal behavioural cues [3]. The importance of social intelligence is stressed by many. For example, Breazeal et al. [4] state that "for agents to become part of human life, they should be able to communicate with people in similar ways people interact with others". Vinciarelli et al. [5] consider generation of social behaviour a "crucial need for agents in human-like interaction". And sociability, the ability to perceive, understand and express social cues of communication in a human understandable way, is believed to be "crucial to facilitate social bonding and long-term engagement" [6]. Embedding social behaviour in SIAs has been proven to influence user experience and outcome. For example, a social robot for children with diabetes was perceived as more empathetic and a better educator, and increased health awareness and adherence to exercises and diary keeping compared to a non-social robot [7].

2.1.1. Behavioural Expressions and Their Perceptions

State-of-the-art systems use generic *role or strategy models* to select agent behaviour (e.g., [8–10]), or adapt their behaviour based on a limited *user model* such as the perceived level of knowledge (e.g., [11–14]) or the user's affective state (e.g., [15–17]). These variations in agent behaviour influence user experience (e.g., [18–22]), user behaviour (e.g., [23–25]), and interaction outcome (e.g., [26–28]). This indicates that people are able to detect and respond differently to variations in agent behaviour, but less attention seems to be given to how variations in role or strategy are displayed through behavioural cues, nor whether the designed role or strategy was perceivable.

People form impressions of and make inferences about other people's thoughts, feelings, and beliefs. Observable non-verbal behaviour and appearance inform these *perceptions*. Cuddy et al. [29] state that warmth and competence are universal dimensions, believed to be accountable for 82% of variance in our perceptions of others [30]. Another factor explaining the behaviour of someone else is personality attribution [31]. Thus, communication is not limited to content, but also informs about the intent and social attitude towards someone or a situation. This is expressed in non-verbal behaviour, where people interpret motion cues as signals for the current state in humans, but also in animals and inanimate devices [32]. Perception studies have been conducted studying the effect of verbal behaviour style (e.g, [18, 33]), or appearance such as robot anthropomorphism [34], and the rendering style of virtual agents [35]. Moreover, cues specific to artificial agents have been studied, for example, lighting appeared to influence user perception of a virtual agent's level of dominance [36].

2.1.2. Social-Affective Expressions and Their Perceptions

The field of social signal processing (SSP) attempts to improve social intelligence of artificial agents by understanding and modelling social interactions [5, 37]. A social signal is an (deliberate or unintended) expression of social attitude towards someone or a situation (e.g., hostility or agreement), displayed by behavioural cues (e.g., gestures and facial expressions), and affecting the interaction. Previous work includes analysis of human behaviour (e.g., [38, 39]), automated behaviour recognition (e.g., [40, 41]), and behaviour synthesis (e.g., [42–44]). Expressive behaviours are employed in agents interacting with users ranging from an emphatic chess-playing robot [22] to a moody robotic tutor [45]. Underlying behaviour models are often based on human behaviour, but ,needless to say, are a simplification of actual human behaviour. And, although humans do treat inanimate agents as social beings [46], interpretation of cues and therefore perception of the agent are not necessarily the same.

2.1.3. RELATED RESEARCH APPROACHES AND OUTCOMES

Previous examples in sections 1.2 and 1.3 show a great interest in social human-agent interaction and agents' expressive behaviours. Most research seems to focus on emotion (e.g., [45, 47–49]) or personality (e.g., [50–52]). In general, a wide variety of behavioural cues have been studied using various methods (including exposure and outcome measures), applying them to many different agents. The diversity in research of non-verbal behaviour in HRI was recently stressed in a review [53], who discusses both perception and effect results over four main nonverbal communication modalities: kinesics (distinguishing between arm gestures, body and head movements, eye gaze, and facial expressions), proxemics, haptics, and chronemics, as well as multi-modal combinations of these modalities. Another survey on social robotics reviewed social judgements such as deception and politeness, and studies on preference, liking and engagement [54]. To the best of our knowledge, no coherent state-of-the-art overview currently exists on research into the expression and perception of non-verbal social cues of socially interactive agents (SIAs).

This paper presents such a review, starting with two notions. *First*, SIA behaviours should bring about socially appropriate perceptions by the humans. Appropriate means that the perception reflects the underlying intention and dynamic social context accurately. Social expressions are integrated into agents' fluent behaviours and can be viewed as "modulations" of these behaviours. Therefore, this review focuses on humanoid agent's and robot's non-verbal expressions, providing a state-of-the-art overview of research on parameter-based manipulations of bodily shape and motion, and how they are perceived by humans. The *second* notion is that the research in this domain is conducted and described in rather divergent ways. To make progress as a community and provide a state-of-the-art review, we need a shared understanding of the main concepts in this domain. Therefore, we will first provide definitions and a taxonomy of the main concepts in the next section.

2.2. DEFINITIONS AND TAXONOMY

Research on non-verbal human-robot interaction is multidisciplinary with a diverse vocabulary and focus. To relate the different research approaches and outcomes, and to enable shared projections of the research field (e.g., visions, challenges and roadmaps), this section provides definitions and a novel taxonomy, *TAXOMOD*, of the core concepts of parameter-based behaviour modulation. To provide definitions we rely on foundations from psychology and other social sciences. We will first discuss the various non-verbal behaviours and possible manipulations thereof, followed by the relation to social signals and perceptions of behavioural cues.

2.2.1. Non-verbal Behaviour Concept Foundations

Non-verbal behaviours are a non-linguistic means of communication, central and essential in the communication process. Non-verbal cues are both universal and cultural and contribute to the meaning derived from communication, thereby enhancing (mis)understanding of the verbal communication, but non-verbal cues (with or without verbal communication) also function to form impressions of others, display identity, display relational state, and attempt to influence others [55]. Historically, non-verbal behaviour functions were viewed to clarify and amplify verbal communication by redundancy, substitution, complementation, emphasis, and contradiction [56]. Later, the affective function was acknowledged in the following list of functions: 1) emblems that carry a linguistic meaning; 2) illustrators that reinforce a verbal message; 3) affect displays; 4) regulators managing turn-taking; 5) and manipulators for tension release [57].

Non-verbal communication is a complex dynamic process; multiple cues in parallel may be needed to serve a single function, and might serve another function at the same time, further, previous interactions may influence subsequent ones [55]. There are different channels of non-verbal communication which we discuss below.

Channels of non-verbal behaviour

Kinesics informally referred to as body language, is the study of physical bodily movements as communicative means [58]. This behaviour entails bodily movements, positioning, facial expressions, and (arm- and/or head) gestures [59], which can communicate extensive contextual, social, and interpersonal information (e.g., situation awareness, social intent, emotional state) [60].

Proxemics is about the perception and use of space related to communication, namely, the conscious or unconscious setting of distances between various objects, others and oneself [61]. Social distances —or ‘personal space’— were categorised into four zones: public (<12 ft); social (4–12 ft); personal (1.5–4 ft); and intimate (0–1.5 ft) [62], contributing to comfort levels, affiliation, and intimacy [63].

Haptics deals with communicative signals through touch (i.e., tactile interactions), which is seen as the “earliest and most elemental mode of communication” [64].

Vocalics / Paralinguistics is about communication through voice in non-linguistic form. Prosody, pitch, volume, and intonation modify meaning, or display emotion. Paralinguistics plays an important role in communication since speech requires the presence of a voice that can be modulated and all modulations of voice as such are paralinguistic.

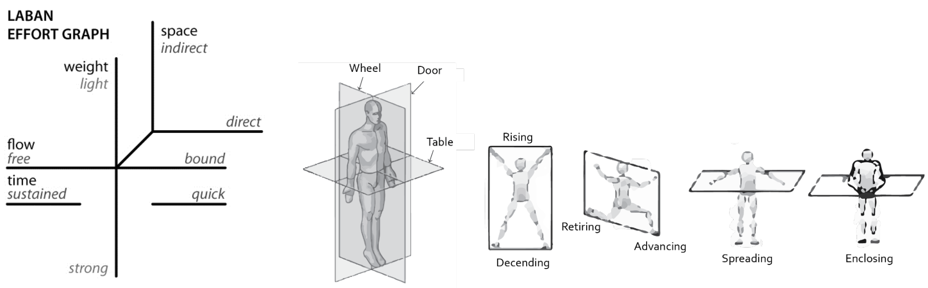
Chronemics concerns the role of time in communication. It looks at the tempo of human interaction and the pace at which we expect communication to occur [65].

CATEGORIES OF KINESICS

In early research, kinemes were proposed to analyse bodily movements, analogous to phonemes and morphemes in linguistics [66]. Later, research in sign language showed that any hand-sign can be described by the handshape, location, orientation, and movement [67]. Which ultimately led to the belief that all bodily movements can be described by the parameters *amplitude*, *fluidity*, *power*, *acceleration*, and *repetition* [19]. Alternatively, rhythm, tempo, sequence, and direction have been proposed as formal descriptions of movement [68]. These classifications bear some overlap with Laban Movement Analysis (LMA), describing performance of bodily movements by four qualities: *body*, *effort*, *shape*, and *space* [69]. The body quality describes how people move, where movement is initiated and how it spreads through the body. The space quality describes the motion in terms of form (e.g., ball-like) and change in form (e.g., shivering). The shape quality describes changes in the form of the body on three planes (dimensions): table (horizontal), door (vertical), and wheel (sagittal) (see Figure 2.1b). The effort quality describes the dynamics of movement on a continuum (between two extremes) of four factors: space (direct-indirect), weight (light-strong), time (sustained-quick), and flow (free-bound) (see Figure 2.1a). LMA was initially designed for dance but (mainly the effort and shape qualities) is nowadays widely applied —also in human-computer interaction (e.g., [70–72]).

Head Movement involves movement restricted to the head such as gaze behaviour. Head movements are associated with specific communication functions such as the semantic meanings of nodding and shaking for indicating referents in conversations [73]. Arguably nodding and shaking are (head only) gestures, but also the result of respectively vertical or horizontal head movement.

Posture is the body orientation and the (static) position of the limbs, the initial posture modifies bodily movement. Posture reveals the structure, content, and interrelationships of human interactions [74]. Expansiveness/openness is seen as an important informative signal. Whereas in a closed posture sensitive body parts (i.e., throat, abdomen and genitals) are mostly obscured, in an open posture these body parts are exposed and feet, hands and fingers are spread, and the head is straight.



(a) Effort motion factors (b) Shape dimensions

Figure 2.1: Laban Movement Analysis qualities Effort and Shape.

Facial Expression involves movement of the facial muscles to form coherent expressions and is believed to be the most important cue for emotion expression [75].

Gesture involves movements of the hands and/or head that are communicative (i.e., intentionally and meaningful to assist speech, such as pointing) or informative (i.e., signals about the state of the speaker, such as scratching indicating discomfort) [76]. The high level of detail and dexterity of gestures support various communicative gestures: deictic, or indexic (e.g., pointing); iconic, or lexical (universal representative of objects and actions); metaphoric, or symbolic also called emblems (representative of abstract concepts with a conventional meaning e.g., waving); and beat, or motor (short, repetitive, rhythmic movement aligned with prosody for punctuating other modes) [75, 77].

2.2.2. Non-verbal Behaviour Manipulation: Review Scope & Taxonomy

For practical reasons[1] we limit our review to kinesics, excluding facial expressions because its limited applicability to humanoid robots such as NAO). In other words, we look at bodily motion (including gesture) and shape (i.e., posture).

We acknowledge the body language categories, but do not differentiate between them to allow for a broader comparison (by stacking results, regardless of the categories, we have a larger set of studies). Instead, we differentiate between *type* of manipulation and *body part* to which the manipulation was applied[2]. Below we present the resulting taxonomy of behaviour manipulation (See Figure 2.2).

Type

describes the main characteristic of the change in shape or movement of the body. We have defined four types: *expansiveness*, *approach*, *dynamics*, and *repetition*. Expansiveness and approach manipulations can be applied to either static postures or movements. Motion dynamics and number of repetitions manipulate bodily movement.

Expansiveness manipulations adjust the bodily shape and/or motion to increase/decrease the total volume in terms of, for example, body openness and gesture scale. Expansiveness manipulations change bodily shape on the vertical (door plane) and/or horizontal (table plane) axes. An exception is head tilt which involves changes of the sagittal (wheel plane) axis as well as the vertical (door plane) direction. Nonetheless head tilt has been classified as expansiveness due to the manipulation resulting in a more open, enlarged bodily shape.

[1]The proposed limitation follows from our research interest —modelling behavioural cues, expressing social signals, for a humanoid robot or virtual agent— posing limitations on possible behaviour manipulation and requiring differentiation between anatomic movements and expressions.

[2]Studies specify manipulation on different levels of detail ranging from no specification to three degrees of freedom of individual joint or individual body parts such as knee or finger, and even differentiate between position and movement. We decided to differentiate on a high level of detail (i.e., full-body, head, torso, arms, and legs) because we can translate more detailed description and not the other way around, and it will provide some insight into differences in effects for individual body parts.

Approach / avoidance manipulations change bodily shape and/or motion in the sagittal (wheel plane) direction, advancing or retiring to/from the user —resulting in approaching or avoiding the user— such as a chest lean. Once more the head is an exception; head orientation changes (e.g., pan) and gestures, such as gaze, change shape mainly in the horizontal direction (door plane), but have been classified as approach because this type of change influences the attention directed at the user.

Dynamics manipulations change the LMA effort qualities of bodily movement. The effort qualities (i.e., time, weight, space, and flow) are based on LMA and are formed by a combination of changes in temporal-extent and shape. Examples of time manipulations include gesture speed and motion acceleration. Examples of weight manipulations include adjustments to body stability and motion power. Examples of space manipulations include motion tension and body relaxation. Examples of flow manipulations include motion fluidity and gesture fluency.

Repetition manipulates the number of movements and is also referred to as rate or frequency. This type of manipulation was applied to either all movements, a certain type of behaviours such as all gestures, or to specific behaviours such as gaze or self-contact.

BODY PART

describes the body part to which a manipulation was applied. We differentiate between full-body and individual body parts arms, head and torso. We acknowledge the existence of other body parts such as —obviously— legs, but limit our taxonomy to those body

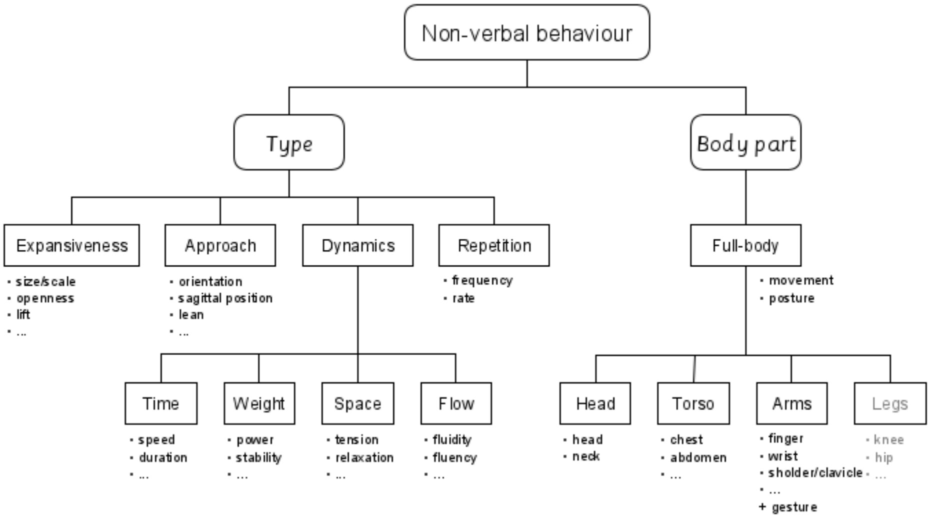


Figure 2.2: Taxonomy of the behaviour manipulation presenting the possible types and modalities including examples.

parts presented in the review results. Manipulations on all modalities can be applied either to static postures or motions.

Full-body manipulations include those in which the combined set of manipulations applied in parallel to more than the arms and head or more than one other body part results in one holistic manipulation. This may include the application of one manipulation to multiple body parts (e.g., expansiveness of head, arms and legs in [78]), or parallel application of various manipulations to individual body parts (e.g., a diverse set of manipulations each applied to either full-body, arms and/or head in [79]).

Head manipulations adjust the neck bend and thereby head position. For example, head tilt and pan, but also head gestures such as gaze. On some occasions head movements are part of arms- and head gestures.

Torso manipulations adjust the bodily bend such as chest lean.

Arms includes manipulations of the limbs from shoulder to hand. For example, arm position, gesture scale, clavicle lift, wrist extension or finger rigidness. Arm motions are often gestures and in some cases combined with head movement.

Legs manipulations adjusting the legs, from hip to toe such as knee bend. None of the reviewed studies evaluated leg manipulations on an individual basis (only as part of full-body manipulations).

2.2.3. Social Perception Concepts

Interpersonal communication goes beyond merely sending and receiving messages between individuals. The messages conveys communicative meaning and relational information, and is delivered through verbal and non-verbal behaviours. Also body language (e.g., gestures) has a communicative, literal meaning (e.g., illustrators, emblems) and/or an informative function of the agent's internal state (e.g., affect, state). We are mainly interested in the latter, specifically *someone else's* (social) perception about these social signals. Humans make inferences, attributions and perceptions about many different (affective) constructs appearing on different levels with personality traits being the most stable, style being context dependent, and emotions as short and intense feelings.

In the field of social signal processing (SSP) a social signal is defined as "a communicative or informative signal that, either directly or indirectly, provides information about social facts, namely, interactions, social emotions, social attitudes, or social relations" [5, 37, 80]. Social signals are expressed by display of non-verbal behavioural cues such as those described in Section 2.2.1 which are produced during interactions and provide information about or influence the actor, relationship and/or interaction [37]. Research demonstrates that robots and virtual agents can be programmed to express social signals, despite limitations in bodily movement (e.g., [81, 82]). Figure 2.3 illustrates some examples of non-verbal cues in different channels and categories that express a social signal in human-robot interaction.

Personality

Personality is a dynamic and organised set of traits and behaviour patterns that define how someone is perceived by others. A distinction is made between personality style, traits and stance, where style and traits are relatively stable over time, opposed to stance that describes the current interpersonal behaviour, traits and motives [83]. Personality style entails individuals' inclinations and preferences across context.

2

Personality traits are what we mostly refer to when we speak about personality. Traits are consistent, habitual patterns of behaving, thinking and feeling [84]. A commonly used model is the *five factor model* (FFM), also know as the Big Five or OCEAN personality traits (i.e., openness to experience, conscientiousness, extroversion, agreeableness, and neuroticism) [85]. Extraversion and conscientiousness are perceived most accurately [86] because these traits are more visual [87, 88].

Openness to experiences describes a person's general interest in and appreciation for a variety of experiences. People high on openness are by nature curious and willing to try new things. Others may perceive a person as creative and unpredictable.

Conscientiousness describes a person's impulse control. Conscientious people have a tendency to display self-discipline and strive for achievement, and a preferences for planned actions. Others may perceive a person as stubborn and focused.

Extraversion describes a person's tendency for external activation. Extraverts tend to engage in interactions and stand out in groups. Others may perceive a person as full of energy, enthusiastic or dominant.

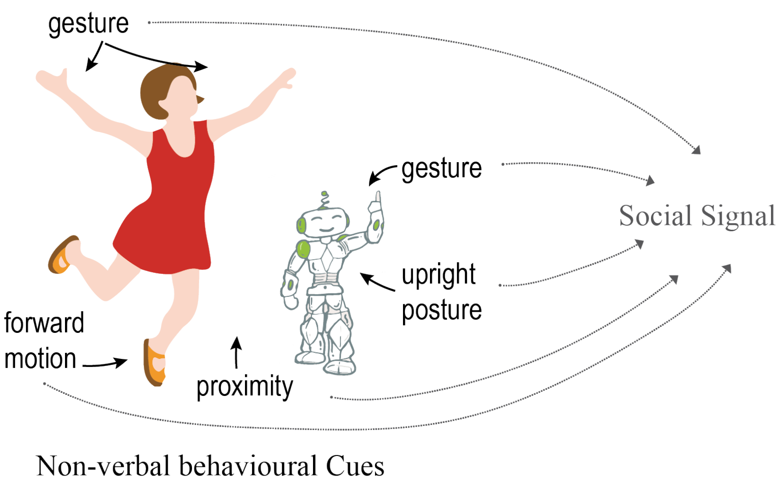


Figure 2.3: Non-verbal behavioural cues are displayed, and express a social signal.

Agreeableness describes a person's concern for social harmony. Agreeable people value getting along and are optimistic about other people. Others may perceive a person as kind, trustworthy and willing.

Neuroticism describes a person's response to negative emotions —it is the inverse of emotional stability. Neurotic people are vulnerable to stress and show strong emotions. Others may perceive a person as pessimistic and anxious.

Personality types are an alternative model that characterises people by their preferences and attitudes. Myers-Briggs model for personality [89] proposed two perceiving functions (sensation and intuition) and two judging functions (thinking and feeling) that are expressed in either an introvert or extravert way [90].

Social Dimensions

Social dimensions are universal, interpersonal concepts that can only exist for a person with a reference to another person. The social dimensions originate from the interpersonal circumplex (Leary's rose) and the stereotype content model (SCM) (see below). Although "researchers tend to agree on two primary dimensions that underlie relational communication —affiliation (or intimacy) and dominance (or status/competence)" [91], the two models named the axes differently. The different terms of Leary's rose and SCM for warmth and affiliation entail similar key concepts: judging respectively good-ill, and affection-hostility (or love-hate). Further, Mast et al. [92] defined *horizontality* as affiliatedness, warmth and friendliness, and *verticality* as power and control —just as dominance is.

Warmth & Competence are seen as two universal dimensions of social signals/cognition [29, 91, 93]. Perception of warmth (kindness, empty, friendliness, trustworthiness) is the result of evaluation of the valence of (social) intent, judgement of competence (intelligence, ability, efficacy), and the assessment of someone's ability to act upon their intentions. The *Stereotype Content Model* (SCM, see Figure 2.4b) warmth and competence framework allows for social judgements into four quadrants which predict active vs passive and facilitative vs harmful behaviour patterns [29].

Dominance & Affiliation , or interpersonal stance, is the position of a person towards the interaction or the relationship. The *interpersonal circumplex* —or Leary's Rose— (see Figure 2.4a is a two dimensional model for interpersonal stance defined by dominance (degree of power vs submission) and affiliation (willingness to cooperate) [94, 95] divided in octants of high/low dominance and affiliation. The horizontal affective-axis is symmetrical and vertical dominance-axis is complementary. According to interpersonal theory, these two dimensions are primary to describe human relations [94].

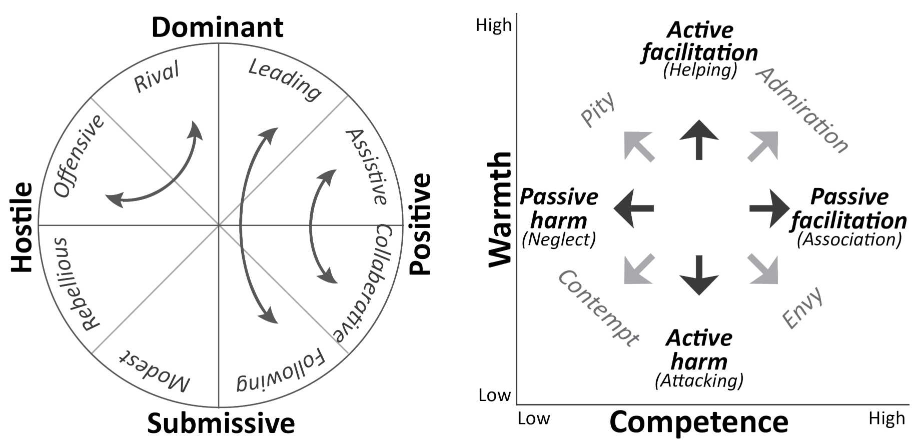
Emotions

Emotions are high-intensity, short-lived, and directed responses to stimuli. Emotions are described either as discrete (basic) categories or in a multidimensional model. However, it is possible to map discrete emotions to a multidimensional model [24].

Discrete emotions are often identified or classified in studies on universal mental state expressions (e.g., in the face), such as Ekman's six well-known basic emotions: happiness, anger, fear, surprise, sadness, and disgust [96]. However, more "hidden" discrete emotions can also be felt, such as experiencing feelings for someone or conveying information about a social relationship [97]. Ekman [98] proposed an expanded listing of these social emotions (e.g., hate, envy, contempt, admiration, pride, shame, embarrassment). The following emotions appear most often in Socially Interactive Agent (SIA) research:

- *Happy*, including joy, describes a person who is in a positive stage, showing a feeling of pleasure or contentment.
- *Sad* describes a person in a negative, unpleasant state.
- *Anger* is a negative, strong feeling of annoyance, displeasure, or hostility.
- *Fear* an unpleasant emotion caused by the threat of danger, pain, or harm.
- *Surprise* is a strong feeling occurring in response to something unexpected, this can be either pleasant or unpleasant.

Multi-dimensional emotions are an alternative way to describe emotions, typically in a circumplex. A complex "wheel of emotions" was developed by Plutchik [99]. On this wheel are placed eight primary bipolar emotions (i.e., joy-sorrow, anger-fear, acceptance-disgust, and surprise-expectancy) that may blend into complex emotions, such as contempt. SIA studies often use a simpler Valence, Arousal, Dominance (VAD) model; Russell and Mehrabian [100] showed that many emotion words and affective situations can be represented in this multidimensional model, by classifying each emotion by three independent bipolar dimensions of pleasure-displeasure (i.e., valence), high-low arousal, and dominance-submissiveness.



(a) Interpersonal Circumplex (b) Stereotype Content Model

Figure 2.4: Models of social interaction evaluation and prediction.

- *Valence*, or pleasure, describes a person's attitude towards another person or situation in terms of positive-negative. For example, joy is a positive emotion while anger negative.

- *Arousal* describes the physical activity and bodily activation associated with the feeling in terms of high-low. For example, boredom is a low arousal state while rage is a high arousal emotion.

- *Dominance* describes whether or not someone feels in control of the situation. Anger is seen as a dominant emotion, fear as submissive.

2.2.4. Social Perception Taxonomy

In SIA studies, a variety of behaviour manipulations have been applied to express a certain social signal. How users perceive such behavioural cues has been measured using various scales and instruments. The review and taxonomy focus on measures of personality, social dimensions, style, and emotion (see Figure 2.5). Up to now, style proved to be used scarcely as an outcome measure in the concerning research, and will therefore not appear in the review. Note that the taxonomy and review focus on the perception of robot and agent expressions and, consequently, user opinion (such as likeability) is out of scope.

We adhere to the concept definitions as presented previously and, therefore, we only briefly highlight the main classes. Additionally, we added the 'other' class clustering relevant measure for concepts that have been studied scarcely up to now.

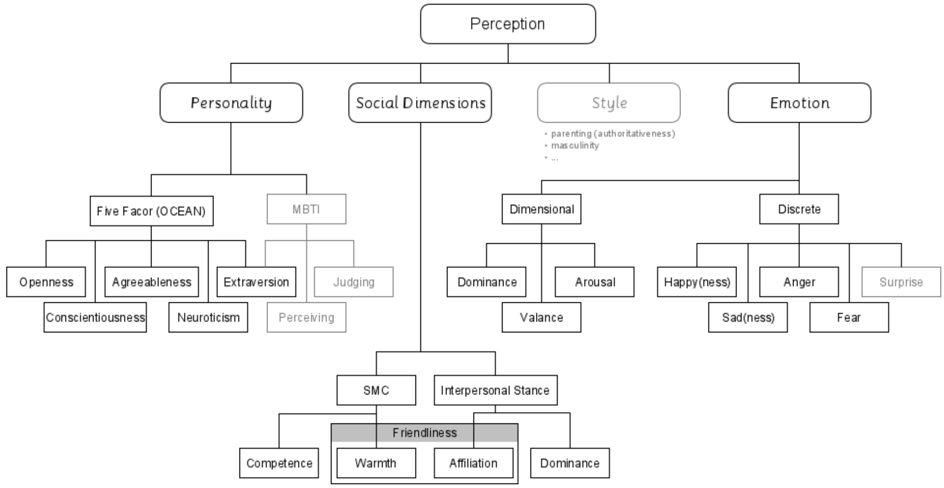


Figure 2.5: Taxonomy of the perception measures taken. Items depicted in grey will not appear in the review results, because they have so far hardly been investigated.

Personality measurements included up-to five traits from the five factor model: *openness, conscientiousness, extraversion, agreeableness*, and *neuroticism*. Alternatively, MBTI personality type measures have been reported, but these were reported only scarcely and therefore placed under 'other' perceptions.

Social dimensions measures were differentiated into three factors:

- *Friendliness* describes a person's intentions towards another in terms of both warmth and affiliation, also referred to as horizontality.
- *Competence* describes a person's capabilities to act upon one's intentions and includes also measures of intelligence.
- *Dominance* describes a person's power and control over another person or situation, also referred to as verticality.

Emotion measures include factors from the multidimensional VAD-model and discrete (basic) emotions that were subject of a few of the studies included in this review. Scarcely reported emotions such as disgust or pride have been listed under 'other'. We strive to stay as close as possible to the factors as presented in the reviewed studies; therefore, we included both discrete and multidimensional emotions resulting in the following list of emotion factors: *valence, arousal* and *dominance*, and *happy, sad, anger* and *fear*.

Other measures include all social constructs not classifiable under personality, social dimensions and emotion as presented above. These measures may include more high level constructs such as robot role or parenting style, but also once or scarcely reported measures such as disgust or masculinity.

2.3. REVIEW METHOD

This review presents an up to date overview of bodily shape and motion cues and manipulations thereof designed to communicate an agent's personality, social dimensions or emotion, and validated perception thereof by humans. It focuses on the effects of variations in bodily shape and motion of a SIA and how the user perceives this agent (e.g., a user's perception of the SIAs level of extraversion due to modulation of head tilt). We exclude recognisability of expressions displayed by undocumented, pre-designed behaviours (i.e., motions that have been created in advanced and do not involve parameter-based real-time modulation). For example, a particular bodily animation of a virtual agent being recognised as anger. Studies evaluating well-documented pre-designed behaviours and presenting the correlations between manipulated parameters and perceptions have been included (e.g., [101]). In the remainder of this section we present the search strategy, selection process and criteria, and the characteristics of studies meeting these criteria.

2.3.1. Review Procedure

We used Google Scholar between May 2nd — May 17th 2017 and January 8th — January 29th 2023 to perform a series of search queries each attempting to retrieve the effect of variations in bodily shape and motions displayed by a SIA on users' perceptions of the agent. We ran 11 search queries to ensure inclusion of work from various fields, all using specific vocabulary to describe relevant constructs, concepts and factors (see Table 2.1). As even our most narrow search query returned over 300 matches, we limited the results of each query to the first 100 publications, amounting to a total of 2200 publications used as raw data for this review. Publication selection from this initial set followed a four-step procedure. This procedure was executed in mid-2017 and again early 2023. In the first step, publications were filtered based on their title and abstract. Publications that appeared to meet the inclusion criteria were kept. This resulted in 175 publications is 207 and 115 in 2023. In the second step, we filtered out 73 and 71 duplicates (due to the 11 search queries), two publications published in a language other than English, and one not available to us, resulting in 101 and 44 publications. In the third step, the remaining articles were fully read. Those articles meeting the inclusion criteria were kept as the final data set to be reviewed, resulting in 36 and 22 unique publications. Some publications present multiple studies, while some studies are described across multiple papers; in total we included 38 and 25 unique studies.

2.3.2. In- and Exclusion Criteria

For inclusion in this review, publications are required to be scientific peer-reviewed contributions published in English. Selected publications must present: 1) a detailed description of displayed non-verbal behaviour and applied (parameter-based) manipulations in bodily shape and/or motion (e.g., head tilt or motion speed) as displayed by a SIA); 2) the result of a user study evaluating how humans perceive the agent displaying these behaviours. Articles presenting only agent features that cannot be manipulated in real-time, such as appearance or gender are excluded. Likewise, articles limited to manipulations of facial features, proximity, or gesture synchronisation are excluded. Moreover, articles presenting a model of non-verbal behaviour without validation of user perception (i.e., theoretical or computational models and effect studies) are excluded. All publications discussing a human's (affective, social or interpersonal) perception of a SIA—either with physical or virtual embodiment, independent of application domain—are included in this review.

From each included publication we listed individual studies separately. For each study, we extracted applied manipulations and outcome measures and classified them according to our taxonomy (see Section 2.2).

2.3.3. Study Characteristics

A total of 58 publications, presenting 63 studies evaluating the effect of variation in bodily shape or motion as displayed by a SIA on humans' perceptions of this agent personality, social dimensions, emotion or other social construct, have been included in this review. Of these studies, 32 report a robot perception study (15 NAO robot) and 31 a virtual agent perception study (mainly human avatars (20), followed by wooden mannequins (7)). The reported results were obtained from 4 real-world interactions, 16 lab interactions, 10 physical lab observations, and 33 passive screen observations. An overview of the included studies and their characteristics[3] is presented in Table 2.2.

[3] We only report characteristics within the scope of the review, co-reported independent variables that are not bodily shape or motion (e.g., facial expression, paralinguistic cues and verbal behaviour), and dependent variables that are not affective or interpersonal perception factors (e.g., naturalness, likeability, or social presence) are not included.

Table 2.1
THE SEARCH QUERIES WITH FOR EACH THE NUMBER OF PUBLICATIONS SELECTED BASED ON TITLE AND ABSTRACT FROM THE FIRST 100 RESULTS.

	Search query	**select 2017**	**select 2023**
1.	“user study” perception style (“social robot” OR “virtual agent” OR “embodied conversational agent”)	4	9
2.	experiment perception style (“nonverbal behaviour” OR gesture OR posture) (“social robot” OR “virtual agent” OR “embodied conversational agent”)	17	10
3.	(gesture OR posture) (robot OR “virtual agent” OR “embodied agent”) (style OR mood OR personality OR attitude)	13	8
4.	(animation OR modulation OR manipulation) (gesture OR posture) (robot OR “virtual agent” OR “embodied agent”) (style OR mood OR personality)	28	13
5.	(“user study” OR experiment) perception (gesture OR posture) (robot OR “virtual agent” OR “virtual human” OR “embodied conversational agent”) (style OR mood OR personality OR stance OR warmth OR competence OR ~dominance OR ~affiliation)	26	19
6.	(“user study” OR experiment) (“user perception” OR evaluation) (gesture OR posture) (animation OR modulation OR manipulation OR modelling OR synthesis) (robot OR virtual OR agent) (style OR personality OR stance OR “social signal”)	12	7
7.	(“user study” OR experiment) perception (gesture OR posture) (style OR personality OR stance OR role) (“virtual human” OR “virtual character”)	27	8
8.	(“user study” OR experiment) perception (gesture OR posture) (style OR personality OR stance OR role) (“virtual agent” OR “embodied * agent” OR “pedagogical agent” OR “conversational agent”)	23	15
9.	(“user study” OR experiment) perception (gesture OR posture) (style OR personality OR stance OR role) robot	15	17
10.	(“user study” OR experiment) (perception OR impression) “social signal” (gesture OR posture) (style OR role OR character OR personality OR persona OR state OR strategy OR mood) (robot OR “* character” OR “* agent”)	3	5
11.	“user perception” (animation OR modulation OR manipulation OR model OR synthesis) (gesture OR posture) (style OR personality OR stance OR role OR mood)	7	4

2

Table 2.2

Characteristics of the studies reported in selected articles. Each row presents one study. Columns describe the characteristics: **Citation (C)** reference to bibliography; **Participants (N)** number of participants (S=students, A=academic staff); **Embodiment (E)** embodiment of the agent (R=physical robot + platform, V=virtual + H=humanlike, R=robotlike, W=wooden mannequin, S=stickfigure); **Stimuli (S)** presentation form (1=picture observation, 2=video observation, 3=lab observation, 4=lab interaction, 5=interaction 'in the wild'); **Exposure (T)** duration of exposure to stimulus; **Manipulation (IV)** modality and type of manipulation; **Parallel (P)** measure effect of manipulations applied in parallel (N=no, Y=yes, +=manipulations outside the scope of this review applied as well); and **Perception (DV)** measured social/affective construct.

ID	Citation	Participants	Embodiment	Stimuli (presentation form)	Exposure (duration per stimulus)	Manipulation (IV)	Parallel	Perception (DV)
1	[102]	244	V (w)	2	n/a	**Full-body** (motion) Dynamics (LMA)	Y	Personality (OCEAN)
	result:	Correlations reported for openness with indirect space and free flow, conscientiousness with direct, sustained and bound, extraversion with indirect, sudden and free flow, agreeableness with light weight, and neuroticism with indirect, sudden and free motion. All independent of action (picking or pointing).						
3	[103]	435	V (w)	2	10sec	**Arms** (gesture) Expansiveness Dynamics (time, fluency)	N	Personality (OCEAN)
	result:	Separate positive correlations reported for extraversion by Y translation, Z translation, scale, finger extension and average velocity, openness by scale, tension and wrist extension. Separate negative correlations reported for openness with dis-fluency, both agreeableness and emotional stability with tension, arm swivel, dis-fluency, clavicle lift, velocity warp and average velocity, and conscientiousness with dis-fluency and clavicle lift.						
4	[103]	n/a	V (w+h)	2	n/a	**Arms** (gesture) Expansiveness Dynamics (fluency)	Y	Personality (OCEAN)
	result:	Positive correlation reported for extraversion with plasticity adjustment (Y-translation, Z-translation, stroke-scale and finger extension). Negative correlations reported for emotional stability, conscientiousness and agreeableness with stability adjustment (dis-fluency, clavicle lift, and arm swivel). Results for openness were inconclusive depending on a specific agent and behaviour.						
5	[104]	160	V (h)	2	90sec	**Arms** (gesture) Expansiveness Dynamics (time) Repetition	Y	Personality (OCEAN)
	result:	Positive correlations reported for extraversion and agreeableness, and negative for neuroticism with amplitude (scale), horizontal and saggital direction, rate (frequency) and speed. No significant results for conscientiousness and openness.						
6	[105]	17a	R (icat)	3	n/a	**Head** Expansiveness Approach Dynamics (time) Repetition	Y+	Personality (OCEAN)
	result:	Positive correlations for extraversion, agreeableness and openness, and negative for conscientiousness, with speed, frequency, and vertical and horizontal direction. No correlations found for neuroticism.						
7	[106]	59	V (h)	2	15sec	**Arms** (gesture) Expansiveness Approach Dynamics (time) **Torso** (motion) Repetition	Y	Personality (OCEAN)

Continued on next page

Table 2.2 – *Continued from previous page*

ID	C	n	E	S	t	IV	p	DV
	result:	Positive correlation for extraversion with scale, vertical and horizontal position, speed, shoulder position (clavicle lift) and elbow position (arm swivel). Negative correlation for extraversion with duration. No significant results for other personality traits.						
8	[106, 107]	61	V (h)	2	15sec	**Full-body** (posture & gesture) Expansiveness Approach Dynamics (time, fluency)	Y	Personality (OCEAN)
	result:	Positive correlations for extraversion and agreeableness -and negative for neuroticism- with smooth gestures (scale, horizontal direction, vertical retraction-position and motion-connection smoothness, collarbone saggital position (backwards). Negative correlations for extraversion and agreeableness -and positive for neuroticism- with horizontal retraction-position, collarbone vertical position (down), arm swivel and knee bend (leg swivel).						
9	[108]	30	V (h)	2	n/a	**Arms** (gesture & position) Expansiveness Dynamics (time, fluency) Repetition Other (presence)	Y	Personality (Neuroticism, Agreeableness)
	result:	No correlations for neuroticism and agreeableness with expansiveness, rate, size (scale), shift-speed, collarbone horizontal and vertical, elbow horizontal (arm swivel). Negative correlation for neuroticism with presence of self-adaptors (not for agreeableness)						
10	[109]	40s	V (h)	2	n/a	**Full-body** Expansiveness Approach Dynamics (time) Repetition	Y	Extraversion
	result:	Positive correlation for extraversion with motion scale, gesture scale, position, and duration, clavicle lift, arm swivel, forward-lean, and gesture rate.						
11	[110]	48s	R (AIBO)	4	25min	**Full-body** (motion) Expansiveness Dynamics (time) Repetition	Y+	Extraversion, Intelligence
	result:	Positive correlation for extraversion with motion angle (wideness), speed and frequency. No correlation found for perceived intelligence.						
12	[111]	45s	R (NAO)	3	127-195sec	**Arms & Head** Expansiveness **Arms** Dynamics (time) Repetition	Y+	Extraversion, Intelligence
	result:	Positive correlation for extraversion with head and arm motion size, speed (velocity), and frequency. Negative correlation for extraversion with head down. No correlation found for perceived intelligence.						
13	[112]	21s	R (NAO)	2	n/a	**Full-body** (motion) Repetition	+	Extraversion, Intelligence
	result:	Small positive correlation for extraversion with increased motion frequency. No correlations found for perceived intelligence.						
14	[113]	194	R (TOOMAS)	5	n/a	**Head** Approach Repetition	Y+	Extraversion
	result:	Positive correlation for extraversion with increased head nod, wink and gaze switching.						
15	[114]	40s	V (w)	4	n/a	**Full-body** (posture) Expansiveness Other (proximity)	Y+	Extraversion
	result:	Positive correlation for extraversion with posture expansiveness (spreading).						
16	[115]	10s	V (r)	2	n/a	**Full-body** (posture) Dynamics (strength) **Head** Approach	N	Extraversion

Continued on next page

Table 2.2 – *Continued from previous page*

ID	C	n	E	S	t	IV	p	DV
	result:	Separate positive correlations for extraversion with head orientation towards the user and a strong posture.						
17	[116]	31s	R (Amiet)	3	n/a	**Arms** (gesture) Expansiveness Dynamics (time) Repetition	N	Extraversion, Other (MBTI: Thinking, Performance)
	result:	Separate positive correlations for extraversion with large and frequent gestures.Perception of thinking-feeling were dependent on the extraversion dimension. Positive correlation for intelligence with gesture frequency. No correlations with gesture speed found.						
18	[117]	78s	V (h)	2	10min	**Arms** (gesture) Other (presence)	+	Role (Expert, Mentor, Motivator)
	result:	Positive correlations for motivator and mentor with emotional gestures. Positive correlations for expert and mentor with deictic gestures.						
19	[117]	71s	V (h)	4	n/a	**Arms** (gesture) Other (presence)	+	Role (Expert, Mentor, Motivator)
	result:	Positive correlations for expert-like (expert and mentor) with deictic gestures. No correlations for emotional gestures.						
21	[118]	16	V (s)	2	n/a	**Full-body** (motion) Expansiveness Approach	Y	Affect (happy, sad) Other (energetic, tired, masculine, feminine)
	result:	Correlations for happy with increased swing shoulder z, wrist z, x, knees y, x, head x, hip x,; increased tilt neck y, head y, and decreased tilt head z. Correlations for sad with decreased swing shoulders z, wrist y; decreased tilt shoulders y, head y, neck y, torsoU y, elbow y, hip y, and increased tilt head z, and neck z. Positive correlations for energetic with increased swing knees y, elbows x, y, shoulders y, z, wrist z, x, thighs z, y; and tilt head y. Correlations for tired with decreased swing ankles z, y; decreased tilt shoulders y, head y, wrist y, neck y and increased tilt head z, neck z, and torso Lz and Uz. Correlations for feminine with increased swing hip x, torsoL x, thighs, elbows x, wrists x; and decreased tilt shoulders y. Positive correlations for masculine with increased swing shoulders z, ankles y, knees y, head z, neck z, and elbows z.						
22	[119]	30s	R (ROLLY)	3	n/a	**Full-body** (motion) Expansiveness Dynamics (time) Repetition Other (proximity)	Y	Dominance, Friendliness
	result:	Positive correlation for friendliness with motion frequency, speed and forward direction. Negative correlation for friendliness with motion size (volume). Positive correlation for dominance with speed, and forward direction.						
24	[120, 121]	93	R (NAO, Reeti)	2	n/a	**Full-body** (posture, gestures) Expansiveness Dynamics (relaxation) Repetition	Y+	Dominance, Competence, Other (authoritativeness)
	result:	Positive correlations for authoritative and effective with postural openness and relaxation, increased gaze, nod, self-contact hips, illustrator and emblem gesture frequency. Negative correlations for authoritativeness and frequency of blink and self-contact hands or face.						
26	[122]	26s	R (NAO)	3	n/a	**Arms & Head** (gesture) Expansiveness **Arms** Dynamics (time) Repetition **Head** Approach	Y	Mood (valance, arousal)
	result:	Positive correlations for valance and arousal with amplitude, speed, decay-speed, head tilt (vertical head position) hand height, finger rigidness, and palm direction.Negative correlation for valance and arousal with hold-time. Positive correlations for valance and arousal with motion amplitude, repetition, speed, head tilt and hand height. No correlations for valence and arousal with decay-speed, hold-time, finger rigidness, and palm direction.						

Continued on next page

Table 2.2 – *Continued from previous page*

ID	C	n	E	S	t	IV	p	DV
27	[123, 124]	36s	R (NAO)	4	6-10min	**Arms & Head** (gesture) Expansiveness **Arms** Dynamics (time) **Head** Approach	Y	Mood (valence, arousal)
	result:	Positive correlations for valance and arousal with motion amplitude, speed, head tilt, palm direction (extraverted), and following gaze. Negative correlations for valance and arousal with hold-time and finger bend.						
28	[45]	34s	R (NAO)	5	25min	**Arms & Head** (gesture) Expansiveness **Arms** Dynamics (time) Repetition **Head** Approach	Y	Mood (VAD)
	result:	No correlations found for valance, arousal and dominance with gesture size (angle), speed, hold-time (fluency), repetition, head tilt, head orientation, hand height, palm direction and finger rigidness.						
29	[32]	18	R (Icat)	3	n/a	**Head** Expansiveness (tilt) Approach Dynamics (time)	Y	Emotion (VAD, positive, negative)
	result:	Separate positive correlations for arousal with acceleration and curvature. Positive correlations for dominance and valance with curvature. Negative correlation for dominance with acceleration.						
31	[125]	18s	V (h)	2	10sec	**Full-body** (motion) Dynamics (time) Approach Other (view angle)	N	Emotion (valence, arousal)
	result:	Positive correlations for valance and arousal with head-torso inclination (tilt). Positive correlation for arousal with speed. No correlation for valance with speed.Reported results were independent of observer angle.						
33	[126]	26s	R (NAO)	3	n/a	**Head** Expansiveness (tilt)	N	Emotion (VAD, anger, sadness, fear, happy) Other (pride, excited)
	result:	Positive correlations for valance, arousal and dominance with head tilt.						
34	[127]	32	V (w)	2	n/a	**Full-body** (motion) Expansiveness Dynamics (fluidity, time)	N	Emotion (anger, joy, fear, sadness)
	result:	Positive correlation for angry with body expansiveness. No correlations for angry with acceleration speed and motion speed. Separate positive correlations for joy with acceleration speed, motion speed and expansiveness. Separate negative correlations for sadness and fear with acceleration speed, motion speed and expansiveness.						
35	[127]	34	V (w)	2	n/a	**Full-body** (motion) Expansiveness Dynamics (fluidity, time)	Y	Emotion (anger, joy, fear, sadness)
	result:	Positive correlation for angry with expansiveness.No correlation for angry with acceleration speed and motion speed. Negative correlation for sadness and fear with acceleration speed, motion speed and expansiveness. Positive correlation for joy with acceleration speed, motion speed and expansiveness.						
37	[128]	49s	V (h)	2	n/a	**Full body** (posture) Approach Expansiveness Dynamics (weight transfer)	Y	Emotion (anger, sad)
	result:	No correlation found for sadness with head bend (forward), chest bend, abdomen twist and arm position. Small correlation for angry with backward head bend, absence of a backward chest bend, no abdominal twist, and arms raised forward and upward.						
42	[79]	98	V (h)	2	30-60sec	**Full-body** (posture, gesture) Expansiveness Approach Dynamics (weight) Other (presence, synchronisation)	Y+	Comtence, Warmth

Continued on next page

Table 2.2 – *Continued from previous page*

ID	C	n	E	S	t	IV	p	DV
	result:	Positive correlations for warmth and competence with gesture openness, non-upright posture, and presence of semantic gestures. Correlation for competence with gesture synchronisation, hand position centre-centre, and a stable posture.						
43	[129]	80s/a	V (h, r)	2	45sec	**Arms** (gesture) Other (presence)	N	Competence, Warmth
	result:	Positive correlations for warmth and competence with gesture presence.						
44	[130]	90	R (NAO)	4	5+13min	**Arms & Head** (gesture) Other (presence)	+	Competence, Agreeableness
	result:	No correlations found for competence and agreeableness with gesture presence.						
45	[131]	40s/a	R (Honda)	4	n/a	**Arms & Head** (gesture) Other (presence)	+	Competence, Friendliness
	result:	Small positive correlation for competence with gesture presence.						
46	[131]	41	R (Honda)	4	n/a	**Arms** (gesture) Other (presence)	N	Competence, Friendliness
	result:	No correlations for competence and friendliness with gesture presence.						
47	[132]	32s	R	4	15min	**Arms & Head** (gesture) Other (presence, proximity)	+	Intelligence
	result:	Small positive correlation for intelligence with gesture presence. The result was only found in male participants.						
48	[133]	20s	R (NAO)	4	n/a	**Head** (gesture) Other (presence)	N	Intelligence
	result:	Small positive correlation for intelligence with presence of head nod.						
50	[101]	15	V (h)	2	n/a	**Head** Expansiveness Approach Dynamics (time) Repetition (frequency)	Y	Other (Attentiveness)
	result:	Positive correlation for attentiveness with max down amplitude, max down speed, duration, number of oscillations and amplitude.						
51	[134]	101	R (NAO)	5	10min	**Full-body** (posture, gestures) Expansiveness Approach Dynamics (weight) Repetition Other (presence)	Y+	Competence, Warmth, Dominance, Affiliation
	result:	Positive correlation for competence with body stability, gesture frequency and gaze directed at audience. Positive correlations for warmth and affiliation with head tilt, body stability, gesture openness and frequency, gaze and body directed at audience, and presence of semantic gestures. Positive correlation for dominance with head tilt, gaze and body directed at audience, open gestures, and presence of semantic gestures.						
52	[134]	72	R (NAO)	5	10min	**Full-body** (posture, gestures) Expansiveness Approach Dynamics (weight) Repetition Other (presence)	Y+	Competence, Warmth, Dominance, Affiliation
	result:	Positive correlations for competence, warmth, and affiliation with body stability, gaze and body directed at audience, head tilt, frequent, open gestures, hand position peripheral and centre-centre, and presence of semantic gestures. Dominance correlation was dependent on the context (positive at school, negative at a heath-camp).						
53	[135]	772	R (NAO)	1	free	**Full-body** (posture) Expansiveness Other (view angle)	Y	Dominance
	result:	Positive correlation for dominance with head tilt, shoulder angle, leg expansiveness and extension. Results were independent of posture and observer angle.						
54	[78]	30	R (NAO)	3	free	**Full-body** (posture, gesture) Expansiveness	Y	Dominance, Valence, Arousal

Continued on next page

Table 2.2 – *Continued from previous page*

ID	C	n	E	S	t	IV	p	DV
	result:	Positive correlations for dominance and arousal with head tilt, shoulder angle, leg expansiveness and extension. No correlation for valance with head tilt, shoulder angle, and leg expansiveness and extension. Results were independent of gesture.						
55	[136]	71	V	4	3min	**Arms** (gesture, posture) Expansiveness Repetition Other (presence)	Y+	Competence, Warmth
	result:	No correlations for warmth and competence with gesture frequency, arm position and presence of semantic and beat gestures.						
56	[137]	29	V (h)	2	n/a	**Arms** (gesture) Expansiveness	N	Friendliness
	result:	Open waving more fiendly than closed waving.						
57	[137]	20s	V (h)	4	n/a	**Arms** (gesture) Expansiveness	Y+	Friendliness
	result:	Familiar and modelled friendly agent perceived as more friendly.						
58	[138]	30s	V (h)	4	20min	**Full-body** (posture, motion) Expansiveness Dynamics	Y+	Happy, Sad, Angry
	result:	Percentatge of correct recognition: happy 60.41%, angry 52.08%, sad 100%.						
71	[139]	51sa	V (h)	2	n/a	**Full-body** (posture, motion) Expansiveness Dynamics	Y	Dominance
	result:	Both gait and scenario have a significant effect on perceived dominance.						
59	[140]	204sa	R (NAO)	1	10sec	**Full-body** (posture) Expansiveness	Y	Dominance, Competence
	result:	Expansiveness significant for dominance and competence, not for treat and eeriness.						
60	[141]	400	V (h)	2	n/a	**Arm** (gesture) Expansiveness Approach	Y+	OCEAN Personality
	result:	Openness, conscientiousness, extraversion, agreeableness and neuroticism perception significant difference between positive and negative extremes when body movement is included alongside voice and/or facial expression. Body movement alone did not result is correct perception of personality.						
61	[142]	87	V (h)	4	15-20min	**Arm and Head** (gesture) Expansiveness	Y+	Dominance, Intelligence, Other (Cooperativity, Autonomy)
	result:	Significant effect of modelled dominance on perceived dominance and competence. No effect on perceived autonomy and cooperativity.						
62	[143–145]	30a	R (Pepper)	3	n/a	**Arm** (gesture) Expansiveness, Dynamics (speed)	N	Personality, Intelligence
	result:	Significant effects for both speed and amplitude on perceived extraversion, agreeableness and neuroticism, but not for openness and conscientiousness. Higher speed and amplitude associated with higher extraversion (except for the disengaging gesture) and higher neuroticism (except for the head-touching gesture), and lower agreeableness for only the disengaging gesture. No significant effect on Intelligence.						
63	[146]	n/a	R (NAO)	4	n/a	**Arm** (gesture) Expansiveness, Dynamics (speed)	Y	Extraversion, Intelligence
	result:	Extroverted robot perceived significantly more extroverted. Effect of extroversion of robot on perceived intelligence was not significant.						
64	[147]	36	R (ERICA)	2	n/a	**Arm** (gesture) Expansiveness	N	Other (Politeness)
	result:	Palm-based pointing gestures were perceived more polite compared to index-finger gestures. No significant differences were found between downward and forward gestures.						
65	[147]	26	R (ERICA)	2	n/a	**Arm** (gesture) Dynamics (Speed)	N	Other (Politeness)

Continued on next page

Table 2.2 – *Continued from previous page*

ID	C	n	E	S	t	IV	p	DV
	result:	Faster and shorter gestures were perceived least polite. No significant differences were found between default and slow gestures and default and longer hold time.						
66	[147]	31s	R (ERICA)	2	n/a	**Arm** (gesture) Expansiveness	N	Other (Politeness)
	result:	Hands facing up where perceived more polite than hands to the side.						
67	[148]	20sa	R (Ottoman)	4	n/a	**Full-body** (motion) Expansiveness, Approach, Dynamics	N	Dominance
	result:	Significant main effect of motion style on perceived dominance status.						
68	[149]	n/a	R (EMYS)	2	n/a	**Head** (tilt) Expansiveness	N	Other (Assertiveness)
	result:	Pride posture associated with perceived assertiveness.						
69	[150]	30s	R (Pepper)	3	5 sec	**Arm** (gesture) Expansiveness, Dynamics (Jerk)	Y	Valance, Arousal, Dominance
	result:	Arousal was well conveyed in 7 of the 8 motions, dominance in 5 and pleasure in 4. Overall results not obvious.						
70	[151]	85s	V (w)	2	10 sec	**Full-body** (motion) Expansiveness, Dynamics	Y	Fear, Sadness, Joy, Anger
	result:	Gestures expressing sadness, joy, and anger were perceived as their intended emotions. Fear was perceived as fear or sadness, and the difference was not significant.						
72	[152]	103	V	2	12 sec	**Full-body** (gesture, posture) Expansiveness, Dynamics	Y+	Arousal, Valence
	result:	Arms outward perceived as high valence. Forward lean decreased perceived valance.						
73	[152]	273	V	2	12 sec	**Full-body** (gesture, posture) Expansiveness, Dynamics	N	Arousal, Valence
	result:	All factors significant for perceived valance. For arousal arm gestures were significant, body lean not.						
74	[153]	134s	V	2	n/a	**Full-body** (motion) Expansiveness, Dynamics	N	Extraversion
	result:	characters that walk with a larger stride length and with a more pronounced belt line tilt are perceived as more extroverted. Characters that walk with a lower beat (e.g. higher speed) are perceived as more extroverted when the value of upper body twist is not high.						
75	[154]	84s	R (Cozmo)	4	n/a	**Full-body** (motion) Approach, Dynamics	Y+	Valence
	result:	Strong effect of movement on perceived valance.						

2.4. REVIEW RESULTS

In this section we present a quantitative overview of the studies surveyed, as well as discuss findings with considerable support for effects of manipulation of bodily shape or motion on perception of personality, social dimensions, and emotion.

We focus on effects of modulation of bodily shape or motion performance (e.g., postures and gestures, see Section 2.2.2) on the social perception factors listed in our taxonomy (see Section 2.2.4). Other manipulations and perceptions included in the studies but outside of the scope of this review are omitted. For example, likeability of the robot is excluded because it is not a perception of personality, social dimensions or emotion. Likewise, we omitted the effect of paralinguistic cues and facial expressions because they are not modulations of bodily movement or posture. All included studies are listed in Table 2.2. An overview of studied manipulations and perceptions is given in Table 2.3[4].

[4]This table is also available online https://docs.google.com/spreadsheets/d/121u4TZlnBGw_ii4-oXeVjhfIuPStqixUHrZY2aH1CdA/edit?gid=585335540

Bodily manipulations that did not *modulate* shape or motion (e.g., presence of a certain behaviour) are not reported in the table and only briefly discussed here.

2.4.1. STATISTICS

We considered a total of 58 papers, presenting 63 studies. Each study provides a detailed description of 1) modulation of bodily shape or motion, and 2) evaluation of effects of these manipulations on humans' perception of the SIAs personality, social dimensions, style, or emotion.

Manipulations These studies reported a total of 59 distinctive manipulations modulating either static posture or motion of one or more body parts. We classified these manipulations by type and differentiated between body parts based on the Taxonomy in section 2.2.2. The distributions of studies over the different types of manipulations is as follows: *expansiveness* (n=48); *approach* (n=25); *dynamics* (n=37); and *repetition* (n=19). The distributions of studies over the different body parts is as follows: *full-body* manipulations (n=28); *arms* (n=22); *head* (n=13); and *torso* (n=1). Full-body manipulation studies include those in which the combined set of manipulations applied in parallel to more than the arms and head results in one holistic manipulation. For example, **study *24*** manipulated, in parallel, body posture (expansiveness and dynamics), arm gesture (frequency) and head (gesture frequency); this study is listed in the table under full-body expansiveness, dynamics and repetition.

The majority of studies performed multiple (types of) manipulations in parallel (n=38). All such effects are listed in the table in the designated cells. Of these, fourteen addition-

Table 2.3
RESULTS OF EFFECTS OF MODULATION OF BODILY SHAPE OR MOTION ON USER'S PERCEPTION OF AN AGENT. LIGHT SHADED CELLS INDICATE PRESENCE OF STUDIES REPORTING SIGNIFICANT EFFECTS, DARK SHADED CELLS INDICATE SIGNIFICANT EFFECTS IN THE SAME DIRECTION REPORTED BY THE MAJORITY OF STUDIES IN THAT CELL. GREEN OUTLINES INDICATE A POSITIVE EFFECT, RED OUTLINES A NEGATIVE EFFECT.

Manipulation (Body part & Type)	Personality					Social Dimensons			Emotion							
Full-body	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism	Competence / Intelligence	Friendliness	Dominance	Dominance	Valance	Arousal	Happy	Sad	Anger	Fear	Other
Expansiveness	8	8	8, 10, 11, 15+, 74*	8	8	11, 24, 42, 51, 52	22, 42, 51, 52	22, 51, 52, 59, 67	24, 53*, 54, 71	54, 72	54	21, 34*, 35, 58, 70	21, 34*, 35, 37, 58- 70	34*, 35, 37, 70	34*, 35, 70	21, 24
Approach	8	8	8, 10, 74*	8	8	42, 51, 52	42, 51, 52	51, 52, 67		31*, 72, 73*, 75	31*, 73*	21,	21, 37	37		21
Dynamics	1, 8	1, 8	1, 8, 10, 11, 16*, 74*	1, 8	1, 8	11, 24, 42, 51, 52	22, 42, 51, 52	22, 51, 52, 67	24, 71	31*, 75-	31*	34*, 35, 58, 70	34*, 35, 37, 58, 70	34*, 37, 58, 70	34*, 35, 70	24
Repetition			10, 11, 13+			11, 13, 24, 42, 51, 52	22, 42, 51, 52	22, 51, 52	24							24
Arms	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism	Competence	Friendliness	Dominance	Dominance	Valance	Arousal	Happy	Sad	Anger	Fear	Other
Expansiveness	3*, 4, 5, 7, 60, 62*	3*, 4, 5, 7, 60, 62*	3*, 4, 5, 7, 12, 17*, 60, 62*, 63	3*, 4, 5, 7, 9, 60, 62*	3*, 4, 5, 7,9, 62*	12, 55, 61, 62*, 63	55, 56, 57	61	28	26, 27, 28, 73*	26, 27, 28, 69					17, 61, 64*, 66
Approach	7	7	7, 60	7	7, 60						73*					64*
Dynamics	3*, 4, 5, 7, 62*	3*, 4, 5, 7, 62*	3*, 4, 5, 7, 12, 17*, 62*, 63	3*, 4, 5, 7, 9, 62*	3*, 4, 5, 7, 9, 62*	12, 62*, 63			28	26, 27, 28, 69	26, 27, 28, 69					17, 65*
Repetition	5	5	5, 12, 17*	5, 9	5, 9	12, 55	55		28	26, 28	26, 28					17
Head	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism	Competence	Friendliness	Dominance	Dominance	Valance	Arousal	Happy	Sad	Anger	Fear	Other
Expansiveness	6	6	6, 12	6	6	12, 61		61	28, 29, 33*	26, 27, 28, 29, 33*	26, 27, 28 , 29, 33*	29, 33*	29, 33*	33*	33*	33*, 50, 61, 68
Approach	6	6	6, 14, 16*	6	6				28, 29, 69	26, 27, 28, 29	26, 27, 28, 29	29	29			50
Dynamics	6	6	6	6	6				29	29	29	29	29			50
Repetition	6	6	6, 14	6	6											50
Torso	Openness	Conscientiousness	Extraversion	Agreeableness	Neuroticism	Competence	Friendliness	Dominance	Dominance	Valance	Arousal	Happy	Sad	Anger	Fear	Other
Expansiveness																
Approach																
Dynamics																
Repetition	7	7	7	7	7											

*Studied isolated effects of individual manipulations.

ally included 'non-bodily' parameters (e.g., proximity, paralinguistic cues, facial expression). If a study investigated multiple manipulations but looked at individual effects, that effect is marked * in Table 2.3). For example, **study *3*** investigated the effect of dynamics and the effect of expansiveness of the arms as two independent manipulations on all five OCEAN personality traits. If a study investigated multiple manipulations of different body parts in parallel they are reported as full-body. Only if individual effects are studies results are reported per body part.

Finally, bodily manipulations that did not *modulate* shape or motion have been reported in 19 studies. Eleven of the 19 studies reported such manipulations *additional* to modulation of bodily shape or motion. The remainder (n=8) reported only non-modulating manipulations, specifically presence of one or more behaviour(s), in which case the study is only listed in Table 2.2 and not discussed further throughout this paper.

Overall we observed that most studies performed manipulations in parallel, and that expansiveness was the most manipulated across the different body parts.

Perceptions Reported outcome measures are divided in three classes, each with multiple factors, as described in Section 2.2.4. The distribution of studies looking at perception of robot *personality* (Five Factor Model [85]) is: openness (n=9); conscientiousness (n=9); extraversion (n=19); agreeableness (n=10); and neuroticism (n=10). The distribution of studies looking at *social dimensions* is: competence (n=12); friendliness (n=7); and dominance (n=5). The distribution of studies investigating perceived *emotion* is: dominance (n=8); valence (n=11); arousal (n=9); and the discrete emotions happy (n=7); sad (n=8); anger (n=6); and fear (n=4). Ten (10) studies manipulating bodily shape or motion investigate *other* outcome measures. This includes, for example, the perception of parenting *style* (**study *24***).

Overall we observed that perception of extraversion and positive/negativeness (in the form of either affective dimensions or happy/sad) were studied most frequently.

2.4.2. Effects of Manipulations on the Perception of Personality, Social Dimensions and Emotion

Visual inspection of Table 2.3 clearly shows that not all combinations of manipulations and perceptions have been investigated equally frequent. In fact, many manipulation-perception combinations have never been studied (58 out of 180 possible combinations for full-body, arms and head) and the majority of reported combinations have been studied only once or twice (71 out of 180). The remaining 51 combinations have been studied three or more times, but not necessarily with consisted reports on effects. We now report those effects with considerable support. Support is derived from the number of studies that show the same manipulation-perception effect. Considerable support is defined as follows: 1) an effect is found by two or more labs[5] and 2) agreement exists over the results in the majority of the studies with a minimum of three. These two criteria were met by 28 manipulation-perception pairs (highlighted in dark grey in Table 2.3). Please note that presence of a study in the table does not mean a significant effect was found,

[5] As presented in supplementary online material: https://docs.google.com/spreadsheets/d/121u4TZlnBGw_ii4-oXeVjhfIuPStqixUHrZY2aH1CdA/edit?usp=sharing

only that the study investigated the effect. Also note that study results including a more detailed description of the manipulation and whether or not the effect was significant can be found in Table 2.2. Only 20 pairs with considerable support showed significant positive or negative effects and will be discussed in detail below.

Personality

Evidence for effects of manipulation seems to converge to effects on the personality trait extraversion. It was found to positively correlate with increased full-body expansiveness and arm expansiveness, as well as increased frequency of bodily motion and arm gestures. Further, the dynamics of arm- or body movements influenced how extravert the agent was perceived by participants.

Expansiveness There is considerable support for a positive correlation between *expansiveness* and perceived *extraversion*. This holds for both full-body manipulation as well as arm gesture manipulation. The University of California studied the influence of *expansiveness* in a virtual human in six studies and reported positive correlations for increased scale of full-body motion (**study *10***), increased gesture stroke scale (**study *3***, ***4***, ***5***, ***7***, ***8***, ***10***), raised arm position (**study *4***, ***7***, ***8***, ***10***), outward arm position (**study *3***, ***4***, ***5***, ***7***, ***8***, ***10***), finger extension (**study *3***, ***4***), and reduced knee bend (**study *8***). In all but **study *3*** manipulations from multiple types were applied in parallel. Other labs reported similar effects. In study***15*** (Stanford) it was found that virtual wooden mannequins with more expansive postures were perceived as more extravert than those displaying enclosed postures, and virtual characters stride length led to increased perceived extraversion in **study *74***. Perceived extraversion of a humanoid robot was shown to correlate with gesture scale manipulated both individually in **study *17*** (KAIST) and **study *62***, and combined with speed in **study *63*** and with speed and frequency in **study *12*** (University of Twente). However, **study *60*** found that bodily movement alone was *not* enough to influence perceived personality, it was only successful combined with voice and/or facial expression.

Dynamics There is considerable support for a positive correlation between movement *speed* and perceived *extraversion*. This support mainly comes from arm gesture manipulation. The University of California (**study *3***, ***5***, ***7***) found that a virtual human was found to be perceived more extravert when arm gesture speed was increased. Similar correlations for arm gesture speed were reported with a NAO robot by the University of Twente in **study *12*** and University of Hamburg in **study *63***, and a Pepper robot by University of Glasgow in **study *62***. Further, **study *74***found extraversion was positively correlated with fast speed in full-body motion of a virtual agent. However, **study *17*** did not find any correlation between speed adjustments and users' perception of robot's level of extraversion.

Further evidence for the influence of *dynamics* on *extraversion* includes work on full-body dynamics but all these studies investigate different parameters within the *dynamics* category: **study *11*** reports effects of faster interpolation between full-body motions on the perception of extraversion of an AIBO robot; effects of strong postures compared to weak postures (**study *16***), sudden and free bodily movement (**study *1***), more relaxed gesture retraction position and rounded gesture phase connections versus tensed (**study *8***), and motion duration (**study *10***).

Repetition There is considerable support for a positive correlation between motion *frequency* and perceived *extraversion*. This comes from repetition of full-body motion as well as arm gestures. An increased frequency of bodily motions or gestures resulted in higher extraversion scores for virtual humans (**study *5***, ***10***) and various robots (**study *11***, ***12***, ***13***, ***17***). All but **study *17*** obtained results from parallel manipulation.

Discussion of other effects on personality Arm *expansiveness* and *dynamics* and its effect on perceived *openness* or *consciousness* has been studied frequently as well. However, for *openness* there was no agreement in the effect between the different studies, and for *consciousness* the majority of the studies reported no significant effects from either expansiveness or dynamics manipulations.

Social Dimensions

The influence of full-body manipulations was investigated for all three social dimensions. Considerable support was found for a positive correlation between approach manipulations and perceived competence, expansiveness manipulations and perceived dominance, and for the absence of a relation between manipulation of the frequency of bodily motions and perception of the agent's level of dominance.

Overall there is a large body of support for style being influenced by full-body movement (all cells contain at least one supporting study, see light-grey cells), but individually these other effects do not have considerable support, either because they are only from one lab, or because the results are inconsistent.

Approach Considerable support was found for a positive correlation between *approach / avoidance* manipulation and perceived *competence*. In combination with other manipulations, Nguyen et al. (**study *42***) manipulated the full-body behaviour of virtual humans and found that agents who directed their gaze at the audience were perceived as more competent than an 'inattentive' agent with gaze diversion, and a forward leaning agent was perceived more competent than agents with an upright posture. Gaze orientation was manipulated simultaneously with gesture openness, horizontal- and vertical arm position, posture stability and lean, and gesture frequency, but gaze diversion was compared against three conditions with varying values for all other parameters except gesture repetition. Posture lean was also manipulated simultaneously with other parameters, and values changed in parallel with posture stability, gesture openness and presence of specific type of gestures. Peters et al. adapted this model for a NAO robot and validated that gaze direction and body positioning towards the audience, modulated in parallel with posture stability, gesture frequency and presence of fiddling, positively influenced children's perception of the robot's competence in **study *51*-*52***.

Dynamics Considerable support was found for a positive correlation between full-body *dynamics* manipulation and perceived *competence*. This originates from studies of posture weight: **study *42*** found that a more stable posture (manipulated in parallel with other parameters) of a virtual human was correlated with increased perceived competence and a comparable model applied to a NAO robot also showed increased perceived competence in **study *51*** and ***52***by TU Delft. Further, a more relaxed posture (manipulated in parallel with other parameters) was also related to perceived competence of a NAO robot in **study *24***. However, speed of bodily motion was shown to *not* influence perceived competence of an AIBO robot in **study *11***. Similar, the speed of arm gestured showed not to influence perceived competence of a NAO robot in **study *12*** and ***63***, and a Pepper robot in **study *62***.

2

Expansiveness Considerable support for the absence of a relation between *expansiveness* arm manipulation and perceived *competence* was found in **study *12***, ***55***, ***62*** and ***63***. Only **study *61***reported a positive effect of expansiveness on perceived competence of a virtual human.

Considerable support was found for a positive correlation between *expansiveness* manipulation and perceived *dominance*. The influence of body expansiveness on user's perception of a robot's dominance expression was studied repeatedly at Delft University of Technology. In **study *51*** they found that schoolchildren perceived a NAO robot with tilted head and open gestures as more dominant than a robot with its head down and closed gestures. These perceptions resulted from parallel manipulation of gaze direction, body orientation and presence of specific type of gestures. In a follow-up study (***52***) conducted at a children's medical camp, conflicting results were found. Subsequently the individual effect of full-body expansiveness was studied in **study *53***, and showed positive correlations. Also other labs found expansiveness of a NAO's posture (**study *59***) and a robotised ottoman's motion length (**study *67***) to relate with perceived dominance. Contrary, **study *22*** found no effect for expansiveness of the Rolly robot's motion on perceived dominance.

Repetition Considerable support for the absence of a relation between *repetition* and perceived *dominance* was found in **study *22***, ***51***, and ***52***.

Discussion of other effects on social dimensions As can be seen in the table, overall several studies investigated the influence of *expansiveness*, *approach* and *dynamics* manipulations and *repetition* on perception of all social dimensions. However, when focusing on individual effects, there is not enough support to claim specific relationships between manipulations and perception. We highlight several studies.

A positive correlation between expansiveness in robots and perceived competence was found in **study *24***, while **study *11*** did not find any differences, and **study *42*** reported conflicting results.

While **study *22*** consistently reported positive effects for approach, dynamics, and repetition adjustments and negative effects for expansiveness on the perception of friendliness (warmth), **study *42***, ***51***, and ***52*** reported conflicting results. For example, in **study *42*** a virtual human perceived as high in warmth displayed open gestures while the hands were positioned centred before the body or more outward and lower, depending on the expressed level of competence. In other words, it is not clear what the effect of the hands is on the perception of warmth.

Further, presence of certain gesture (functions) has been studied repeatedly for the effects on social dimensions. We haven't done a structural analysis of this manipulation because it is not a modulation of non-verbal behaviour. We highlight some findings. Presence of (co-verbal) arm gestures was found to positively influence perceived competence (**study *43***, ***45***, and ***47***) and friendliness (**study *42***, ***43***, and ***46***). However, others did not find any (competence: **study *44*** and ***47***, friendliness: **study *45***) or inconclusive (competence: **study *42***) effects. Presence of head motions and/or gestures showed some positive results on perceived competence (**study *45***, ***47***, and ***48***), while others reported no significant results (**study *42*** and ***44***). Results for friendliness remain inconclusive.

Emotion

There is considerable support for a positive correlation between perceived *positiveness/negativeness* of the agents' behaviour and *expansiveness* manipulation. First, full-body expansiveness positively correlates with perceived joy (and negatively with perceived sadness). Second, head expansiveness (i.e., head tilt) positively correlates with perceived valence. Further, there is evidence that expansiveness influences perceived arousal and dominance, and there is evidence that manipulation of approach is related to high valence and arousal and vice versa. The relation between repetition and emotion has been studied remarkable infrequently.

Expansiveness Considerable support was found for a positive correlation between *expansiveness* manipulation and perceived *dominance*. This effect was shown as a result of postural openness of a NAO robot in **study *24*** and ***54***, and a virtual human in **study *71***. In all studies the stimuli displayed full body motion in where the openness manipulation was applied to the posture and thus influenced the motion in total.

Considerable support was found for a positive relation between *expansiveness* manipulation and perceived *happiness* and a negative relation between *expansiveness* and *sadness*.

Increased *expansiveness* was found to positively influence perceived *happiness*, while decreased expansiveness positively influenced perceived sadness. Lin et al. reported significant and individual correlations between expanded motions and perceived joy, and between contracted motions and perceived sadness in a virtual wooden mannequin in **study *34***. In a followup study (***35***) they defined a joyful motion as expanded, loose and fast, and a sad motion as contracted, stiff and slow, and applied these modulations to behaviours that are emotional by nature (e.g., flourishing or crestfallen posture). The recognition of emotions improved when expressive behaviour and its modulation were congruent. In **study *21***, a virtual stick figure's full-body expansiveness influenced perceived emotions of joy and sadness. Expansiveness was manipulated by the parameters vertical head position and direction, vertical- and horizontal arm position and direction, and leg spread and stretch. These manipulations were applied in parallel with *approach* adjustments (i.e., horizontal head position and direction, and saggital arm- and torso position and direction). **study *58*** and ***70*** studied recognition of happy and sad emotions displayed by postural expansiveness amongst other manipulations. Sadness was well perceived in both studies. Happiness was recognised just above chance in **study *58***with 60.41%. However, sadness expressed by head bend and enclosed arms as well as forward lean was not recognised in a virtual human in **study *37***.

Considerable support was found for a positive correlation between *expansiveness* manipulation and perceived *valence* and *arousal*. In **study *26*-*27*** Xu et al. showed that a NAO that raised its head, combined with display of more expanded and quicker arm gestures, was perceived more positive than a NAO with bend head and enclosed, slow gestures. This 'positiveness' manipulation effected perceived arousal as well. In **study *26***, they found that modulation of only 'important' expansiveness parameters (i.e., motion scale, vertical hand position, and head tilt) was equally capable of expressing valence as all parameters (i.e., motion scale, vertical hand position, head tilt, finger extension, and palm direction); however, the difference between happy and very-happy was not clear with important parameters only. In a subsequent interactive scenario (**study *28***), the same manipulation did not result in any differences in perceived valence or arousal between the positive and the negative robot. Positive results on perceived robot valence were found by others as well for the influence of head tilt uniquely (***33*** and ***31***) and combined with head approach and movement speed (***29***). Both studies also reported a correlation with perceived arousal and dominance. Positive results on perceived arousal were reported in **study *69*** as well where arm expensiveness modulated together with motion jerk and gaze behaviour of a Pepper robot resulted in arousal being conveyed in 7 out of 8 motions. Pleasure was only conveyed in 4 of these motions. Positive results on perceived valence were reported in **study *73*** as well where arm expansiveness was related to perceived valence of a virtual agent. In this study, however, arousal showed related to forward arm gestures opposed to open or sideways gestures.

Approach Considerable support was found for a positive correlation between *approach* manipulation and perceived *valence* and *arousal*. Head *approach* (pan, orientation) modulations were found to positively influence perceived *valence* and *arousal* when evaluated in parallel with head tilt as presented above. Xu et al. found positive effects for head orientation, manipulated in parallel with head tilt and arm gesture expansiveness and dynamics, on perceived valence and arousal in a NAO robot in **study *26*-*27***, but could not confirm these findings in **study *28***. Similarly, **study *29*** found positive effects for head pan, manipulated in parallel with head tilt and head motion speed, on perceived valence, arousal and dominance in an iCat. Full body *approach* modulations were found to positively influence perceived *valence*. Movement approach combined with arm lift and motion jerk displayed while interacting with a Cozmo robot was shown to related to perceived valence in **study *75***. Cheng found in **study *73***that a forward body lean increased perceived valence, however, when combined with forward or sideways arm gestures the forward body lean whas negativly related to perceived valence (**study *72***).

Dynamics Considerable support was found for a positive correlation between *dynamics* manipulation and perceived *arousal* and *happiness*.

Xu et al found an effect of motion speed, fluency and finger rigidness (when combined with other parameters as presented above) in **study *26*** (observation) and ***27*** (lab interaction). However, in a follow-up study where participants interacted with the NAO robot in the wild (**study *28***) these findings were not confirmed. Another lab (**study *69***) studied the effect of motion smoothness on perceived arousal and found that arousal was conveyed in smooth motions displayed by a Pepper robot. The two labs studied different dynamics, thus the criteria for considerable support were not met for each individual dynamic.

In a first study of individual modulations (**study *34***) Lin found negative relations of motion stiffness and motion speed on perceived happiness, and no significant effect of motion fluidity were reported. In a follow-up study (**study *35***) with combined parameters fluidity, stiffness, speed, power and expansiveness Lin reported a positive relation between perceived joy and motions that are loose, fast and expanded. Happy when modelled as direct gaze (approach), backward neck bed (expansiveness), body expansiveness and rapid movements showed a 60% recognition rate in **study *58***. Joy modelled as backward neck bend, frontal or upward arm stretch, opening/closing of arms, expansive, high activity and high dynamics was perceived as intended in **study *70***, this study did not elaborate on the implementation of high dynamics. Again, these studies studied different dynamics, per individual dynamics the criteria for considerable support have not been met.

Considerable support for the absence of a relation between *dynamics* and perceived *anger* was found in **study *34***, ***37*** and ***58***. Only **study *70*** reported a positive relation between perceived anger and high motion dynamics combined with expansiveness modulations.

2.5. Discussion

We considered 2200 articles derived from 2 × 11 Google Scholar queries (once in 2017 and again in 2023). We reviewed 58 articles that presented 63 studies that met our inclusion criteria: 1) present a detailed description of displayed non-verbal behaviour and applied (parameter-based) manipulations in bodily shape and/or motion as displayed by a SIA; and 2) present the result of a user study evaluating how humans perceive the agent displaying these behaviours. The studied manipulations show low consistency in terminology and vocabulary, and are scattered over 59 unique parameters adjusting bodily shape and/or motion (ranging from fixed postures to single joint modulations) and 26 (5 personality traits, 3 social dimensions, 7 emotion, and 13 ‘other’) outcome measures, measured using standardized instruments for humans (e.g., SAM, Godspeed, BFI-K) or various Likert- or differential scales. To compare results, and create a conclusive model of behavioural manipulations to express specific social signals, first a unified language needed to be developed to represent behavioural manipulations. As a first step towards this direction we created a taxonomy of manipulations (as independent variable) and perceptions (as dependent variable). Of the resulting manipulation-perception pairs we found considerable support for 30 pairs (highlighted and outlined in Table 2.3).

2.5.1. General Outcomes

Evidence suggests that the effects of bodily manipulation on perceived personality converge primarily on the trait of extraversion. It was found to positively correlate with increased full-body expansiveness and arm expansiveness, as well as increased frequency of bodily motion and arm gestures. Further, there is an effect of the dynamics of arm movements on perceived extraversion. This finding is consistent with psychological literature on perception, which claims that extraversion and conscientiousness are perceived most accurately [86] because these traits are more visual than the others [87, 88]. Extraversion is the most reliable observable trait from facial expressions in people [155], as well as the trait with the most recognisable differences between anonymous stick figures animated from male versus female speakers [156]. This also makes sense from a conceptual point of view, because this trait is most related to observable behaviour. There is in fact a conceptual problem with trying to express the other traits as these are either mental (neuroticism, conscientiousness) or relational (openness, agreeableness). It seems unlikely that simple manipulations of the body will results in reliable differences on these dimensions, especially if context and dialogue are not taken into account as an influencing factor. For perception of relational traits, we argue that interaction is necessary; impressions need to be formed based on interpersonal behaviour. Whereas intrapersonal perceptions may exist solely in someones mind and can be based on observing someone [157]. This implies that the design of most studies —including observational stimuli— is insufficient to study perception of inter-personal constructs.

The influence of full-body manipulations were investigated for all three social dimensions. Considerable support was limited and found a positive correlation between approach manipulation and perceived competence, expansiveness modulations and perceived dominance, posture weight and perceived competence and, for the absence of a relation between manipulation of the frequency of bodily motions and perception of the agents' level of dominance. Overall there is a large body of support for social dimensions being influenced by full-body movement (all pairs contain at least one supporting study, see light-grey cells in Table 2.3), but individually these other effects do not have considerable support, either because they are only from one lab, or because the results are not consistent. Therefore it seems that there is potential for more research into this direction. In particular trying to disentangle individual relations between bodily manipulations and style perception. When looking at the perception of dominance in both style and emotion, there is in both considerable support for a positive correlation between expansiveness and perceived dominance. Conceptually it makes sense to group these two constructs together as it is not clear how, behaviourally, dominance in interaction should be seen as different from dominance in emotion. One of the studies reviewed (**study *54***) explicitly looked at the effect of only expansiveness on dominance and found a clear effect that was not present for valence, suggesting a selective effect on dominance perception for expansiveness manipulation.

There is considerable support for a positive correlation between perceived positive/negativeness of the agent's behaviour and expansiveness of the gestures and postures. Support comes from multiple effects. First, full-body expansiveness positively correlates with perceived happiness (and negatively with perceived sadness). Second, head expansiveness positively correlates with perceived valence. Further, there is evidence that expansiveness influences perceived arousal, and there is evidence that manipulation of approach is related to high valence/arousal and vice versa. Overall it can be seen that whenever perception of valence is influenced, arousal is also influenced. As this was also the case with the manipulation of expansiveness and dominance perception it seems that arousal as a separate factor seems to be hard to influence. This is interesting as arousal conceptually is the most observable factor. More research is needed to disentangle effects of bodily manipulation on perception of valence, arousal and dominance.

Some of the most influential manipulations we found have effects on multiple perception outcomes. It seems that head tilt and orientation, expansiveness (the total volume occupied by the body, but particularly the arms) and movement speed influence the perception of personality, social dimensions and emotion. However, manipulations of these factors are rarely very specific. For example, head tilt influences valence but also dominance and extraversion all in the same direction. Arm expansiveness also influences extraversion, dominance and valence, all in the same direction. We are tempted to interpret this as follows: *apparently the robot or virtual agent body can signal very important information, but the context may be essential in interpreting that signal.* Context here refers to not only the interaction context but also the experimental context. Asking a participant to rate on a P, A, D scale, will automatically collapse the interpretation of the behaviour on that scale as well. Therefore, a lot of the variation we see in our review may in fact be artificially introduced due to the different experimental setups (see also [158] for a related argument in affect detection databases). Direct evidence for context dependence is for example presented in **study *51*** and ***52***: children in schools and camps perceive robots differently. Further, in **study *24*** different effects were found depending on the type of robot; a NAO robot was perceived more dominant with increased gaze and a Reeti robot with decreased gaze. Further, inconsistent results may be due to different effects for different behaviours, this effect not found for very specific gestures (e.g., [78, 102, 135]) but supported by others for certain type of behaviours (e.g., [79, 117]).

The current study consists of articles derived from a search including works up to 2017 and a search conducted early 2023. In 2023 we see a continuation of the trends of the 2017 search. For personality expression consistent effects have still been reported for extraversion only. For emotion expression the results have been strengthened after 2023, but no new insights have been discovered. In studies of perception of social dimensions for a few independent-dependent variables pairs the addition studies resulted in reaching the threshold for considerable support, however, many pairs do not reach this threshold even though having been studied up to six times. This highlights the complexity of these valuable expressions.

2.5.2. Related Reviews

We have reviewed existing literature that studies the effect of parameter-based manipulation of bodily motion and shape on people's perceptions of a SIAs personality, social dimensions, style, and emotion. This work is different from other recent reviews in it's focused scope on non-verbal expressive behaviour, more specifically bodily movement and shape (i.e., kinesics) applicable to SIAs with a humanoid body, and social perceptions thereof by humans. Recently Saunderson & Nejat [53] published a survey on non-verbal communication in human-–robot interaction. They presented work on various ways humans are influenced by robots, including emotion recognition (and response) as well as cognitive framing, behaviour response and task performance. This review differs from ours in that they have a broader scope gaining insight about certain aspects of non-verbal behaviour including diversity of research in HRI. They focus on behavioural effects rather than validation of expressive behaviour, and whenever looking at perception they limited their search to emotion recognition. Further, they looked at a wide range of behaviours including kinesics as well as proximity, haptics, and chronemics. Nonetheless, they did present some work relevant and additional to the current review. For example, on recognition of 'emotive robot body movements and postures' [159], emotion conveyed in robot head movements [81], and 'interpretation of emotional intent [...] of only simple head movements and arm gestures' [82]. Although Saunderson & Nejat presented relevant studies on perception of non-verbal behaviour, more specifically kinesics, their works seem exemplary rather than a structured overview of existing knowledge on the topic.

Another recent work reviewed personality in HRI [160]. The work looked at both the user's and robot's personality and at perception as well as preferences and behavioural responses. Further it considered the big five and MBTI personality models as well as constructs such as intelligence, dominance, persuasiveness and credibility. With respect to robot personality it looked at whether people perceive the robot personality as intended, and the impact of robot personality on user's attitudes toward robots. The authors noted that extraversion is the most used personality trait and found that perceived extraversion was influenced by manipulating arm gestures, speech rate and volume, and speed and frequency of movements and that extraverted robots elicit higher acceptance. But studies showed contradicting perception results, possibly resulting from variation in context of use such as robot role and task. This work shows, like the current review, a disposition for extraversion within the personality research in HRI, possibly due to it's visible nature and salience in shorter interactions. Further it argues for context dependency of perceptions as well as we do.

2.5.3. Methodological Issues

One of the main findings of this review is that the field is diverse and fragmentary, so that it is difficult to derive more generic conclusive models of behavioural manipulations to express specific social signals.

An important first observation is that there is major inconsistency in the terminology used for both the (parameter-based) manipulations as well as the constructs measuring user perception. This is likely related to the multidisciplinary community (HRI vs IVA/ECA vs SSP), variation in embodiment, and the level of abstraction describing behavioural manipulations (e.g., reporting the manipulation as degree of expansiveness vs individual joint-angle values). This inconsistency makes interpretation, comparison and re-use of findings difficult.

A second observation is that the specificity with which the social signal is attempted to be expressed using bodily cues and manipulations varies significantly between studies. For example, valence was modelled in arm- and/or head gestures, full-body motion, and finger extension. This variety results in uncertainty about the underlying factors that drive the difference in perception.

Additionally, limited crossover between research groups, combined with significant variations in measurement, methods, and terminology, hinders the generalisation of findings. Following scattered results hinder exclusion of contextual influences and other biases, thus care should be taken to generalise over settings and agents.

Lastly, the validity of perception effects for individual manipulations is hindered by the fact that most studies apply, and thus measure the effect of, multiple manipulated parameters in parallel. Further, manipulations were applied to different body parts, but when applied in parallel were treated as 'full-body' because we can not separate the effects. Moreover, the ecological validity and construct validity are influenced by methodological decisions. Interpersonal perceptions requires interaction, nonetheless, studies often rely on brief passive observations and these observations are measured using instruments that were not designed and validated for reporting perceptions about SIAs.

2.5.4. Limitations

This review tries to give a complete picture of the effect of bodily manipulation of humanoid bodies of SIAs on the perception of affect: personality, social dimensions, style and emotion. To do this we had to abstract away from the many details in the studies. Sometimes, studies seem to present conflicting or inconclusive results. This can be due to several reasons. First, different parameters within one manipulation type may be non-comparable, for example in the dynamics type, fluency and power may have different effects but are still grouped in one cell in our analysis. So, the individual studies should always be checked if questions arise. Our tables will help doing so. Second, different manipulation types may have been done in parallel which makes them suitable to validate a specific expression within a specific project but makes it impossible to separate the effect. For example, expansiveness may overshadow effects of head orientation if these two manipulation were not tested as separate factors. However, the effect of head orientation is still reported in such a case (and this might not be correct). Third, although some studies report a particular manipulation, in fact the manipulation was co-varied with another experimental variable on purpose to enhance the communicative holistic effect. For example, when hand position is not systematically varied between conditions but is 'optimised' in each condition for better expressing an overall social signal.

Further, we were interested in perceived style expressions, but found hardly any studies on this matter. There is a body of literature on inter-human teaching style (e.g., [161, 162]), yet none of the studies directly measured teaching style. Teaching styles were subject of interest, but were disentangled into the social dimensions [134] and one author studied parenting styles by measuring of dominance and authoritativeness [120]. Presumably, this is related to the limited capabilities of directly measuring perceived (teaching) style. There are some instruments measuring teaching style (e.g., [163, 164], but none were validated for perception by children (the typical target group of mentioned human-robot interaction studies - opposed to observational studies) neither for perception of robots. Additionally, we consider actual interactive scenarios necessary for perception of style, whereas the vast majority of studies (43 out of 63) rely on passive observations (i.e., display of a picture, video, or physical robot).

Finally, there are studies that we had to exclude that do manipulate non-verbal behaviour and do measure perception, but do not make clear *how* the manipulations were done in detail. This makes it impossible to analyse or replicate them (e.g., [165, 166]. This also includes data-driven approaches (e.g., [167, 168]) and model-based approaches that do not (fully) describe their model (e.g., [169, 170]), as well as studies with unclear outcome measures (i.e., [133]).

2.6. CONCLUSION & RECOMMENDATIONS

This review focuses on non-verbal expressions through parameter-based manipulations of bodily shape and motion of humanoid robots and virtual agents, and how these manipulations are perceived by humans. We present a comprehensive survey of work published until January 2023 in this area. We analyse and summarise the available design knowledge to be used in the research and development of SIA's parameter-based behaviour modulation with intentional expression of social signals. Overall conclusions include: 1) research in this field shows a wide variety in concept definitions, behavioural manipulations and evaluation methodology. Therefore, we derived a taxonomy from the literature and structured existing literature accordingly. This way, we could create an overview of what has been studied and relate the studies to provide insight into outcomes that have 'considerable support'. 2) some social signals can be reliably displayed by behaviour manipulation in the form of posture- or motion modulation or designed key-expressions. However, for most manipulation-perception pairs the number of studies is limited (and variation within studies too large) to establish considerable support for correlations. 3) In particular, expansiveness manipulations seem the most effective, specifically for expression of extraversion and positiveness (i.e., valence, and happy or sad).

2.6.1. KEY FINDINGS

Clarity in and agreement on terminology and vocabulary is missing but is important, as well as standardised evaluation methods and validated measurement instruments.

In terms of guidelines for modulating the humanoid bodies of SIAs we observed the following:

First, the expression of personality traits using virtual or robot bodies seems limited to the trait extraversion.

Second, the expression of social dimensions seems possible, but only when using the whole body, and more research is needed to disentangle individual effects on friendliness, competence and dominance.

Third, the expression of emotion seems possible, but restricted to generic positive versus negative signals.

Fourth, context, including setting and verbal behaviour, seems important for users for the correct interpretation of the expressive behaviour.

Our taxonomy, TAXOMOD, is a starting point to develop a shared understanding and interpretation of research objectives and outcomes, to position and compare each other's research, to explicate progress in this area, and to formulate a road-map.

2

REFERENCES

[1] O. Mubin, C. J. Stevens, S. Shahid, A. A. Mahmud, and J.-j. Dong, *A review of the applicability of robots in education,* Technology for Education and Learing **1**, 209 (2013).

[2] B. Lugrin, C. Pelachaud, and D. Traum, *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application* (ACM, 2022).

[3] T. Fong, I. Nourbakhsh, and K. Dautenhahn, *A survey of socially interactive robots,* Robotics and Autonomous Systems **42**, 143 (2003).

[4] C. Breazeal, *Role of expressive behaviour for robots that learn from people,* Philosophical Transactions of the Royal Society of London B: Biological Sciences **364**, 3527 (2009).

[5] A. Vinciarelli, M. Pantic, D. Heylen, C. Pelachaud, I. Poggi, F. D'Errico, and M. Schroeder, *Bridging the Gap Between Social Animal and Unsocial Machine: A Survey of Social Signal Processing,* Affective Computing **3**, 69 (2011).

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

[7] O. a. Blanson Henkemans, V. Hoondert, F. Schrama-Groot, R. Looije, L. L. Alpay, and M. a. Neerincx, *" I just have diabetes ": children ' s need for diabetes self-management support and how a social robot can accommodate their needs,* Patient Intelligence **4**, 51 (2012).

[8] A. L. Baylor, *Expanding preservice teachers' metacognitive awareness of instructional planning through pedagogical agents,* Educational Technology Research and Development **50**, 5 (2002).

[9] R. Looije, M. a. Neerincx, and V. D. Lange, *Children's responses and opinion on three bots that motivate , educate and play,* Physical Agents **2**, 13 (2008).

[10] M. Nalin, M. Tabor, L. Bergamini, and A. Sanna, *Children's perception of a Robotic Companion in a mildly constrained setting: How children within age 8 – 11 perceive a robotic companion Categories and Subject Descriptors,* in *Human-Robot Interaction 2011 Conference Workshop on Robots with Children* (2012).

[11] C. Buche, R. Querrec, P. D. Loor, and P. Chevaillier, *MASCARET : Pedagogical Multi-Agents System for Virtual Environment for,* in *International Conference on Cyberworlds* (2003) pp. 423–430.

[12] T. Smith, G. Affleck, B. Lees, and C. Branci, *Implementing a generic framework for a web-based pedagogical agent,* in *Annual Australassian Society for Computers in Learning in Tertiary Education* (Brisbane, Queensland, 1999).

[13] J. B. Janssen, C. C. Van Der Wal, M. a. Neerincx, and R. Looije, *Motivating children to learn arithmetic with an adaptive robot game,* in *Social Robotics*, Vol. 7072 LNAI (Springer, 2011) pp. 153–162.

[14] T. Belpaeme, P. E. Baxter, R. Read, R. Wood, H. Cuayáhuitl, B. Kiefer, S. Racioppa, I. Kruijff-Korbayová, G. Athanasopoulos, V. Enescu, R. Looije, M. Neerincx, Y. Demiris, R. Ros-Espinoza, A. Beck, L. Cañamero, A. Hiolle, M. Lewis, I. Baroni, M. Nalin, P. Cosi, G. Paci, F. Tesser, G. Sommavilla, and R. Humbert, *Multimodal Child-Robot Interaction: Building Social Bonds,* Journal of Human-Robot Interaction **1**, 33 (2012).

[15] I. R. Nourbakhsh, J. Bobenage, S. Grange, R. Lutz, R. Meyer, and A. Soto, *An affective mobile robot educator with a full-time job,* Artificial Intelligence **114**, 95 (1999).

[16] M. Bruijnes, *Social and emotional turn taking for embodied conversational agents,* in *International Conference on Social Computing and International Conference on Privacy, Security, Risk and Trust* (IEEE Computer Society, 2012) pp. 977–978.

[17] S. D'Mello, A. Graesser, and R. W. Picard, *Toward an affect-sensitive autotutor,* IEEE Intelligent Systems **22**, 53 (2007).

[18] M. L. Walters, M. Lohse, M. Hanheide, B. Wrede, D. S. Syrdal, K. L. Koay, A. Green, H. Hüttenrauch, K. Dautenhahn, G. Sagerer, and K. Severinson-Eklundh, *Evaluating the Robot Personality and Verbal Behavior of Domestic Robots Using Video-Based Studies,* Advanced Robotics **25**, 2233 (2011).

[19] B. Hartmann, M. Mancini, and C. Pelachaud, *Implementing expressive gesture synthesis for embodied conversational agents,* in *GW2005: Gesture in Human-Computer Interaction and Simulation*, Vol. 3881 LNAI (Springer, 2006) pp. 188–199.

[20] A. Aly and A. Tapus, *A model for synthesizing a combined verbal and nonverbal behavior based on personality traits in human-robot interaction,* in *HRI* (IEEE, 2013) pp. 325–332.

[21] M. Alseid and D. Rigas, *Users' views of facial expressions and body gestures in e-learning interfaces: an empirical evaluation,* 8th WSEAS International Conference on Software Engineering , 121 (2009).

[22] I. Leite, G. Castellano, A. Pereira, C. Martinho, and A. Paiva, *Empathic Robots for Long-term Interaction: Evaluating Social Presence, Engagement and Perceived Support in Children,* International Journal of Social Robotics **6**, 329 (2014).

[23] L. Riek, T.-C. Rabinowitch, P. Bremner, A. Pipe, M. Fraser, and P. Robinson, *Cooperative gestures: Effective signaling for humanoid robots,* in *HRI* (IEEE, 2010) pp. 61–68.

[24] M. Tielman, M. Neerincx, J. J. Meyer, and R. Looije, *Adaptive emotional expression in robot-child interaction,* in *HRI* (ACM, Bielefeld, 2014) pp. 407–414.

[25] J. Mori, H. Prendinger, and M. Ishizuka, *Evaluation of an embodied conversational agent with affective behavior,* in *Proceedings of the AAMAS03 Workshop on Embodied Conversational Characters as Individuals, Melbourne, Australia* (2003).

[26] N. Wang and J. Gratch, *Can virtual human build rapport and promote learning,* Frontiers in Artificial Intelligence and Applications **200**, 737 (2009).

[27] C. Zaga, M. Lohse, K. P. Truong, and V. Evers, *The Effect of a Robot's Social Character on Children's Task Engagement: Peer Versus Tutor,* in *ICSR*, Vol. 9388 LNAI (Springer, Paris, 2015) pp. 704–713.

[28] R. Looije, M. a. Neerincx, J. K. Peters, and O. a. B. Henkemans, *Integrating Robot Support Functions into Varied Activities at Returning Hospital Visits,* International Journal of Social Robotics (2016), 10.1007/s12369-016-0365-8.

[29] A. J. Cuddy, S. T. Fiske, and P. Glick, *Warmth and Competence as Universal Dimensions of Social Perception: The Stereotype Content Model and the BIAS Map,* Advances in experimental social psychology **40**, 61 (2008).

[30] B. Wojciszke, R. Bazinska, and M. Jaworski, *On the dominance of moral categories in impression formation,* Personality and Social Psychology Bulletin **24**, 1251 (1998).

[31] G. B. Moskowitz, *Social cognition: Understanding self and others* (Guilford Press, New York, 2005).

[32] M. Saerbeck and C. Bartneck, *Perception of affect elicited by robot motion,* Proceeding of the 5th ACM/IEEE international conference on Human-robot interaction - HRI '10 , 53 (2010).

[33] J. Linssen, M. Theune, and D. Heylen, *Taking things at face value: How stance informs politeness of virtual agents,* CEUR Workshop Proceedings **1119**, 71 (2013).

[34] B. T. T. Chee, P. Taezoon, Q. Xu, J. Ng, and O. Tan, *Personality of social robots perceived through the appearance,* Work **41**, 272 (2012).

[35] K. Zibrek and R. McDonnell, *Does render style affect perception of personality in virtual humans?* in *Proceedings of the ACM Symposium on Applied Perception* (ACM, 2014) pp. 111–115.

[36] A. Heloir, K. H. Kipp, and M. Kipp, *Using individual light rigs to control the perception of a virtual character's personality,* in *International Conference on Affective Computing and Intelligent Interaction* (Springer, 2011) pp. 117–124.

[37] M. Pantic, R. Cowie, F. D'Errico, D. Heylen, M. Mehu, C. Pelachaud, I. Poggi, M. Schroeder, and A. Vinciarelli, *Social Signal Processing: The Research Agenda,* in *Visual Analysis of Humans* (Springer London, 2011) Chap. 26, pp. 511–538.

[38] R. op den Akker, M. Bruijnes, R. Peters, and T. Krikke, *Interpersonal stance in police interviews: Content analysis,* Computational Linguistics in the Netherlands Journal 3 **3**, 193 (2013).

[39] M. Chollet, M. Ochs, and C. Pelachaud, *A multimodal corpus for the study of non-verbal behavior expressing interpersonal stances,* in *IVA Workshop Multimodal Corpora: Beyond Audio and Video* (2013).

[40] J. Botzheim, J. Woo, N. T. N. Wi, N. Kubota, and T. Yamaguchi, *Gestural and facial communication with smart phone based robot partner using emotional model,* in *World Automation Congress (WAC), 2014* (IEEE, 2014) pp. 644–649.

[41] G. Saponaro, G. Salvi, and A. Bernardino, *Robot anticipation of human intentions through continuous gesture recognition,* in *Collaboration Technologies and Systems (CTS), 2013 International Conference on* (IEEE, 2013) pp. 218–225.

[42] W. Breitfuss, H. Prendinger, and M. Ishizuka, *Automated generation of non-verbal behavior for virtual embodied characters,* in *Proceedings of the 9th international conference on Multimodal interfaces* (ACM, 2007) pp. 319–322.

[43] B. Ravenet, M. Ochs, and C. Pelachaud, *From a user-created corpus of virtual agent's non-verbal behavior to a computational model of interpersonal attitudes,* in *Proc. Int. Conf. on Intell. Virtual Agents*, Vol. 8108 LNAI (Springer, 2013) pp. 263–274.

[44] M. Destenhe, K. Hashimoto, and A. Takanishi, *Emotional gait generation method based on emotion mental model - Preliminary experiment with happiness and sadness,* 2013 10th International Conference on Ubiquitous Robots and Ambient Intelligence, URAI 2013 , 86 (2013).

[45] J. Xu, J. Broekens, K. Hindriks, and M. Neerincx, *Effects of Bodily Mood Expression of a Robotic Teacher on Students,* in *International Conference on Intelligent Robots and Systems* (IEEE, 2014) pp. 2614–2620.

[46] R. Hortensius and E. S. Cross, *From automata to animate beings: the scope and limits of attributing socialness to artificial agents,* Annals of the new York Academy of Sciences **1426**, 93 (2018).

[47] M. L'hommet and S. Marsella, *Expressing Emotion Through Posture and Gesture,* in *The Oxford Handbook of affective computing* (Oxford University Press, 2015) pp. 273–285.

[48] G. Ball and J. Breese, *Emotion and Personality in a Conversational Agent,* Embodied Conversational Agents , 189 (2001).

[49] I. Cohen, R. Looije, and M. A. Neerincx, *Child's recognition of emotions in robot's face and body,* in *Proceedings of the 6th international conference on Human-robot interaction* (ACM, 2011) pp. 123–124.

[50] M. L. Walters, D. S. Syrdal, K. Dautenhahn, R. Te Boekhorst, and K. L. Koay, *Avoiding the uncanny valley: Robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion,* Autonomous Robots **24**, 159 (2008).

2

2

[51] A. Cafaro, H. H. Vilhjálmsson, T. Bickmore, D. Heylen, K. R. Jóhannsdóttir, and G. S. Valgarosson, *First impressions: Users' judgments of virtual agents' personality and interpersonal attitude in first encounters,* Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **7502 LNAI**, 67 (2012).

[52] D. S. Syrdal, K. Dautenhahn, S. N. Woods, M. L. Walters, and K. L. Koay, *Looking good? Appearance preferences and robot personality inferences at zero acquaintance,* in *AAAI Spring symposium: multidisciplinary collaboration for socially assistive robotics* (2007) pp. 86–92.

[53] S. Saunderson and G. Nejat, *How robots influence humans: A survey of nonverbal communication in social human–robot interaction,* International Journal of Social Robotics , 1 (2019).

[54] C. Breazeal, K. Dautenhahn, and T. Kanda, *Social robotics,* Springer handbook of robotics , 1935 (2016).

[55] J. K. Burgoon, V. Manusov, and L. K. Guerrero, *Nonverbal communication* (Routledge, 2021).

[56] P. Ekman and W. V. Friesen, *The repertoire of nonverbal behavior: Categories, origins, usage, and coding,* semiotica **1**, 49 (1969).

[57] P. Ekman, *Emotional and conversational nonverbal signals,* in *Language, knowledge, and representation* (Springer, 2004) pp. 39–50.

[58] M. Danesi, *Kinesics,* (2006).

[59] R. L. Birdwhistell, *Background to kinesics,* ETC: A review of general semantics , 10 (1955).

[60] R. L. Birdwhistell, *Kinesics and context: Essays on body motion communication* (University of Pennsylvania press, 2010).

[61] E. T. Hall, R. L. Birdwhistell, B. Bock, P. Bohannan, A. R. Diebold Jr, M. Durbin, M. S. Edmonson, J. Fischer, D. Hymes, S. T. Kimball, *et al.*, *Proxemics [and comments and replies],* Current anthropology **9**, 83 (1968).

[62] E. T. Hall, *The hidden dimension,* (1966).

[63] M. Cook, *Experiments on orientation and proxemics,* Human Relations **23**, 61 (1970).

[64] L. K. Frank, *Tactile communication,* ETC: a Review of General Semantics , 31 (1958).

[65] T. J. Bruneau, *Chronemics and the verbalnonverbal interface,* The relationship of verbal and nonverbal communication **101** (1980).

[66] R. L. Birdwhistell, *Introduction to kinesics: An annotation system for analysis of body motion and gesture* (Department of State, Foreign Service Institute, 1952).

[67] U. Bellugi and E. S. Klima, *The signs of language,* Psychological Science **1**, 6 (1979).

[68] M. Bacigalupi, *The craft of movement in interaction design,* in *Proceedings of the working conference on Advanced visual interfaces* (1998) pp. 174–184.

[69] P. Bennett, N. Ward, S. O'Modhrain, and P. Rebelo, *Damper: a platform for effortful interface development,* in *Proceedings of the 7th international conference on New interfaces for musical expression* (ACM, 2007) pp. 273–276.

[70] K. K. Roudposhti, L. Santos, H. Aliakbarpour, and J. Dias, *Parameterizing interpersonal behaviour with laban movement analysis—a bayesian approach,* in *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on* (IEEE, 2012) pp. 7–13.

[71] D. Chi, M. Costa, L. Zhao, and N. Badler, *The emote model for effort and shape,* in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (ACM Press/Addison-Wesley Publishing Co., 2000) pp. 173–182.

[72] T. Nakata, T. Sato, T. Mori, and H. Mizoguchi, *Expression of emotion and intention by robot body movement,* in *Proceedings of the 5th international conference on autonomous systems* (1998).

[73] E. Z. McClave, *Linguistic functions of head movements in the context of speech,* Journal of pragmatics **32**, 855 (2000).

[74] A. E. Scheflen, *The significance of posture in communication systems,* Psychiatry **27**, 316 (1964).

[75] D. N. Kurien, *Body language: Silent communicator at the workplace,* The IUP Journal of Soft Skills **4**, 29 (2010).

[76] N. Abner, K. Cooperrider, and S. Goldin-Meadow, *Gesture for linguists: A handy primer,* Language and linguistics compass **9**, 437 (2015).

[77] D. McNeill, *Hand and mind: What gestures reveal about thought* (University of Chicago press, 1992).

[78] R. Peters, J. Broekens, K. Li, and M. A. Neerincx, *Robots expressing dominance: Effects of behaviours and modulation,* in *8th International Conference on Affective Computing and Intelligent Interaction (ACII)* (IEEE, Cambridge, UK, 2019) pp. 1–7.

[79] T.-h. D. Nguyen, E. Carstensdottir, N. Ngo, M. S. El-nasr, M. Gray, D. Isaacowitz, and D. Desteno, *Modeling Warmth and Competence in Virtual Characters,* in *Intelligent Virtual Agents*, Vol. 9238 (Springer, 2015) pp. 167–180.

[80] I. Poggi and D. Francesca, *Cognitive modelling of human social signals,* in *Proceedings of the 2nd international workshop on Social signal processing* (ACM, 2010) pp. 21–26.

2

[81] S. Embgen, M. Luber, C. Becker-Asano, M. Ragni, V. Evers, and K. O. Arras, *Robot-specific social cues in emotional body language,* in *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication* (IEEE, 2012) pp. 1019–1025.

[82] J. Li and M. Chignell, *Communication of emotion in social robots through simple head and arm movements,* International Journal of Social Robotics **3**, 125 (2011).

[83] K. Eriksen and V. E. Kress, *Beyond the DSM story: Ethical quandaries, challenges, and best practices* (Sage Publications, 2004).

[84] L. A. Pervin and O. P. John, *Handbook of personality: Theory and research* (Elsevier, 1999).

[85] L. R. Goldberg, *The structure of phenotypic personality traits.* American psychologist **48**, 26 (1993).

[86] D. Watson, *Strangers' ratings of the five robust personality factors: Evidence of a surprising convergence with self-report.* Journal of Personality and Social Psychology **57**, 120 (1989).

[87] O. P. John and R. W. Robins, *Determinants of interjudge agreement on personality traits: The big five domains, observability, evaluativeness, and the unique perspective of the self,* Journal of personality **61**, 521 (1993).

[88] D. C. Funder and K. M. Dobroth, *Differences between traits: properties associated with interjudge agreement.* Journal of personality and social psychology **52**, 409 (1987).

[89] I. Myers and P. Myers, *Gifts differing: Understanding personality type* (Nicholas Brealey Publishing, 2010).

[90] C. G. Jung and H. G. Baynes, *Psychological types: Translated with an introduction.* The Journal of Nervous and Mental Disease **60**, 221 (1924).

[91] C. H. Dupree and S. T. Fiske, *Universal dimensions of social signals: Warmth and competence,* in *Social Signal Processing* (Cambridge University Press, 2017) pp. 23–33.

[92] M. S. Mast and J. A. Hall, *The vertical dimension of social signaling,* in *Social Signal Processing* (Cambridge University Press, 2017) pp. 34–45.

[93] S. T. Fiske, A. J. C. Cuddy, and P. Glick, *Universal dimensions of social cognition: warmth and competence,* Trends in Cognitive Sciences **11**, 77 (2007).

[94] T. Leary, *Interpersonal diagnosis of personality,* American Journal of Physical Medicine & Rehabilitation **37**, 331 (1958).

[95] J. S. Wiggins, *An informal history of the interpersonal circumplex tradition,* Personality Assessment **66**, 217 (1996).

[96] P. Ekman and W. V. Friesen, *Constants across cultures in the face and emotion.* Journal of personality and social psychology **17**, 124 (1971).

[97] M. Lewis, *Self-conscious emotions: Embarrassment, pride, shame, and guilt,* Emotions , 623 (2000).

[98] P. Ekman, *Basic emotions in dalgleish t. e power t.(eds.), the handbook of cognition and emotion,* (1999).

[99] R. Plutchik, *Emotions: A general psychoevolutionary theory,* Approaches to emotion **1984**, 197 (1984).

[100] J. A. Russell and A. Mehrabian, *Evidence for a three-factor theory of emotions,* J. of res. in Personality **11**, 273 (1977).

[101] C. Oertel, J. Gustafson, and A. W. Black, *Towards building an attentive artificial listener: On the perception of attentiveness in feedback utterances,* in *ICMI* (ACM, Tokyo, 2016) pp. 21–28.

[102] F. Durupinar, M. Kapadia, S. Deutsch, M. Neff, and N. I. Badler, *PERFORM: Perceptual Approach for Adding OCEAN Personality to Human Motion Using Laban Movement Analysis,* ACM Transactions on Graphics (TOG) **36**, 6 (2016).

[103] H. J. Smith and M. Neff, *Understanding the Impact of Animated Gesture Performance on Personality Perceptions,* ACM Transactions on Graphics (TOG) **36**, 1 (2017).

[104] C. Hu, M. a. Walker, M. Neff, and J. E. F. Tree, *Storytelling Agents with Personality and Adaptivity,* in *IVA*, Vol. 9238 LNAI (Springer, Delft, 2015) pp. 181–193.

[105] B. Meerbeek, J. Hoonhout, P. Bingley, and J. Terken, *The Influence of Robot Personality on Perceived and Preferred Level of User Control,* Interaction Studies **9**, 204 (2008).

[106] K. Liu, J. Tolins, J. Fox Tree, M. Neff, and M. Walker, *Two Techniques for Assessing Virtual Agent Personality,* IEEE Transactions on Affective Computing **7**, 94 (2016).

[107] K. Liu, J. Tolins, J. E. F. Tree, M. Walker, and M. Neff, *Judging IVA personality using an open-ended question,* Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **8108 LNAI**, 396 (2013).

[108] M. Neff, N. Toothman, R. Bowmani, J. E. Fox Tree, and M. a. Walker, *Don't scratch! Self-adaptors reflect emotional stability,* Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **6895 LNAI**, 398 (2011).

[109] M. Neff, Y. Wang, R. Abbott, and M. Walker, *Evaluating the effect of gesture and language on personality perception in conversational agents,* in *IVA*, Vol. 6356 LNAI (Springer, 2010) pp. 222–235.

2

[110] K. M. Lee, W. Peng, S. a. Jin, and C. Yan, *Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human-robot interaction,* Journal of Communication **56**, 754 (2006).

[111] M. Joosse, M. Lohse, J. G. Perez, and V. Evers, *What you do is who you are: The role of task context in perceived social robot personality,* in *ICRA* (IEEE, Karlsruhe, 2013) pp. 2134–2139.

[112] K. Leuwerink, *A Robot With Personality : Interacting With a Group of Humans,* Tech. Rep. (University Twente, 2012).

[113] Y. Ludewig, N. Doring, and N. Exner, *Design and evaluation of the personality trait extraversion of a shopping robot,* in *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication* (2012) pp. 372–379.

[114] K. Isbister and C. Nass, *Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics,* International Journal of Human-Computer Studies **53**, 251 (2000), 91906217 .

[115] J. C. Park, H. Song, S. Koo, Y. M. Kim, and D. S. Kwon, *Robot's behavior expressions according to the sentence types and emotions with modification by personality,* in *ARSO Workshop* (IEEE, Seoul, 2010) pp. 105–110.

[116] H. Kim, S. S. Kwak, and M. Kim, *Personality design of sociable robots by control of gesture design factors,* in *RO-MAN'17* (IEEE, Munchen, 2008) pp. 494–499.

[117] A. L. Baylor and Y. Kim, *Simulating instructional roles through pedagogical agents,* International Journal of Artificial Intelligence in Education **15**, 95 (2005).

[118] S. A. Etemad and A. Arya, *Expert-Driven Perceptual Features for Modeling Style and Affect in Human Motion,* IEEE Transactions on Human-Machine Systems **46**, 534 (2016).

[119] J. Kim, S. S. Kwak, and M. Kim, *Entertainment robot personality design based on basic factors of motions: A case study with ROLLY,* Proceedings - IEEE International Workshop on Robot and Human Interactive Communication , 803 (2009).

[120] W. Johal, S. Pesty, and G. Calvary, *Towards companion robots behaving with style,* in *Ro-MAN*, Vol. 23 (IEEE, Endinbrugh, 2014) pp. 1063–1068.

[121] W. Johal, G. Calvary, and S. Pesty, *Non-verbal Signals in HRI: Interference in Human Perception,* in *ICSR*, Vol. 9388 LNAI (Springer, 2015) pp. 275–284.

[122] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx, *Bodily mood expression: Recognize moods from functional behaviors of humanoid robots,* in *ICSR*, Vol. 8239 LNAI (Springer, 2013) pp. 511–520.

[123] J. Xu, J. Broekens, K. Hindriks, and M. a. Neerincx, *Robot Mood is Contagious : Effects of Robot Body Language in the Imitation Game,* in *AAMAS* (Paris, 2014) pp. 973–980.

[124] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx, *Mood contagion of robot body language in human robot interaction,* Autonomous Agents and Multi-Agent Systems **29**, 1216 (2015).

[125] M. Inderbitzin, A. Valjamae, J. M. B. Calvo, P. F. M. J. Verschure, and U. Bernardet, *Expression of emotional states during locomotion based on canonical parameters,* 2011 IEEE International Conference on Automatic Face and Gesture Recognition and Workshops, FG 2011 , 809 (2011).

[126] A. Beck, L. Canamero, and K. a. Bard, *Towards an Affect Space for robots to display emotional body language,* 19th IEEE International Symposium on Robot and Human Interactive Communication Principe , 464 (2010).

[127] Y. H. Lin, C. Y. Liu, H. W. Lee, S. L. Huang, and T. Y. Li, *Evaluating emotive character animations created with procedural animation,* in *IVA*, Vol. 5773 LNAI (Springer, 2009) pp. 308–315.

[128] V. Vinayagamoorthy, A. Brogni, A. Steed, and M. Slater, *The Impact of a Character Posture Model on the communication of affect in an immersive virtual environment,* Proceedings - VRCIA 2006: ACM International Conference on Virtual Reality Continuum and its Applications **1**, 229 (2006).

[129] K. Bergmann, F. Eyssel, and S. Kopp, *A second chance to make a first impression? how appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time,* (Springer, 2012) pp. 126–138.

[130] N. Kramer, C. Edinger, and A. Rosenthal-von der Putten, *The Effects of a Robot's Nonverbal Behavior on Users' Mimicry and Evaluation,* in *IVA*, Vol. 5208 (2015) pp. 131–138.

[131] M. Salem, K. Rohlfing, S. Kopp, and F. Joublin, *A friendly gesture: Investigating the effect of multimodal robot behavior in human-robot interaction,* Proceedings - IEEE International Workshop on Robot and Human Interactive Communication , 247 (2011).

[132] V. Chidambaram, Y.-H. Chiang, and B. Mutlu, *Designing Persuasive Robots: How Robots Might Persuade People Using Vocal and Nonverbal Cues,* 7th Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '12) , 293 (2012).

[133] N. Segato, A. Krogsager, D. G. Jensen, and M. Rehm, *The Role of Physical Embodiment of Humanoid Robot Interaction: Focusing on Backchannel Head Nods in Danish First Meeting Encounters,* Communications in Computer and Information Science **435 PART I**, 583 (2014).

[134] R. Peters, J. Broekens, and M. A. Neerincx, *Robots educate in style: The effect of context and non-verbal behaviour on children's perceptions of warmth and competence,* in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (IEEE, Lisbon, 2017) pp. 449–455.

2

2

[135] R. Peters, J. Broekens, K. Li, and M. A. Neerincx, *Robot dominance expression through parameter-based behaviour modulation,* in *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents* (2019) pp. 224–226.

[136] B. Biancardi, C. Wang, A. Cafaro, and C. Pelachaud, *A Computational Model for Managing Impressions of an Embodied Conversational Agent in Real-Time,* in *8th International Conference on Affective Computing and Intelligent Interaction (ACII) Robots* (Cambridge, UK, 2019).

[137] T. Randhavane, A. Bera, K. Kapsaskis, K. Gray, and D. Manocha, *Fva: Modeling perceived friendliness of virtual agents using movement characteristics,* IEEE Transactions on Visualization and Computer Graphics **25**, 3135 (2019).

[138] T. Randhavane, A. Bera, K. Kapsaskis, R. Sheth, K. Gray, and D. Manocha, *Eva: Generating emotional behavior of virtual agents using expressive features of gait and gaze,* (Association for Computing Machinery, Inc, 2019).

[139] T. Randhavane, A. Bera, E. Kubin, K. Gray, and D. Manocha, *Modeling data-driven dominance traits for virtual characters using gait analysis,* IEEE Transactions on Visualization and Computer Graphics **27**, 2967 (2021).

[140] J. P. Stein, P. Cimander, and M. Appel, *Power-posing robots: The influence of a humanoid robot's posture and size on its perceived dominance, competence, eeriness, and threat,* International Journal of Social Robotics **14**, 1413 (2022).

[141] S. Sonlu, U. Güdükbay, and F. Durupinar, *A conversational agent framework with multi-modal personality expression,* ACM Transactions on Graphics **40** (2021), 10.1145/3439795.

[142] A. M. R. von der Pütten, C. Straßmann, R. Yaghoubzadeh, S. Kopp, and N. C. Krämer, *Dominant and submissive nonverbal behavior of virtual agents and its effects on evaluation and negotiation outcome in different age groups,* Computers in Human Behavior **90**, 397 (2019).

[143] B. Craenen, A. Deshmukh, M. E. Foster, and A. Vinciarelli, *Shaping Gestures to Shape Personalities: The Relationship Between Gesture Parameters, Attributed Personality Traits and Godspeed Scores; Shaping Gestures to Shape Personalities: The Relationship Between Gesture Parameters, Attributed Personality Traits and Godspeed Scores* (2018).

[144] A. Deshmukh, B. Craenen, A. Vinciarelli, and M. E. Foster, *Shaping robot gestures to shape users' perception: The effect of amplitude and speed on godspeed ratings,* (Association for Computing Machinery, Inc, 2018) pp. 293–300.

[145] B. G. W. Craenen, A. Deshmukh, M. E. Foster, and A. Vinciarelli, *Shaping gestures to shape personality: Big-five traits, godspeed scores and the similarity-attraction effect socially interactive agents track,* (2018).

[146] A. Andriella, H. Siqueira, D. Fu, S. Magg, P. Barros, S. Wermter, C. Torras, and G. Alenyà, *Do i have a personality? endowing care robots with context-dependent personality traits,* International Journal of Social Robotics **13**, 2081 (2021).

[147] C. T. Ishi, R. Mikata, and H. Ishiguro, *Person-directed pointing gestures and interpersonal relationship: Expression of politeness to friendliness by android robots,* IEEE Robotics and Automation Letters **5**, 6081 (2020).

[148] J. Li, A. Cuadra, B. Mok, B. Reeves, J. Kaye, and W. Ju, *Communicating dominance in a nonanthropomorphic robot using locomotion,* ACM Transactions on Human-Robot Interaction **8** (2019), 10.1145/3310357.

[149] R. Paradeda, M. J. Ferreira, C. Martinho, and A. Paiva, *Communicating assertiveness in robotic storytellers,* (Springer Verlag, 2018) pp. 442–452.

[150] J. A. Claret, G. Venture, and L. Basañez, *Exploiting the robot kinematic redundancy for emotion conveyance to humans as a lower priority task,* International Journal of Social Robotics **9**, 277 (2017).

[151] E. Yang and J. Ryu, *Do we perceive emotional gender gesture of virtual avatar as intended?* (2021).

[152] J. Cheng, W. Zhou, X. Lei, N. Adamo, and B. Benes, *The effects of body gestures and gender on viewer's perception of animated pedagogical agent's emotions,* (Springer, 2020) pp. 169–186.

[153] S. P. Badathala, N. Adamo, N. J. Villani, and H. N. Dib, *The effect of gait parameters on the perception of animated agents' personality,* (Springer Verlag, 2018) pp. 464–479.

[154] T. Law, J. de Leeuw, and J. H. Long, *How movements of a non-humanoid robot affect emotional perceptions and trust,* International Journal of Social Robotics **13**, 1967 (2021).

[155] P. Borkenau, S. Brecke, C. Möttig, and M. Paelecke, *Extraversion is accurately perceived after a 50-ms exposure to a face,* Journal of Research in Personality **43**, 703 (2009).

[156] M. Koppensteiner and K. Grammer, *Body movements of male and female speakers and their influence on perceptions of personality,* Personality and Individual Differences **51**, 743 (2011).

[157] E. Aronson, T. D. Wilson, and R. M. Akert, *Social psychology: The heart and the mind.* (1994).

[158] B. Dudzik, M. Jansen, F. Burger, F. Kaptein, J. Broekens, D. K. J. Heylen, H. Hung, M. A. Neerincx, and K. P. Truong, *Context in human emotion perception for automatic affect detection: A survey of audiovisual databases,* in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)* (2019) pp. 206–212.

2

2

[159] D. McColl and G. Nejat, *Recognizing emotional body language displayed by a human-like social robot,* International Journal of Social Robotics **6**, 261 (2014).

[160] L. P. Robert Jr, R. Alahmad, C. Esterwood, S. Kim, S. You, Q. Zhang, *et al., A review of personality in human–robot interactions,* Foundations and Trends® in Information Systems **4**, 107 (2020).

[161] A. F. Grasha, *A Matter of Style: The Teacher as Expert, Formal Authority, Personal Model, Facilitator, and Delegator,* College Teaching **42**, 142 (1994).

[162] T. Wubbels, M. Brekelmans, P. D. Brok, J. Levy, T. Mainhard, and J. V. Tartwijk, *Let's Make Things Better –Developments in Research on Interpersonal Relationships in Education,* in *Interpersonal relationships in education* (Sense Publishing, Rotterdam, 2012) Chap. 14, pp. 225–249.

[163] E. Robbers, P. V. Petegem, V. Donche, and S. Demaeyer, *Predictive validity of the learning conception questionnaire in primary education,* International Journal of Educational Research **74**, 61 (2015).

[164] J. D. Vermunt and Y. J. Vermetten, *Patterns in student learning: Relationships between learning strategies, conceptions of learning, and learning orientations,* Educational Psychology Review **16**, 359 (2004).

[165] N. Tan, C. Clavel, M. Courgeon, and J.-C. Martin, *Postural expressions of action tendencies,* Proceedings of the 2nd international workshop on Social signal processing - SSPW '10 , 53 (2010).

[166] G. Castellano, M. Mancini, C. Peters, and P. W. McOwan, *Expressive copying behavior for social agents: A perceptual analysis,* IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans **42**, 776 (2012).

[167] M. Destephe, A. Henning, M. Zecca, K. Hashimoto, and A. Takanishi, *Perception of emotion and emotional intensity in humanoid robots gait,* in *2013 IEEE International Conference on Robotics and Biomimetics (ROBIO)* (IEEE, 2013) pp. 1276–1281.

[168] A. Heloir, M. Kipp, S. Gibet, and N. Courty, *Evaluating data-driven style transformation for gesturing embodied agents,* Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) **5208 LNAI**, 215 (2008).

[169] W. Johal, G. Calvary, and S. Pesty, *A robot with style, because you are worth it!* in *CHI* (ACM, Toronto, 2014) pp. 1201–1206.

[170] N. Mirnig, G. Stollnberger, M. Giuliani, and M. Tscheligi, *Elements of Humor: How Humans Perceive Verbal and Non-verbal Aspects of Humorous Robot Behavior,* in *HRI* (ACM, Vienna, 2017) pp. 211–212.

3

The Effects of Synthesised Warmth and Competence Expression on Children's Perception of a Robot

Social robots are entering the private and public domain where they engage in social interactions with non-technical users. This requires robots to be socially interactive and intelligent, including the ability to display appropriate social behaviour. Progress has been made in emotion modelling. However, research into behaviour style is less thorough; no comprehensive, validated model exists of non-verbal behaviours to express style in human-robot interactions. Based on a literature survey, we created a model of non-verbal behaviour to express high/low warmth and competence—two dimensions that contribute to teaching style. In a perception study, we evaluated this model applied to a NAO robot giving a lecture at primary schools and a diabetes camp in the Netherlands. For this, we developed, based on expert ratings, an instrument measuring perceived warmth, competence, dominance, and affiliation. We show that even subtle manipulations of robot behaviour influence children's perceptions of the robot's level of warmth and competence.

Parts of this chapter have been published in proceedings of the 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), (2017) [1].

3.1. INTRODUCTION

Robots are emerging as collaboration partners in training and education (e.g., [2–4]), and the importance of social intelligence has been stressed [5]. Educational robots may take the role of tutor, tool or peer; learning from, about or with robots [6]. Different roles have been linked to learning activities. For example, for basic learning tasks, a peer robot was preferred over a tutor robot [7], but for language learning a tutor was preferred [8]. The quality of human teacher-student interactions is partly determined by the teacher's ability to adapt their style to the student and activity [9]. Thus, roles can be performed with various styles, and, like human educators, effective educational robots should be able to express appropriate styles.

Roles and style are being used in experimental human-robot interaction (HRI) (e.g., [10, 11]). However, these studies seldomly report a validated behaviour model. For example, in [10] a *motivator-robot* was intended to signal *empathy* and *trustworthiness* by listener behaviour (gaze, nod and 'listening expression'), but validation of this model has not been reported. Attempting to bridge the social intelligence gap between humans and machines, social signal processing (SSP) research focused on modelling of and synthesis of agent behaviour based on analysis of human behaviour (e.g., [12, 13]). Up to now, research focused on understanding human behaviour, and exploring effects of robot behaviour. However, the question of *how* to stylize and express this has received less attention. Beneficial personalization of robot behaviour requires understanding of the effect of interaction style, and validation of style perception.

3.2. BACKGROUND

The importance of robot behaviour style is grounded in educational psychology. Human educators adapt both content and style of communication to the learner and task at hand. This ability is believed to be crucial for effective and motivating educational interactions. Style refers to (opposed to more stable personality traits and role) the current observable behaviour—the way a role is performed. This can be trained and used strategically.

3.2.1. TEACHING STYLE

Teaching styles are behaviour patterns that affect information presentation and interaction. Style is considered a technique addressing differences in learning- and cognitive styles [14]. Grasha [9] constructed five teaching styles: *Expert* (transmitting information), *Formal authority* (providing feedback and establishing boundaries), *Personal model* (showing an example), *Facilitator* (encouraging critical thinking), and *Delegator* (available in the background during project work). Teaching style is selected based on student capabilities (e.g., knowledge, responsibility, and motivation), teacher's need for control, and teacher's willingness to build and maintain relationships. Certain clusters of styles appear most frequent in classrooms and were linked to specific situations and strategies. For example, teacher-centred styles (i.e., expert and formal authority) were linked to teacher control, and less capable students, while student-centred styles (i.e., facilitator and delegator) were linked to teachers willing to loosen control, and more capable students.

3.2.2. INTERPERSONAL CIRCUMPLEX

This model—also known as Leary's Rose—defines interaction stance by two axes: *dominance* and *affiliation* [15, 16]. The horizontal affective-axis depicts willingness to cooperate. The vertical dominance-axis depicts the degree of power. Commonly, the circumplex is partitioned into eight octants (Fig. 3.1a). Further, Leary's theory states two interaction rules: dominance is complementary and affiliation symmetric, meaning that, an opposed stance evokes opposed stance and dominance evokes submissiveness. A cooperative style has been linked to maintaining contact, a competitive style to aversion. For example, head nods [17], gaze, and open posture [18] are associated with high affiliation. A dominant style is commonly stereotyped as loud and obtrusive, while submissive style is believed to show discrete, unnoticed behaviour. For example, loudness, vocal control, and gaze are associated with perceived dominance [19, 20]. Moreover, dominant people are believed to lean forward, use more gestures, have open and up-right posture, and orient towards others [21]. However, a meta-analysis of the relation between dominance and non-verbal behaviours argued that previous work has been inconclusive or based on limited data, and concluded that the relation exists to different degrees and even directions, depending on the person and situation [22]. Nonetheless, the model is widely used in social skill training.

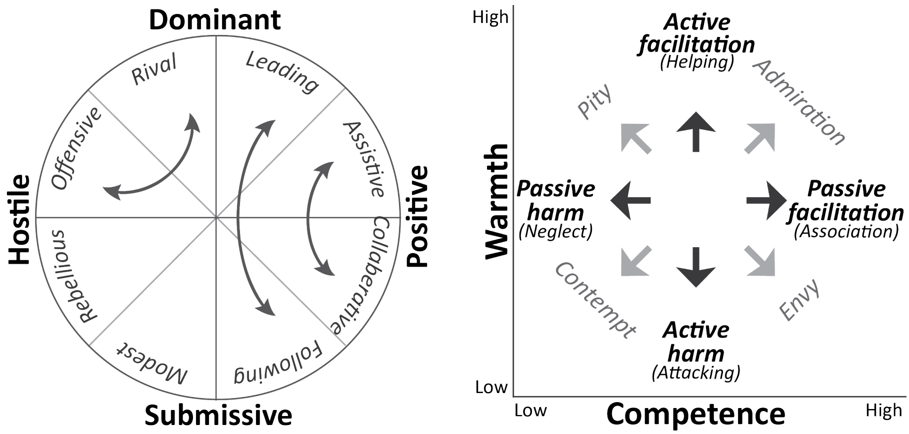
3.2.3. STEREOTYPE CONTENT MODEL

The stereotype content model (SCM) defines *warmth* and *competence* as two fundamental dimensions of social perception [23]. Warmth (kindness, empathy, friendliness and trustworthiness) evaluates valence of intent, competence (intelligence, power, efficacy and skill) assesses the ability to act on these intentions. Perceived warmth and competence generate emotions of admiration, envy, pity, and disgust towards someone, and predict active/passive and facilitative/harmful behaviour patterns (Fig. 3.1b). Perceived warmth is believed to be evoked by sincere smiles, head tilt, nodding, leaning forward, and open gestures [20, 24]. Coldness is expressed by closed hands, cutting motion, chin down, and the body pivoting away [24]. Upright posture and open gestures, are predictors of perceived power and associated with competence [20, 24]. Fiddling was suggested to signal low control and confidence, therefore resulting in low-competence judgement [24]. Additionally, warmth judgements are believed to be made before, and influence, competence evaluations—persons evaluated as warm, are likely to be judged more competent [20].

In summary, style can be defined as a behaviour pattern signalling our attitude towards a person or situation, affecting how others evaluate us and subsequently respond. Four social constructs important to the notion of style are discussed: *warmth*, *competence*, *dominance*, and *affiliation*.

3.3. RELATED WORK

In HRI, *interaction style* has been defined as a combination of behaviours that evoke a perceivable robot role (e.g., [25]). For example, in collaborative play, a NAO robot took the role of *peer* (collaborative behaviour) or *tutor* (scaffolding support) [11]. Recent studies investigated the effect of robot behaviour on user experience or outcome



(a) Interpersonal Circumplex

(b) Stereotype Content Model

Figure 3.1: Models of social interaction evaluation and prediction.

(e.g., [2, 11, 25–28]). However, validation of behaviour models is scarce. At best, evaluation was done with parents [27] or teachers [11]. It remains unclear if children perceived the robot as intended. Consequently, we cannot compare and build upon results.

Studies validating behaviour models have focused on emotion expression (e.g., [3, 29]) or personality (e.g., [30]). Although, recently interest has developed in modelling social perceptions such as competence and dominance in virtual humans (e.g., [31, 32]). Presence of gestures positively influenced perceived warmth and competence [33]. An elaborate model of virtual agent behaviours expressing warmth and competence appeared successful [13].

3.4. Robot Style Model

Based on the non-verbal behaviours associated with warmth and competence described above, we created a model of non-verbal behaviours for a NAO robot expressing four different style configurations: *high-warmth and high-competence* (HwHc); *high-warmth and low-competence* (HwLc); *low-warmth and high-competence* (LwHc); and *low-warmth and low-competence* (LwLc).

First, we selected cues applicable to our set-up. For example, the ‘sync’ cue was rejected because of limitations with respect to timing of behaviours proposed by the software (Section 3.5.1), and the NAO robot is incapable of facial expressions. The resulting model is presented in Table 3.1. Next, we applied our model to the robot by annotating available behaviours, creating a library of animations fitting the style configurations. Lastly, we added fitting behaviours to the text sequence (script). For example, when the robot says ‘Pay attention’, for HwHc the open *StateLeft* behaviour and *head-up* are selected, and for LwLc the closed *CapisceLeft* behaviour and *head-down* (Fig. 3.2).

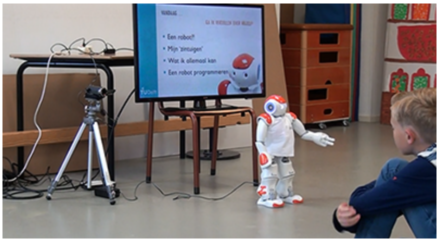
Table 3.1
OVERVIEW OF THE BEHAVIOUR CUES FOR THE FOUR ROBOT STYLE CONFIGURATIONS (WARMTH × COMPETENCE).
⋆ = COMPETENCE, • = WARMTH – = WARMTH × COMPETENCE. *CONDITIONS USED IN SECOND EXPERIMENT

High Warmth, High Competence*	**Low Warmth, High Competence**
Paralinguistic • low pitch • low volume *Body Posture* ⋆ stable • directed at audience • head tilt – gaze fixed at audience *Hand Gestures* • open • semantic & syntactic ⋆ frequent (every sentence) – mid-peripheral – centre-centre	*Paralinguistic* • high pitch • high volume *Body Posture* ⋆ stable • pivot away • chin down – gaze fixed at audience *Hand Gestures* • closed • semantic only ⋆ frequent (every sentence) – low-centre – mid-centre-centre
High Warmth, Low Competence	**Low Warmth, Low Competence***
Paralinguistic • low pitch • low volume *Body Posture* ⋆ wobbling • directed at audience • head tilt – gaze fixed at audience *Hand Gestures* • open • semantic & syntactic ⋆ low frequency – low/mid-peripheral	*Paralinguistic* • high pitch • high volume *Body Posture* ⋆ non-stable • pivot away • chin down – gaze diversion *Hand Gestures* • closed • semantic only ⋆ low frequency – low/high-centre/peripheral – fiddling

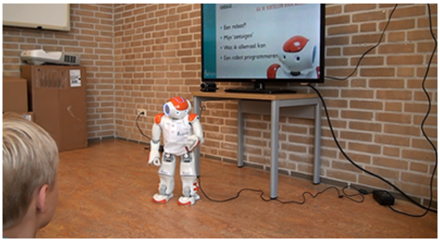
3.5. STUDY 1: PERCEIVED WARMTH AND COMPETENCE

3.5.1. METHOD

Evaluating the effect of non-verbal behaviour on children's perceptions of an educational robot's style, we conducted a 2 × 2 (*warmth* × *competence*) between-subject percep-



(a) High Warmth, High Competence



(b) Low Warmth, Low Competence

Figure 3.2: Video stills illustrating behaviour accompanying a statement in two of the four robot style configurations.

tion study at primary schools. For this, a PowerPoint presentation with text and animation script for the four robot styles was created using the RoboTutor framework (https://github.com/RoboTutor). The 10-minute lecture on robotics, given by a NAO robot (www.alderbaran.com), included three multiple choice questions, which could be answered using the Turningpoint polling system (https://www.turningtechnologies.com). The robot speech was accompanied by non-verbal behaviours, which varied between groups (*HwHc, HwLc, LwHc, LwLc*). Participant's perceptions of the robot's level of *warmth*, *competence*, *dominance*, and *affiliation* were collected after the lecture by rating 20 adjectives on a three-point Likert scale. Face validity was ensured by selecting adjectives at child level, and in correspondence with a developmental psychologist. Construct validity was ensured by expert ratings.

Measurement Perceived *competence*, *warmth*, *dominance*, and *affiliation* were measured by an adjective-based instrument developed for the present study. Children rated 20 adjectives on a three-point Likert scale. Perception score for each dependent variable were calculated by multiplying word-rating with a loading based on expert ratings.

The adjectives (translated from Dutch: *bossy, nagging, clumsy, friend, popular, playful, follower, loner, angry, honest, fights, knowledgeable, boring, nice, listener, confident, educational, helpless, helpful, dumb*) were chosen from words commonly used to describe various positions in Leary's rose or the SCM, and likely to be present in young children's vocabulary. Children rated whether each word would describe the robot (yes, sometimes/maybe, no) by placing a sticker in the corresponding column of a response leaflet. Stickers were used because physical activity was suggested to reduce the primacy effect (selecting extreme high or low for all items), thus increasing reliability [34]. Ratings were coded on a numeric interval scale [2–0], where a higher value is associated with better fitting description. Children did not place words they did not understand, these missing values were replaced by the population mean for that adjective.

The loadings for adjectives on the dependent variables were calculated from expert ratings. Eleven experts in human-computer interaction rated each adjective on four bipolar scales [-2–2]. We assessed the reliability of ratings, excluding words with a standard deviation of one or above. Further, words with a median of 0 were excluded for that construct because the association was weak. From the remaining words (10 for competence, 13 for warmth, 5 for dominance, and 11 for affiliation) we took the mean as loading value. This provides a table of 20 adjectives and their loading, if any, on each dependent variable (Table 3.2).

Participants A total of 101 children, from two primary schools in the Netherlands (S1, n = 40; S2, n = 61), participated in our study. Children at S1 were 10-13 years of age ($M = 11.43, SD = 0.64$), and enrolled in 5th (n = 9) or 6th (n = 31) grade. Children at S2 were aged 5-8 ($M = 6.52, SD = 0.65$), all enrolled in first grade. Gender was evenly distributed (S1 male = 20, female = 20; S2 male = 30, female = 28); three children did not report their gender. All participants were naive to the research aim and had little to no previous experience with a NAO robot. Children within each class were assigned to one of four robot style configurations. Children from the same school, in the same condition were merged into one group. This way, at both schools, all groups contained a minimum of 10 children, controlled for age, class and gender.

Table 3.2
Loadings for 20 adjectives (translated from Dutch) on the dependent variables. The values present the mean expert ratings for the adjectives on four bipolar scales, compliant with our criteria ($SD \geq 1$ and $Mdn \neq 0$).

Adjective	Competence	Warmth	Dominance	Affiliation
Bossy	–	-1.18	2.00	-0.55
Nagging	-0.73	-0.73	–	-1.09
Clumsy	-1.73	–	–	–
Friend	–	1.91	–	1.36
Popular	0.82	0.91	0.55	–
Playful	–	1.36	–	1.82
Follower	–	–	–	–
Loner	–	–	–	–
Angry	–	-1.18	1.09	-1.09
Honest	0.82	1.00	–	1.46
Fight	–	-1.36	1.36	-1.64
Knowledgeable	1.64	–	–	–
Boring	–	-0.64	–	–
Nice	–	1.82	–	1.09
Listener	0.91	1.09	–	1.46
Confident	–	-0.55	1.00	–
Educational	1.64	–	–	0.91
Helpless	-1.45	–	–	–
Helpful	0.64	1.27	–	1.64
Dumb	-1.91	–	–	–

Procedure Before the experiment, the researcher was briefly introduced to the children in the classroom, and children were assigned a group. Afterwards, when all children had participated, children were given the opportunity to ask questions about the robot and experiment, and a demonstration was given. The following steps were repeated for each group:

- children entered, were seated and given a 'clicker';
- researcher introduced the robot and instructed the children to remain seated after the lecture;
- researcher started the selected script;
- robot gave a lecture displaying stylised behaviours;
- researcher explained the questionnaire, accompanied by a brief example rating the popular Disney figure Simba;
- children spread across the room;
- researcher handed the children stickers with 20 adjectives, followed by a response leaflet and pencil; and
- children provided their individual ratings.

3.5.2. Results

We explored differences in children's perceptions of robots displaying high/low warmth and competence-related behaviours. Although, K-S tests indicated non-normal distributions, we decided to perform a MANOVA because we are interested in the interaction effect and the sample size is fair.

Using Pillai's trace, there was a near significant interaction effect of intended warmth × competence on how children perceived the robot, $V = 0.08, F(4,95) = 2.09, p = 0.088$. Separate univariate ANOVAs on the outcome variables revealed a significant interaction effect for intended warmth × competence on perceived affiliation, $F(1,97) = 4.42, p = 0.038$; and warmth, $F(1,97) = 4.09, p = 0.046$. Further, near significant main effects were found for intended competence on perceived competence, $F(1,97) = 3.30, p = 0.072$; and intended warmth on perceived dominance, $F(1,97) = 3.81, p = 0.054$.

In other words, children perceived a robot displaying high-competence behaviours (stable posture, frequent gestures) as more competent ($M = 10.00$) than low-competence (unstable posture, low frequency of gestures) ($M = 9.20$), regardless of the intended level of warmth (Fig. 3.3a). And a robot displaying high-warmth behaviours was perceived slightly warmer ($M = 16.18$) than low-warmth robots ($M = 15.83$). However, competence influenced the effect of warmth manipulations on how children perceived the robot; robots displaying high-warmth behaviours were perceived as more warm than low-warmth robots, but only in the high-competence condition. High-competence robots displaying high-warmth behaviours were perceived warmer ($M = 16.58$) than low-warmth robots ($M = 15.42$), low-competence robots displaying high-warmth behaviours were perceived considerably less warm ($M = 15.76$) than low-warmth robots ($M = 16.21$) (Fig. 3.3b). See also Table 3.3 for mean perception scores.

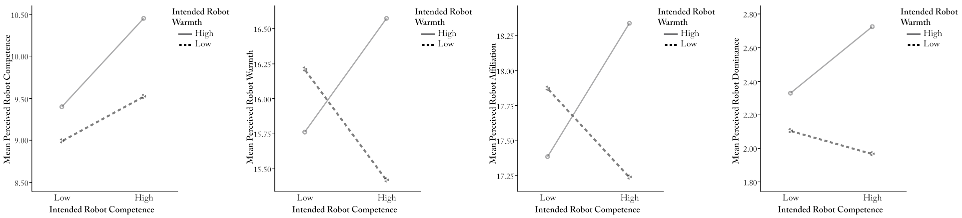
To investigate the bias of individual words on perception scores, we explored differences in word-ratings. We performed a non-parametric Mann-Whitney tests, once for each construct, because normality could not be assumed. Comparison of high and low competence samples showed a significant difference in ratings for 'Follower' ($U = 913.00, z = -2.49, p = 0.013, r = -0.25$) and 'Helpless' ($U = 798.50, z = -2.69, p = 0.007, r = -0.28$); near significant differences were present for 'Helpful' ($U = 933.00, z = -1.83, p = 0.068, r = -0.19$). No difference in word-ratings were found for 'Dumb', 'Nice', 'Knowledgeable', and 'Nagging'. Comparison of high and low warmth samples showed near significant differences for 'Popular' ($U = 1290.00, z = 1.83, p = 0.067, r = 0.19$), 'Loner' ($U = 991.50, z = -1.76, p = 0.078, r = -0.18$), 'Angry' ($U = 1323.00, z = 1.72, p = 0.086, r = 0.17$), 'Confident' ($U = 1305.00, z = 1.65, p = 0.099, r = 0.17$), and 'Helpful' ($U = 1307.00, z = 1.83, p = 0.068, r = 0.19$). Since all children reported exactly the same rating for 'Dumb', there was once again no difference between samples.

3.5.3. Discussion

The results indicate that children did perceive robots differently based on their non-verbal behaviours—the set-up, PowerPoint, textscript and questions were similar for all groups. However, the differences are small. This may be due to subtlety of the manipulation or measurement.

Variations in non-verbal behaviour were subtle, attempting to avoid creating a caricature and keep behaviour believable. However, observers of all four conditions were not able to note any difference, indicating that we may have been too careful selecting and adapting behaviours. This can explain the small differences in perception scores between robot styles. Future studies exploring the effect of enlarged differences in behaviour between styles is needed. Overall perception of warmth and competence was high, indicating that children have a positive stance towards the robot.

The measurement instrument, which we created based on literature and in correspondence with experts, was partly validated using expert ratings. Children confirmed to understand all words, with one exception (Popular) for the youngest children. However, children do not necessarily have the same interpretation of words. Analysis of individual word-ratings revealed that children provided significantly different ratings for words not included in calculation of the outcome measures; ratings of 'Loner' differed between high/low intended warmth, but this word did not contribute to calculation of perceived



(a) competence (range -11.64 to 12.91).
(b) warmth (range -11.27 to 18.72).
(c) affiliation (range -8.73 to 19.46).
(d) dominance (range 0.00 to 12.00).

Figure 3.3: Study 1 results showing mean perception scores. Competence condition displayed on horizontal axis, warmth by separate lines.

warmth. Similar, was the case for 'Follower' on competence. In fact, 'Loner' and 'Follower' did not load to any of the dependent variables. Further, 'Dumb' and 'Knowledgeable' were used to calculate perceived competence with notable loadings of -1.91 and 1.64 respectively. However, no differences in word-ratings were present. This may indicate a ceiling effect; children never think of the robot as being dumb. Or it may result from a priming effect; the robot stated to know about robotics and teach the children about this subject. Discarding 'Knowledgeable' from calculation of competence scores yields a significant difference in perceived competence between high/low competence samples, $F(1,97) = 4.41, p = 0.038$. This indicates that validating and improving the reliability of our measurement instrument are worthwhile.

Our findings on competence are in line with the suggestions from the literature: stable posture and frequent gestures evoke higher perceptions of competence. Our findings on warmth seemed in line with the literature: overall, low pitch and volume, body directed at audience, head tilt, open (semantic and syntactic) gestures evoked higher perceptions of warmth. However, perceived warmth declined in the HwLc condition compared to LwLc. This indicates that low-competence behaviours (unstable posture, infrequent gestures) reversed the effect of warmth behaviours. Alternatively, although opposed to the literature, the fiddling and gaze diversion only present in the LwLc condition could account for the effect. Additional head and hand movements might make the robot more lively and therefore perceived warmer. Our findings partially contradict the theory that warmth is evaluated before competence and therefore characters perceived as warm are likely to be found competent. Although competence scores were slightly higher for high-warmth robots than low-warmth samples, the difference was not significant, $F(1,97) = 2.35, p = 0.129$. The difference in perceived competence between high/low competence samples was larger in the high-warmth (1.04) condition than low-warmth (0.53), and the HwLc robot was perceived marginally less competent than the LwHc robot. This indicates that warmth expression enhances the competence effect rather than biases towards perceived high competence.

3.6. Study 2: Context Dependency

3.6.1. Method

To explore the effect of context (i.e., location, content, and usergroup) on children's perceptions of the robot, we conducted a 2 × 2 (*robot style* × *context*) follow-up study at a camp for children with type 1 diabetes mellitus (T1DM). Only two groups could be formed due to practical limitations, limiting us to two robot style configurations (HwHc and LwLc). The procedure and materials were similar to Study 1, except the lecture was about using MyPAL—an app developed to support learning about T1DM.

Participants A total of 72 children participated, 52 from the first study at schools (HwHc = 26, LwLc = 26), and 20 from the camp. We recruited 21 children with T1DM, of which 6 had experienced interacting with a NAO robot before during earlier studies. One participant was excluded from further analysis because the questionnaire was not understood and completed. This left 20 participants from camp (male = 12, female = 8), aged 8-11 ($M = 9.20, SD = 1.10$), divided in two groups (HwHc = 9, LwLc = 11).

3.6.2. RESULTS

We explored differences in children's perceptions of the robot between the two contexts (school and camp) and between robot styles (HwHc and LwLc). We performed MANOVA analysis to explore main and interaction effects between robot styles and contexts of use.

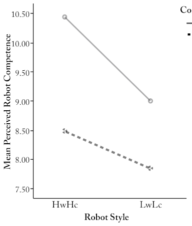
Using Pillai's trace, there was a significant main effect for context on perception, $V = 0.16, F(4,65) = 3.18, p = 0.019$. Separate univariate ANOVAs on the outcome variables revealed significant effects of context on perceived competence, $F(1,68) = 6.61, p = 0.012$, and affiliation, $F(1,68) = 6.99, p = 0.010$, and near a significant effect on warmth, $F(1,68) = 3.52, p = 0.065$. No significant multivariate effect for robot style on perception was found. However, univariate analysis revealed a near significant effect for robot style on competence, $F(1,68) = 3.03, p = 0.086$. No interaction effects were found for context × robot style, $V = 0.07, F(4,65) = 1.14, p = 0.345$.

This indicates that the context (location, users, and content) of the activity (presentation by a NAO robot) influenced perceived warmth, affiliation and competence. Overall, children with T1DM at camp perceived the robot as less warm ($M = 15.28$), affiliated ($M = 16.61$) and competent ($M = 8.14$) than children at school (warmth $M = 16.39$, affiliation $M = 18.11$, competence $M = 9.72$) (Fig 3.4a, 3.4b, 3.4c). Further, we were able to replicate the effect of robot style on perception scores with 20 new participants. Robots displaying high-competence and high-warmth behaviours were perceived more competent ($M = 9.94$) than low-competence and low-warmth robots ($M = 8.65$), independent of the context (Fig. 3.4a).

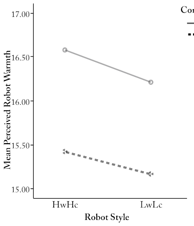
3.6.3. DISCUSSION

Children at camp perceived the robot less warm, affiliated and competent than children at school, independent of the robot style. Thus, changing the users, content and location of an activity can influence how children perceive an educational robot, meaning that results obtained in one context do not necessarily translate to another context.

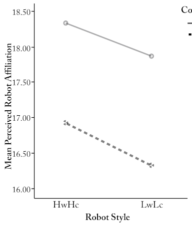
Lower perceived competence may indicate that the setting of Study 1 primed the children to think of the robot as competent because it would teach them. Although we expected the children at camp to see the robot as a friendly helper, and therefore perceive it as more warm and affiliated, this was not the case. Six of the children at camp had



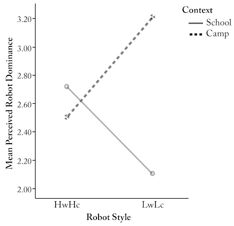
(a) competence
(range -11.64 to 12.91).



(b) warmth
(range -11.27 to 18.72).



(c) affiliation
(range -8.73 to 19.46).



(d) dominance
(range 0.00 to 12.00).

Figure 3.4: Study 2 results showing mean perception scores. Robot condition displayed on horizontal axis, contexts as separate lines.

previously experienced interacting with the NAO robot. Possibly, when novelty wears of, children are more rigorous in evaluating the robot. Or children may compare the robot's behaviour to previous experiences and find the robot less warm in the current activity compared to for example playing a game together. Alternatively, the framing as 'friendly helper' set expectations which the robot could not live up to. Further research is needed to provide solid conclusions on the influence of context on perception scores.

The lack of difference in perceptions between robot styles at camp may result from the small samples or too subtle manipulations. Although this was not the main purpose of the second study, it would be interesting to repeat the study and explore the effect of robot style using a larger sample size and more exaggerated behaviour manipulations.

3.7. CONCLUSION

We evaluated an educational robot displaying non-verbal behaviours expressing high or low warmth and competence with children at primary schools and camp, and showed that even subtle manipulations in robot behaviour influence children's perceptions of the robot's level of warmth and competence. The competence dimension in our model was successfully, but warmth manipulations had the intended effect only for high- competence robots. Moreover, context influenced children's perceptions; at school the robot was perceived warmer than at camp. Although this was a first attempt, and further research is needed replicating the study with enlarged behaviour manipulations and other social constructs, to our knowledge this is the first evidence of style in robot teaching and a first step towards adaptive style.

Table 3.3
MEAN WORD RATINGS AND PERCEPTION SCORES FOR BOTH CONTEXT, GROUPED BY ROBOT STYLE.

	Style	Bossy	Nagging	Clumsy	Friend	Popular	Playful	Follower	Loner	Angry	Honest	Fights	Knowledgeable	Boring	Nice	Listener	Confident	Educational	Helpless	Helpful	Dumb	Dominance	Affiliation	Warmth	Competence
School	HwHc	0.19	0.04	0.19	1.96	1.58	1.92	1.15	0.19	0.04	1.81	0.04	1.92	0.00	2.00	1.88	1.40	2.96	0.85	1.88	0.00	2.72	18.34	16.58	10.44
	HwLc	0.08	0.00	0.27	1.71	1.42	1.79	1.40	0.43	0.08	1.74	0.14	1.87	0.04	2.00	1.78	1.14	1.83	1.14	1.95	0.00	2.33	17.39	15.76	9.40
	LwHc	0.13	0.04	0.21	1.79	1.13	1.83	1.00	0.50	0.00	1.92	0.04	1.75	0.04	1.96	1.67	1.04	1.87	0.71	1.58	0.00	1.97	17.24	15.42	9.52
	LwLc	0.12	0.12	0.38	1.92	1.38	1.85	1.42	0.44	0.00	1.77	0.04	1.79	0.19	1.96	1.85	1.04	1.85	1.17	1.88	0.00	2.11	17.87	16.21	8.99
Camp	HwHc	0.33	0.00	0.56	1.89	1.44	1.67	0.67	0.67	0.11	2.00	0.11	1.89	0.22	2.00	1.44	0.78	1.89	1.22	1.78	0.11	2.51	16.94	15.42	8.49
	LwLc	0.18	0.36	0.55	1.91	1.56	1.91	0.55	0.36	0.36	1.64	0.18	1.73	0.09	2.00	1.73	1.36	1.55	1.00	1.82	0.09	3.22	16.34	15.17	7.86

REFERENCES

[1] R. Peters, J. Broekens, and M. A. Neerincx, *Robots educate in style: The effect of context and non-verbal behaviour on children's perceptions of warmth and competence,* in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (IEEE, 2017) pp. 449–455.

[2] T. Belpaeme, P. E. Baxter, R. Read, R. Wood, H. Cuayáhuitl, B. Kiefer, S. Racioppa, I. Kruijff-Korbayová, G. Athanasopoulos, V. Enescu, R. Looije, M. Neerincx, Y. Demiris, R. Ros-Espinoza, A. Beck, L. Cañamero, A. Hiolle, M. Lewis, I. Baroni, M. Nalin, P. Cosi, G. Paci, F. Tesser, G. Sommavilla, and R. Humbert, *Multimodal Child-Robot Interaction: Building Social Bonds,* Journal of Human-Robot Interaction **1**, 33 (2012).

[3] J. Xu, J. Broekens, K. Hindriks, and M. Neerincx, *Effects of Bodily Mood Expression of a Robotic Teacher on Students,* in *International Conference on Intelligent Robots and Systems* (IEEE, 2014) pp. 2614–2620.

[4] L. Hall, C. Hume, S. Tazzyman, A. Deshmukh, S. Janarthanam, and H. Hastie, *Map Reading with an Empathic Robot Tutor,* in *HRI* (IEEE, 2016) pp. 567–567.

[5] T. Fong, I. Nourbakhsh, and K. Dautenhahn, *A survey of socially interactive robots,* Robotics and Autonomous Systems **42**, 143 (2003).

[6] O. Mubin, C. J. Stevens, S. Shahid, A. A. Mahmud, and J.-j. Dong, *A review of the applicability of robots in education,* Technology for Education and Learing **1**, 209 (2013).

[7] S. Y. Okita, V. Ng-Thow-Hing, and R. Sarvadevabhatla, *Learning together: ASIMO developing an interactive learning partnership with children,* in *RO-MAN* (IEEE, 2009) pp. 1125–1130.

[8] M. Saerbeck, T. Schut, C. Bartneck, and M. D. Janse, *Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor,* in *CHI* (ACM, 2010) pp. 1613–1622.

[9] A. F. Grasha, *A Matter of Style: The Teacher as Expert, Formal Authority, Personal Model, Facilitator, and Delegator,* College Teaching **42**, 142 (1994).

[10] R. Looije, M. a. Neerincx, and V. D. Lange, *Children's responses and opinion on three bots that motivate , educate and play,* Physical Agents **2**, 13 (2008).

[11] C. Zaga, M. Lohse, K. P. Truong, and V. Evers, *The Effect of a Robot's Social Character on Children's Task Engagement: Peer Versus Tutor,* in *ICSR*, Vol. 9388 LNAI (Springer, Paris, 2015) pp. 704–713.

[12] M. Chollet, M. Ochs, and C. Pelachaud, *A multimodal corpus for the study of non-verbal behavior expressing interpersonal stances,* in *IVA Workshop Multimodal Corpora: Beyond Audio and Video* (2013).

3

3

[13] T.-h. D. Nguyen, E. Carstensdottir, N. Ngo, M. S. El-nasr, M. Gray, D. Isaacowitz, and D. Desteno, *Modeling Warmth and Competence in Virtual Characters,* in *Intelligent Virtual Agents*, Vol. 9238 (Springer, 2015) pp. 167–180.

[14] R. M. Felder and L. K. Silverman, *Learning and Teaching Styles in Engeneering Education,* Engr. Education **78**, 674 (1988).

[15] T. Leary, *Interpersonal diagnosis of personality,* American Journal of Physical Medicine & Rehabilitation **37**, 331 (1958).

[16] J. S. Wiggins, *An informal history of the interpersonal circumplex tradition,* Personality Assessment **66**, 217 (1996).

[17] R. Gifford, *The Role of Nonverbal Communication in Interpersonal Relations,* in *Handbook of interpersonal psychology: Theory, research, assessment, and therapeutic interventions* (Wiley, NY, 2011) Chap. 11, pp. 171–190.

[18] D. S. Berry and J. S. Hansen, *Personality, Nonverbal Behaviour, and Interaction Quality in Female Dyads,* PSPB **26**, 278 (2000).

[19] N. E. Dunbar and J. K. Burgoon, *Perceptions of power and interactional dominance in interpersonal relationships,* Social and Personal Relationships **22**, 207 (2005).

[20] L. L. Carli, S. LaFleur, and C. Loeber, *Nonverbal behavior, gender, and influence,* Personality and Social Psychology **68**, 1030 (1995).

[21] D. R. Carney, J. A. Hall, and L. S. LeBeau, *Beliefs about the nonverbal expression of social power,* Nonverbal Behavior **29**, 105 (2005).

[22] J. a. Hall, E. J. Coats, and L. S. LeBeau, *Nonverbal behavior and the vertical dimension of social relations: a meta-analysis,* Psychological bulletin **131**, 898 (2005).

[23] S. T. Fiske, A. J. C. Cuddy, and P. Glick, *Universal dimensions of social cognition: warmth and competence,* Trends in Cognitive Sciences **11**, 77 (2007).

[24] A. J. Cuddy, M. Kohut, and J. Neffinger, *Connect, Then Lead,* Harvard Business Review **91**, 54 (2013).

[25] S. Y. Okita, V. Ng-Thow-Hing, and R. K. Sarvadevabhatla, *Multimodal approach to affective human-robot interaction design with children,* TiiS **1**, 1 (2011).

[26] I. Leite, G. Castellano, A. Pereira, C. Martinho, and A. Paiva, *Empathic Robots for Long-term Interaction: Evaluating Social Presence, Engagement and Perceived Support in Children,* International Journal of Social Robotics **6**, 329 (2014).

[27] W. Johal, S. Pesty, and G. Calvary, *Towards companion robots behaving with style,* in *Ro-MAN*, Vol. 23 (IEEE, Endinbrugh, 2014) pp. 1063–1068.

[28] J. Kennedy, P. Baxter, and T. Belpaeme, *The Robot Who Tried Too Hard: Social Behaviour of a Robot Tutor Can Negatively Affect Child Learning,* in *HRI* (ACM, Portland, 2015) pp. 67–74.

[29] Y. H. Lin, C. Y. Liu, H. W. Lee, S. L. Huang, and T. Y. Li, *Evaluating emotive character animations created with procedural animation,* in *IVA*, Vol. 5773 LNAI (Springer, 2009) pp. 308–315.

[30] M. Neff, Y. Wang, R. Abbott, and M. Walker, *Evaluating the effect of gesture and language on personality perception in conversational agents,* in *IVA*, Vol. 6356 LNAI (Springer, 2010) pp. 222–235.

[31] A. Arya, L. N. Jefferies, J. T. Enns, and S. DiPaola, *Facial actions as visual cues for personality,* Computer Animation and Virtual Worlds **17**, 371 (2006).

[32] A. Cafaro, N. Glas, and C. Pelachaud, *The Effects of Interrupting Behavior on Interpersonal Attitude and Engagement in Dyadic Interactions,* in *AAMAS* (ACM, 2016) pp. 911–920.

[33] K. Bergmann, F. Eyssel, and S. Kopp, *A second chance to make a first impression? how appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time,* (Springer, 2012) pp. 126–138.

[34] R. Ros, M. Nalin, R. Wood, P. Baxter, R. Looije, Y. Demiris, T. Belpaeme, A. Giusti, and C. Pozzi, *Child-robot interaction in the wild: advice to the aspiring experimenter,* in *ICMI* (ACM, 2011) pp. 335–342.

4

ROBOTS EXPRESSING DOMINANCE: EFFECTS OF BEHAVIOURS AND MODULATION

A major challenge in human-robot interaction and collaboration is the synthesis of non-verbal behaviour for the expression of social signals. Appropriate perception and expression of dominance (verticality) in non-verbal behaviour is essential for social interaction. In this paper, we present our work on algorithmic modulation of robot bodily movement to express varying degrees of dominance. We developed a parameter-based model for head tilt and body expansiveness. This model was applied to a variety of behaviours. These behaviours were evaluated by human observers in two different studies with respectively static pictures of key postures (N=772) and real-time gestures (N=31). Overall, specific behaviours proved to communicate different levels of dominance. Further, modulation of body expansiveness and head tilt robustly influenced perceived dominance independent of specific behaviours and observer viewing height and angle. The modulation did not influence perceived valence, but it did influence perceived arousal. Our study shows that dominance can be reliably expressed by both selection of specific behaviours and modulation of behaviours.

This chapter has been published in proceedings of the 8th IEEE International Conference on Affective Computing and Intelligent Interaction (ACII), (2019) [1].

4.1. Introduction

Robots and virtual agents increasingly fulfil functions that require social interaction and communication with people who are not trained to interact with a robot or avatar. In inter-human communication, the interaction consists of both verbal and nonverbal behaviour. Nonverbal behaviour conveys information about the relationship and about likes and dislikes [2]. Based on observable behaviours (e.g., facial expression, proximity) people form impressions of the others and develop expectations [3]. For artificial agents (both virtual and robotic), to engage in meaningful interactions with humans, the importance of social intelligence is widely acknowledged [4, 5].

Research on emotional expressions for virtual agents is abundant (e.g., [6–8]). A common approach for the design of emotional expressions is to develop specific behaviours based on human emotion expression in a particular modality (e.g., posture, gesture, gaze, facial expressions) [4]. Facial expressions are considered important cues for emotions [9], but variations in appearance and functionality of robots as well as line of sight of the user poses limitations on the usefulness of this modality. Emotional expressions are possible with body posture as well; Cohen [10] reported high recognition rates for the emotional expressions of a “facial robot” (iCAT) and a “bodily robot” (NAO). Further, expressing affect *while the robot is busy with its task* (sometimes referred to as mood expression [11]) can be useful in interaction scenarios where the robot is observed by the user while performing a task. Therefore, alternative models for affect expression are needed.

Alternative models for different types of robots have been investigated. For example, Lin [12] applied transformations to walking motions of a virtual avatar and showed that discrete emotions can be expressed by manipulating motion stiffness, pace, and expansiveness. Beck [13] attempted to create an ‘affect space’ able to express emotions on the two dimensional circumplex model of affect [14] by blending key poses of discrete emotions. Xu [15] defined behaviour-specific design patterns of key poses and interpolation parameters targeting specific joints, which showed able to express valence and arousal [16]. We focus on three challenges in affect expression. First, most models and methods require a considerable amount of work in the form of key poses or interpolation patterns for each behaviour before being usable. Second, it is important to assess perceived affect by the users on the affective dimension that is intended to be manipulated, in other words, we need a validated set of “stimuli”. Third, most models of affect expression focus on valence and arousal, or, discrete emotions.

4.2. Dominance

Humans (unconsciously) use social signals to inform others about their affective stance or attitude; based on observations we evaluate someone as, among other things, warm or cold, friendly or hostile, and dominant or submissive (e.g., [17, 18]). Body language serves various communicative functions, amongst which affect displays [19]. It is commonly accepted that power or status, i.e., dominance, is an important factor in interpersonal relations and communications. Dominance is defined as “power and influence over others” [20], but research fields adopt specific notions.

Dominance is a factor in the interpersonal circumplex —or Leary’s Rose— a two dimensional model for interaction stance defined by: dominance–submissiveness and affiliation–hostility [21, 22]. Dominance, in this view, is an interpersonal factor. Further, the complementary mechanism enables strategic use of dominance display as a tool to form interactions by, for example, teachers [23] and police officers [24]. In social signal processing the dominance dimension is referred to as *verticality* [25].

Dominance is also a dimension of affect [2, 26]. Suggested is that affect (including discrete emotions, moods, and attitudes) can be placed in a multidimensional pleasure, arousal and dominance (PAD) model [27]. Emotions are directed responses to internal and external events [28], and can be either self-directed or social. Thus, a dominance relation can exist between a person and the environment or a stimulus as well. Dominance reflects the amount of influence, power and overwhelmingness of a stimulus, be that another human or a magnificent forest.

Both notions of dominance, dominance-as-emotion-dimension and verticality-as-social-stance, relate to similar concepts such as power, control and influence. They are both relational. The first is defined as an affective dimension, and thus has meaning in the context of affect communication. The second is defined as a interpersonal dimension and has meaning in the context of relations. However, both share similar behavioural cues for expression. Therefore, for the purpose of expressing dominance/verticality through the body language of a robot, we propose to consider these dimensions to be equal.

4.2.1. Dominance Expression in Humans

Body shape, or posture, has been associated with dominance display. A dominant posture was stereotyped as forward [29], open [29–31], expansive [31], upright [29], and oriented towards the other [29]. Various studies reported a positive influence of body expansiveness and/or openness on dominance expression (e.g., [31, 31, 32]). Carney [31] defined openness as “keeping limbs open or closed”, and expansiveness as “taking up more space or less space”, but other authors provided less clear definitions of the terms or used them intertwined. In the remainder of this paper *expansiveness* is used to indicate both similar features.

However, the relation between posture and dominance is arguably moderated by other factors such as gender or culture [33], and supposedly previous work was inconclusive, based on limited data, or highly context dependent [30]. Moreover, most studies used pre-recorded exemplary images and compared between distinctive postures. Meaning that the results are only valid for the specific posture under evaluation.

Gestures, co-verbal motions of arms and/or head, are also an important cue for dominance expression. Though, frequency and types are discussed rather than performance. Exception is Kipp [34] who found that an open hand shape is associated with submissiveness while a pointing hand shape is associated with dominance. However, this may result from the sign function of both gestures (pointing at someone vs holding your hands to receive something), and as such does not have to do with modulation of openness.

Pointing was repeatedly associated with dominance expression [35–37]. Pointing is believed a signal of aggression, anger, or arrogance —emotions with a high dominance value in the PAD model [36]. Other types of gestures were associated with dominance as well for there aggressive (e.g., [36]) or attention gaining (e.g., [38]) feature. Head position has received specific attention, multiple studies unanimously concluded that upward head tilt is associated with dominance [29, 39, 40].

To summarise, key poses, head tilt, and body expansiveness are widely acknowledged bodily cues for expressions of dominance/submissiveness. However, concrete models for dominance expression are unavailable.

4.2.2. Dominance Expression in Conversational Agents

There is growing interest in synthesis of expressive behaviour for virtual agents and robots. Often however, studies are application specific and evaluate pre-designed behaviours, often mirroring human behaviour, rather than parameter-based models (e.g., [6, 41, 42]).

Nonetheless, models for dominance expression were subject of research before. Studies may focus on behaviour selection (e.g., [43]), but this line of work gives no insight on the features of the behaviours that convey the signal. It may indicate which behaviours are appropriate, but not how these behaviours should be performed. Study on the 'manner' of performance was done by, for example, Lance [44]. Reportedly, gaze behaviours expressing dominance are performed with head tilted up and a higher body compared to submission. This study, like others, used specific (gaze) behaviours, the effect for other behaviours remains unknown. Ravenet [45] combined behaviour (type) selection and gesture performance parameters (i.e., power, amplitude). They obtained, from co-design, parameters for dominance display similar to studies in human behaviour which served as input to model avatar behaviour. However, validation of this model was not presented. A recent study did perform an exhaustive analysis of valence and arousal perception of gestures and performance parameters and found interesting interaction effects between gestures and performance [46].

Studies that evaluate parameter-based models to control dominance expression often modify multiple parameters in parallel (e.g., [47, 48]). This approach makes it impossible to derive the effect of unique parameters, and thereby hard to reuse. Evaluation of unique parameters is laborious and results seem context dependent. For example, while Kim [49] reported a positive effect on dominance expression for direction and speed, Saerbeck [50] reported a positive effect for motion direction and a negative effect for motion speed. So far, dominance was found positively correlated with forward sagittal motion direction [49, 50], forward gaze direction [41], head position tilted upward [41, 44], and gesture direction [49, 50]. Results for gesture speed were inconclusive [49, 50]. Because it is currently unknown how dominance perception can be manipulated using specific behavioural parameters, in this work we focus on a detailed study to identify the effect of body expansiveness (head tilt as a part of expansiveness) on perceived dominance. We try to isolate the effect of expansiveness but test this on a range of behaviours and user viewing angles to ensure that the effects we find are generic. We also identify the effect of the behaviours themselves on perceived dominance, as a possible confound but also as a means to express dominance.

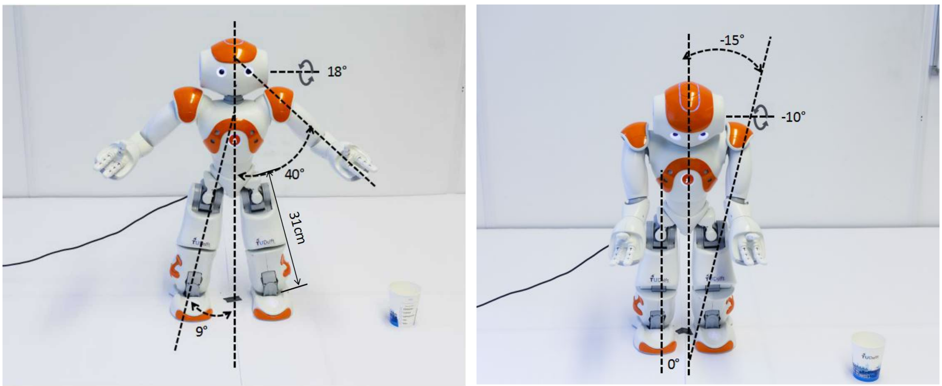
4.3. Robot Dominance Model

Based on behavioural cues for dominance expression in human-human interaction (see Section 4.2), we developed a parametric model of body expansiveness for dominance expression in a humanoid robot. We selected five parameters manipulating body expansiveness: vertical head angle, horizontal shoulder angle, horizontal and sagittal hip angle, and vertical leg stretch (Fig. 4.1).

The modulations are applied to existing (neutral) NAO behaviours. We define a factor f as the dominance level which is positively correlated with the dominance expression, and ranging $[-1.00, 1.00]$, where 0 is the neutral stance, and -1 and 1 represent the most submissive and the most dominant stance respectively. Based on the f value the movement trajectory of affected joints is adapted. For head and arm movement a linear modulation between the neutral and limit position (see Table 4.1) is executed on joints `HeadPitch`, `LShoulderRoll`, and `RSholderRoll`. A time adjustment is applied to maintain a consistent speed. The legs are adapted accordingly following one of three standing patterns (i.e., neutral, dominant, submissive). Although the following implementation is NAO specific, the principles explained therein can be used on any humanoid with a similar structure.

4.3.1. IMPLEMENTATION

We implemented the parameter-based model for dominance expression in an existing framework for NAO behaviour management. To execute a behaviour on the NAO robot, it expects two arrays with the path and execution time. The path information for all original (neutral) behaviours is stored in an XML file. To manipulate dominance expression we apply joint modulations relative to the neutral path and based on the dominance level. The affected joints and their adjustment patterns are given in Table 4.1.



(a) dominance

(b) submissiveness

Figure 4.1: Example manipulations for the most dominant and most submissive stance as implemented on the NAO robot in a standing position. Maximum dominance: head tilt 18°, arm spread 40°, leg spread 9°, and leg stretch 30cm (Fig. 4.1a). Maximum submissiveness: head tilt -10°, arms enclosing 10°, leg angle 0°, and leg stretch 26.5cm (Fig. 4.1b).

Table 4.1
MODULATED JOINTS, AND JOINT ANGLE VALUES FOR MOST DOMINANT, NEUTRAL AND MOST SUBMISSIVE STANCE. (LEFT JOINT VALUES ARE GIVEN, RIGHT JOINT VALUES WERE REFLECTED.)

Parameter	Joint	Dom	Neutral	Sub
Head tilt	`HeadPitch`	0.51	0.00	0.67
Shoulder angle	`LShoulderRoll`	1.33	0.00	-0.31
Leg angle	`LHipYawPitch`	-0.17	0.00	0.00
	`LHipRoll`	0.09	0.00	0.00
Leg stretch	`LHipPitch`	0.13	0.00	-0.44
	`LKneePitch`	-0.08	0.00	0.69
Stability correction	`LAnklePitch`	0.08	0.00	-0.35
	`LAnkleRoll`	-0.13	0.00	0.00

4

Motion Trajectory For a specific behaviour, multiple joints are moving in parallel. The path of joint i is described as:

$$\begin{Bmatrix} x_i = & (x_{i0}, x_{i1}, \cdots, x_{in_i}) \\ t_i = & (t_{i0}, t_{i1}, \cdots, t_{in_i}) \end{Bmatrix} \tag{4.1}$$

where x_{ij} is the trajectory value of joint i at t_{ij} time, and n is the maximum value of n for all m joints. The path for the entire behaviour can be described as two $m \times n$ matrices: $X_{m\times n}$ and $T_{m\times n}$.

Parameter Insertion For each k modulated joints, an array will be inserted in the path matrix. If we define the inserted arrays as $N_{k\times n}$ and $L_{k\times n}$, the path matrix will be:

$$X'_{(m+k)\times n} = \begin{pmatrix} X_{m\times n} \\ Y_{k\times n} \end{pmatrix}, T'_{(m+k)\times n} = \begin{pmatrix} T_{m\times n} \\ L_{k\times n} \end{pmatrix} \tag{4.2}$$

Non modulated joint movement remains the point-to-point path specified in the XML.

Head and Arm Movement Linear modulation is applied to the path trajectories of selected joints (i.e., `HeadPitch`, `LShoulderRoll`, and `RShoulderRoll`) as follows:

$$J(f) = \begin{cases} x_{Neutral} + (x_{Max} - x_{Neutral}) \times f & , f > 0 \\ x_{Neutral} - (x_{Neutral} - x_{Min}) \times f & , f < 0 \end{cases} \tag{4.3}$$

x_{Max} and x_{Min} are the reference values for the limit positions (i.e., most dominant and submissive), corresponding with the joint angles listed in Table 4.1. The dominance factor f is the relative proportion of modulation applied between the neutral and limit position.

We define the three modulated joints as $x_0 \sim x_2$, from (4.3), their new trajectory is derived as:

$$x'_i = \left(J(x_{i0}, f), J(x_{i0}, f), \cdots, J(x_{in}, f)\right) \tag{4.4}$$

Let

$$X_1 = \left(x'_0, x'_1, x'_2\right)^T \tag{4.5}$$

X_1 is the first three rows of the trajectory matrix.

Time Adjustment A time-stamp adjustment is applied to maintain the velocity and acceleration over shortened or prolonged trajectories. For a positive factor f the time is increased. To increase time by the same portion as the trajectory the percentage increase for each time interval should be the same as the percentage increase of the trajectory. From (4.3) we can get:

$$\begin{aligned} x'_{ij} &= x_{ij} + \left(x_{max} - x_{ij}\right) \times f \\ x'_{i(j-1)} &= x_{i(j-1)} + \left(x_{max} - x_{i(j-1)}\right) \times f \end{aligned} \tag{4.6}$$

Then we get the angle displacement:

$$x'_{ij} - x'_{i(j-1)} = (x_{ij} - x_{i(j-1)}) \times (1-f) \tag{4.7}$$

The change in proportion $\frac{x'_{ij}-x'_{i(j-1)}}{x_{ij}-x_{i(j-1)}}$ will be $(1-f)$. Therefore, the new j time for joint i is calculated by:

$$t'_{ij} = t'_{i(j-1)} + \left(t_{ij} - t_{i(j-i)}\right) \times \left(1.00 - f\right) \tag{4.8}$$

Leg Movement We created three patterns of standing poses that vary in expansiveness and extension instead of continuous path trajectories due to balance constraints. The patterns were created manually, and relate to the dominance factor as follows:

$$SPattern(f) = \begin{cases} Pattern_{sub} & , -1.00 \le f < -0.33 \\ Pattern_{neu} & , -0.33 \le f \le 0.33 \\ Pattern_{dom} & , 0.33 < f \le 1.00 \end{cases} \tag{4.9}$$

Upon change of dominant factor f above the specified thresholds, a transition between standing patterns takes place in parallel with behaviour execution. Assume the 3rd to 12th rows in X' and T' are the path for the legs joints, then these values will be replaced by those of the new leg pattern. These new values are defined as two $10 \times n$ matrices X_2 and T_2. The final path for this behaviour is:

$$X'' = \begin{pmatrix} X_1 \\ X_2 \\ X_r \end{pmatrix}, T'' = \begin{pmatrix} T_1 \\ T_2 \\ T_r \end{pmatrix} \tag{4.10}$$

X_r and T_r are paths of unchanged joints.

4.4. PILOT: MODEL VALIDATION

In a first study, we validated whether manipulation of posture expansiveness influenced how dominant people perceive the robot. Further, we explored the covariates vertical viewing angle and horizontal viewing angle, and investigated posture dependency.

4.4.1. METHOD

We conducted a 2 (expansiveness) x 2 (horizontal angle) x 2 (vertical angle) between-subject, factorial posture perception study to evaluate the effect of body expansiveness on perceived dominance expression and explore covariates view angle.

Stimuli The expansiveness modulation described in Section 4.3 was applied to 11 distinctive postures performed by a NAO robot. Pictures were taken from a 0° and 30° horizontal angle and with the robot standing on ground level and a 110cm height table.

Measurement Perceived dominance of the stimulus was measured on a 5-point Likert scale (dominant to submissive).

Procedure An on-line survey was set-up at Amazon Mechanical Turk. A Human Intelligence Task (HIT) consisted of 11 subsequent pictures presented in random order and depicting postures within one condition. A HIT could be started anytime. Each participant could complete the HIT only once. First, demographic data was collected and the task was explained. Then, two trail questions were presented, showing iconic human expressions of dominance and submissiveness. Finally, participants rated the robot images one by one.

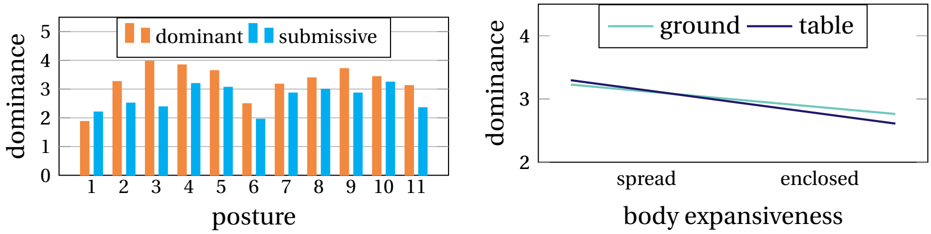
Participants A total of 835 Mechanical Turk workers completed the HIT and were compensated monetarily. Of these, 45 participants failed the trial question and were excluded from further analysis. Another 18 participants were excluded due to low credibility based on a reported age of 1 or 113. The remaining 772 participants were self-identified mostly Americans (n=558) or Indian (n=110), aged between 20 and 83 (Mean=35.42, Std=10.69), and 60% male. Participants where fairly evenly distributed among conditions with a minimum of 89 and maximum of 109 participants per group.

4.4.2. RESULTS

We performed a MANOVA with repeated measures, with within-subject factors the 11 postures, and between-subject factors horizontal view angle, vertical view angle, and body expansiveness.

The tests of between-subject effects showed a significant main effect for body expansiveness on perceived dominance, $F = 262.52, p < 0.001$. This effect was consistent between postures, except for posture 1 (Fig. 4.2a). Both horizontal and vertical view angle did not significantly affect perceived dominance (respectively $F = 1.56, p = 0.212$, $F = 0.38, p = 0.535$). But, there was a significant, positive, interaction effect for vertical view angle and expansiveness, $F = 9.79, p = 0.002$ (Fig. 4.2b). In other words, only body expansiveness influenced perceived dominance and did so in the expected direction, this effect was increased when the vertical view angle was decreased (i.e., robot on a table).

The within-subject tests showed a significant main effect of posture on perceived dominance, $F = 182.52, p < 0.001$. Further, there were small though significant interaction effects between posture and each of the factors horizontal angle, vertical angle, and expansiveness. Most prominently posture and expansiveness, $F = 45.23, p < 0.001$.



(a) Mean perceived dominance for both submissive and dominant stimuli, given per posture.

(b) Mean perceived dominance for both ground and table placed stimuli, given per body expansiveness.

Figure 4.2: Mean perceived level of dominance expressed by the robot (range 0–5).

4.4.3. DISCUSSION

First, the effect of manipulation of body expansiveness on perceived dominance shows that expansiveness is an important factor for dominance expression. Although there was an interaction effect with view height this did not hinder perception as intended. Indicating that body expansiveness manipulation can be used to influence perceived dominance, and that this is robust against variations in view angle. Further, different postures have different associated levels of dominance. This is of no surprise given that some postures by nature are more expansive than others. Nonetheless, for all but one (sitting) postures the body expansiveness manipulation influenced dominance perception in the desired direction. This indicates that body expansiveness modulation can be used to influence perceived dominance independent of specific behaviours.

To summarise, we have identified two methods controlling dominance expression by a robot: posture selection and body expansiveness manipulation. The latter by modulation of the parameters head tilt, arm angle, leg angle, and leg stretch.

4.5. STUDY 2: SYNTHESISED GESTURES

In a follow-up study, we evaluated our model of dominance expression (see Section 4.3) in real-time gesture performance. We evaluated whether increased body expansiveness applied to motions resulted in higher observer perception of the robot's level of dominance. Further, we explored the effect of body expansiveness on valence and arousal.

4.5.1. METHOD

To evaluate the effect of expansiveness modulation applied to motions on observers' perception of the robot's affect expression, we set up an experiment with between-subject variable body expansiveness, and within-subject variable gesture.

Stimuli The proposed modulation can be applied to any robot behaviour. For the purpose of this experiment we selected 10 distinctive behaviours and applied the maximum dominant and submissive modulations. The behaviours were created in Choreograph and designed to express a 'neutral' stance. The resulting submissive, neutral, or dominant versions were shown on a NAO placed in front on a table.

Measurement Perceived dominance was measured on a 9-point Likert scale. We used the Self Assessment Manikin (SAM) [26] because it is a widely acknowledged, validated, instrument measuring affective responses to a wide variety of stimuli. We include all three items (i.e., dominance, valence, arousal) to control for correlations with these factors.

Procedure Participants participated individually and were seated 1.5 meters from the robot. Each participant was assigned to one condition (i.e., neutral, dominant, submissive), and presented the 10 gestures, modulated accordingly, one by one in randomised order. After each gesture the participant completed the SAM questionnaire. Gestures could be viewed repeatedly upon request. A researcher was present in the room to control the robot.

Participants A total of 31 participants were recruited at the University premises, all students or staff. Participants were aged 23–62 (Mean=32.22, std=9.88), mostly male (n=17), and predominantly of Chinese (n=10) or Dutch (n=12) nationality. Eight participants did not provide their age, of these six withheld their nationality, and four their gender as well. Participants were equally balanced between conditions.

4.5.2. Results

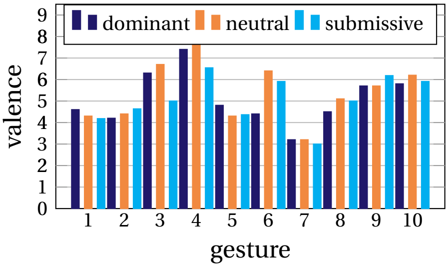
We performed a MANOVA (within-subject factors the 10 gestures and between-subject factor expansiveness) and compared outcomes for perceived dominance, valence and arousal.

Using Pillai's trace, the multivariate test showed a moderate tendency of body expansiveness to affect overall perception, $V = 0.35, F(6,54) = 1.90, p = 0.098$. No interaction effect between gesture and body expansiveness was found. However, separate univariate ANOVAs on the outcome variables revealed significant positive body expansiveness effects on perceived dominance, $F(2,28) = 4.41, p = 0.022$; and arousal, $F(2,28) = 5.10, p = 0.013$. In other words, participants perceived the robot displaying spreading gestures as more dominant ($\mu = 5.6$) and aroused ($\mu = 5.93$) than a robot showing more enclosed gestures (dominance, $\mu = 4.58$; arousal, $\mu = 4.76$), and that these differences were independent of specific gestures.

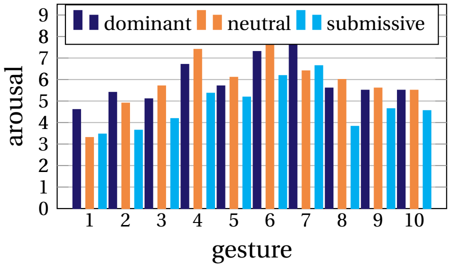
Tests of within-subject effects revealed a small but significant overall effect of gesture on perception, $V = 0.83, F(27,756) = 10.78, p < 0.001$. Univariate tests on the outcome variables revealed significant effects of gesture for all three factors (dominance, $F = 16.56, p < 0.001$; valence, $F = 13.92, p < 0.001$; arousal, $F = 12.78, p < 0.001$). For none of the factors, an interaction effect between gesture and expansiveness was found. In other words, different gestures elicited different perceptions, but these differences were consistent over body expansiveness conditions (Fig. 4.3).

4.5.3. Discussion

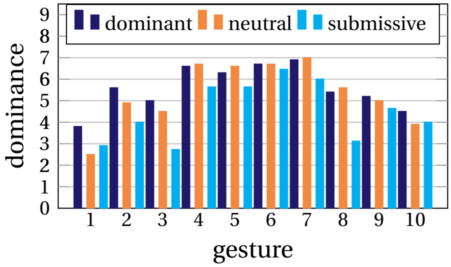
This study shows that, also in motions, body expansiveness influences dominance perception. The effect of body expansiveness is independent of a specific gesture, and correlated with arousal. This might be explained by more dominant behaviours being more expressive, and expressiveness being a factor for arousal display (e.g., [51]). Further, one



(a) Mean perceived valence (range 0–9)



(b) Mean perceived arousal (range 0–9)



(c) Mean perceived dominance (range 0–9)

Figure 4.3: Estimated means for user perception of the robot, numbers 1–10 indicate the individual gestures, coloured bars depict the body expansiveness conditions.

of the parameters, head tilt, was associated with arousal display (e.g., [41, 52]). Arousal was also found to correlate with valence [16]. However, in our study, body expansiveness did not affect perceived valence (Fig. 4.3a). This means that body expansiveness could be used to selectively manipulate dominance display without influencing perceived valence (pleasure).

As with postures, the body expansiveness effect is consistent over multiple gestures, however, different gestures do convey different levels of dominance. Thus, parameter-based dominance control is limited to relative differences within a behaviour's affective tendency. Considerate behaviour selection is necessary to convey a consistent dominance display over a prolonged interaction sequence. The affective tendencies of behaviours (i.e., postures and gestures) can be used to our advantage. For example, adjusting the frequency of specific gestures to control dominance [47], or selecting types of gestures to express certain roles (e.g., [53, 54]).

We have shown the validity of body expansiveness modulation for dominance expression in both postures and gestures. We show that with a limited set of parameters we can express various degrees of dominance. The joint angles are specific to NAO, but the body expansiveness modulation can be applied to any robot or embodied agent, making the model applicable to other systems and scenario's as well. Further, we discovered distinctive affective patterns in individual gestures. However, extensive evaluation of this effect in interaction is required to support expression of certain roles in interactive scenario's.

4.6. CONCLUSION

In conclusion, we showed that dominance perception of robot gestures and postures can be controlled by behaviour selection and parameter-based body expansiveness manipulation. We found that specific postures and gestures have a natural tendency towards being perceived as more or less dominant. Further, the manipulation effect was consistent for a variety of behaviours except a sitting pose. A clear view on the robot may increase the effect. Manipulation was based on inter-human interaction cues for dominance display: vertical head angle, horizontal shoulder angle, horizontal and sagittal hip angle, and vertical leg stretch. Our results are limited by the number of behaviours evaluated and application on a NAO robot and should be replicated with other robot types.

REFERENCES

[1] R. Peters, J. Broekens, K. Li, and M. A. Neerincx, *Robots expressing dominance: Effects of behaviours and modulation,* in *8th International Conference on Affective Computing and Intelligent Interaction (ACII)* (IEEE, Cambridge, UK, 2019) pp. 1–7.

[2] A. Mehrabian, *Nonverbal communication,* in *symp. on motivation* (University of Nebraska Press, Nebraska, 1971) pp. 107–161.

[3] G. B. Moskowitz, *Social cognition: Understanding self and others* (Guilford Press, New York, 2005).

[4] T. Fong, I. Nourbakhsh, and K. Dautenhahn, *A survey of socially interactive robots,* Robot. and autonomous syst. **42**, 143 (2003).

[5] A. Vinciarelli, M. Pantic, D. Heylen, C. Pelachaud, I. Poggi, F. D'Errico, and M. Schroeder, *Bridging the Gap Between Social Animal and Unsocial Machine: A Survey of Social Signal Processing,* Affective Computing **3**, 69 (2011).

[6] G. Castellano, M. Mancini, C. Peters, and P. W. McOwan, *Expressive copying behavior for social agents: A perceptual analysis,* Trans. on Syst., Man, and Cybern. **42**, 776 (2012).

[7] C. Clavel, J. Plessier, J.-C. Martin, L. Ach, and B. Morel, *Combining facial and postural expressions of emotions in a virtual character,* in *Int. Workshop on Intell. Virtual Agents* (Springer, 2009) pp. 287–300.

[8] M. Destephe, A. Henning, M. Zecca, K. Hashimoto, and A. Takanishi, *Perception of emotion and emotional intensity in humanoid robots gait,* in *Proc. 2013 Int. Conf. on Robot. and Biomimetics* (IEEE, 2013) pp. 1276–1281.

[9] D. N. Kurien, *Body language: Silent communicator at the workplace,* The IUP Journal of Soft Skills **4**, 29 (2010).

[10] I. Cohen, R. Looije, and M. A. Neerincx, *Child's perception of robot's emotions: effects of platform, context and experience,* Int. J. of Social Robot. **6**, 507 (2014).

[11] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx, *Mood expression through parameterized functional behavior of robots,* in *Proc. Int. Symp. on Robot and Human Interactive Commun.* (IEEE, Gyeongiu, Korea, 2013) pp. 533–540.

[12] Y.-H. Lin, C.-Y. Liu, H.-W. Lee, S.-L. Huang, and T.-Y. Li, *Evaluating emotive character animations created with procedural animation,* in *Int. Workshop on Intell. Virtual Agents* (Springer, Amsterdam, The Netherlands, 2009) pp. 308–315.

[13] A. Beck, A. Hiolle, A. Mazel, and L. Cañamero, *Interpretation of emotional body language displayed by robots,* in *3rd Int. Workshop on Affect. interaction in natural environments* (2010) pp. 37–42.

[14] J. A. Russell, *A circumplex model of affect.* J. of personality and social psychology **39**, 1161 (1980).

4

[15] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx, *The relative importance and interrelations between behavior parameters for robots' mood expression,* in *Proc. Conf. on Affect. Comput. and Intell. Interaction* (IEEE, 2013) pp. 558–563.

[16] J. Xu, J. Broekens, K. Hindriks, and M. Neerincx, *Bodily mood expression: Recognize moods from functional behaviors of humanoid robots,* in *Proc. Int. Conf. on social robot.* (Springer, Bristol, UK, 2013) pp. 511–520.

[17] S. T. Fiske, A. J. Cuddy, and P. Glick, *Universal dimensions of social cognition: Warmth and competence,* Trends in cogn. sci. **11**, 77 (2007).

[18] S. L. Ellyson and J. F. Dovidio, *Power, dominance, and nonverbal behavior: Basic concepts and issues,* in *Power, dominance, and nonverbal behavior* (Springer, 1985) pp. 1–27.

[19] P. Ekman, *Emotional and conversational nonverbal signals,* in *Language, knowledge, and representation* (Springer, 2004) pp. 39–50.

[20] Oxford dictionaries, *Definition of dominance in English,* https://en.oxforddictionaries.com/definition/dominance, accessed: 2019-02-19.

[21] J. S. Wiggins, *An informal history of the interpersonal circumplex tradition,* J. of personality assessment **66**, 217 (1996).

[22] T. Leary, *Interpersonal diagnosis of personality,* Amer. J. of Physical Med. & Rehabilitation **37**, 331 (1958).

[23] S. O. Ebbens and S. Ettekoven, *Effectief leren: basisboek* (Noordhoff Uitgevers, 2016).

[24] A. Amelsvoort, I. Rispens, and H. Grolman, *Handleiding verhoor* (Stapel & De Koning, 2010).

[25] M. Schmid Mast and J. A. Hall, *Social signal processing,* (Cambridge University Press, Cambridge, 2017) Chap. The vertical dimension of social signaling, pp. 34–45.

[26] M. M. Bradley and P. J. Lang, *Measuring emotion: the self-assessment manikin and the semantic differential,* J. of behavior therapy and experimental psychiatry **25**, 49 (1994).

[27] J. A. Russell and A. Mehrabian, *Evidence for a three-factor theory of emotions,* J. of res. in Personality **11**, 273 (1977).

[28] D. Schacter, D. Gilbert, D. Wegner, and B. M. Hood, *Psychology: European Edition* (Macmillan Int. Higher Education, 2011).

[29] D. R. Carney, J. A. Hall, and L. S. LeBeau, *Beliefs about the nonverbal expression of social power,* J. of Nonverbal Behavior **29**, 105 (2005).

[30] J. A. Hall, E. J. Coats, and L. S. LeBeau, *Nonverbal behavior and the vertical dimension of social relations: a meta-analysis.* Psychological bulletin **131**, 898 (2005).

[31] D. R. Carney, A. J. Cuddy, and A. J. Yap, *Power posing: Brief nonverbal displays affect neuroendocrine levels and risk tolerance,* Psychological sci. **21**, 1363 (2010).

[32] A. J. Cuddy, C. A. Wilmuth, A. J. Yap, and D. R. Carney, *Preparatory power posing affects nonverbal presence and job interview performance.* J. of Appl. Psychology **100**, 1286 (2015).

[33] O. Bialobrzeska and M. Parzuchowski, *Size or openness: expansive but closed body posture increases submissive behavior,* Polish Psychological Bulletin **47**, 186 (2016).

[34] M. Kipp and J.-C. Martin, *Gesture and emotion: Can basic gestural form features discriminate emotions?* in *Int. Conf. on affect. comput. and intell. interaction* (IEEE, Amsterdam, Netherlands, 2009) pp. 1–8.

[35] M. Argyle, *Bodily communication* (Routledge, 2013).

[36] nicolasfradet, *13 revealing body language hand gestures,* https://nicolasfradet.com/hand-body-language, accessed: 2017-01-17.

[37] scienceofpeople, *20 hand gestures you should be using,* http://www.scienceofpeople.com/2015/08/how-to-speak-with-your-hands/, accessed: 2017-01-17.

[38] T. Leonard and F. Cummins, *The temporal relation between beat gestures and speech,* Language and Cogn. Processes **26**, 1457 (2011).

[39] A. Mignault and A. Chaudhuri, *The many faces of a neutral face: Head tilt and perception of dominance and emotion,* J. of nonverbal behavior **27**, 111 (2003).

[40] N. O. Rule, R. B. Adams Jr, N. Ambady, and J. B. Freeman, *Perceptions of dominance following glimpses of faces and bodies,* Perception **41**, 687 (2012).

[41] A. Beck, L. Cañamero, and K. A. Bard, *Towards an affect space for robots to display emotional body language,* in *Proc. 19th Int. symp. in robot and human interactive commun.* (IEEE, Viareggio, Italy, 2010) pp. 464–469.

[42] C. Breazeal, *Emotion and sociable humanoid robots,* Int. j. of human-comput. stud. **59**, 119 (2003).

[43] M. Chollet, M. Ochs, and C. Pelachaud, *From non-verbal signals sequences mining to bayesian networks for interpersonal attitude expression,* in *Proc. Int. Conf. on Intell. Virtual Agents*, Vol. 2792 (Springer, Boston, United States, 2014) pp. 120–133.

[44] B. Lance and S. C. Marsella, *Emotionally expressive head and body movement during gaze shifts,* in *Proc. Int. Conf. on Intell. Virtual Agents*, Vol. 2792 (Springer, 2007) pp. 72–85.

[45] B. Ravenet, M. Ochs, and C. Pelachaud, *From a user-created corpus of virtual agent's non-verbal behavior to a computational model of interpersonal attitudes,* in *Proc. Int. Conf. on Intell. Virtual Agents*, Vol. 8108 LNAI (Springer, 2013) pp. 263–274.

[46] G. Castillo and M. Neff, *What do we express without knowing? emotion in gesture,* in *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019),* edited by N. Agmon, M. E. Taylor, E. Elkind, and M. Veloso (2019) pp. 702–710.

[47] W. Johal, S. Pesty, and G. Calvary, *Towards companion robots behaving with style,* in *Proc. 23rd Int. Symp. on Robot and Human Interactive Commun.* (IEEE, Edinburgh, Scotland, 2014) pp. 1063–1068.

[48] J. Xu, J. Broekens, K. Hindriks, and M. A. Neerincx, *Effects of bodily mood expression of a robotic teacher on students,* in *Proc. Int. Conf. on Intell. Robots and Syst.* (IEEE, Chicago, Illinois, 2014) pp. 2614–2620.

[49] J. Kim, S. S. Kwak, and M. Kim, *Entertainment robot personality design based on basic factors of motions: A case study with rolly,* in *Proc. Int. Symp. on Robot and Human Interactive Commun.* (IEEE, Toyama, Japan, 2009) pp. 803–808.

[50] M. Saerbeck and C. Bartneck, *Perception of affect elicited by robot motion,* in *Proc. Int. Conf. on Human-robot interaction* (IEEE, Osaka, Japan, 2010) pp. 53–60.

[51] N. Dael, M. Goudbeek, and K. R. Scherer, *Perceived gesture dynamics in nonverbal expression of emotion,* Perception **42**, 642 (2013).

[52] M. Inderbitzin, A. Väljamäe, J. M. B. Calvo, P. F. Verschure, and U. Bernardet, *Expression of emotional states during locomotion based on canonical parameters,* in *Proc. 9th Conf. on Automatic Face and Gesture Recognition* (IEEE, Santa Barbara, CA, 2011) pp. 809–814.

[53] A. L. Baylor and Y. Kim, *Simulating instructional roles through pedagogical agents,* Int. J. of Artificial Intell. in Edu. **15**, 95 (2005).

[54] T.-H. D. Nguyen, E. Carstensdottir, N. Ngo, M. S. El-Nasr, M. Gray, D. Isaacowitz, and D. Desteno, *Modeling warmth and competence in virtual characters,* in *Proc. Int. Conf. on Intell. Virtual Agents* (Springer, Delft, The Netherlands, 2015) pp. 167–180.

5

Effect of Robot Interaction Style on Children's Learning Approach and Cognitive Outcomes

Social robots are entering the educational domain where their aim is to facilitate children in their learning process. This requires the robot to display behaviour supportive to the task at hand and the child's learning style and motivation —both verbally and non-verbally. In inter-human educational interactions the importance of strategic behaviour and expression thereof (i.e., teaching style) is acknowledged. Theoretical models exist, as well as studies of the effects. In human-robot interaction robot role and style expression have been studied. A comprehensive, validated model of pedagogical interaction style for a robot does not yet exists. Our aim is to study the effect of robot stylised behaviour on children's learning approach and learning gain. To this aim we used an existing "rolling-rolling" exhibit where children could learn a scientific phenomenon by experimenting, and designed instructional behaviour for the robot. Robot style is implemented as variations in verbal strategy and non-verbal style expression. We show that the presence of an educational robot improves children's learning approach; children receiving feedback from a robot during execution of a learning task showed to engage with the task longer and conduct more informative experiments. Further, children who interacted with a robot who provided explanations (compared to verbalising observations) perceived the robot as more competent. However, no differences in learning approach nor gain were observed between the different robot styles. Further, our iterative approach allowed us to investigate the experiment setup, which showed that children's attentiveness towards the robot is crucial.

This paper is submitted to International Journal of Social Robotics

5.1. Introduction

Learning through exploration can lead to a better understanding of phenomena and their causes. Consider, for example, the speed of a ball rolling down a grass hill, and the effects of the resistance of the grass and the weight of the ball on this speed. By experimenting with balls of different weight and grass of different lengths, one can learn what these effects are exactly. Such educational "playgrounds" have been developed for children (see the 'rolling-rolling', RR, task of Figure 5.1), and the interesting question is if a social robot can be added to the setting to support this exploratory learning.

More specifically, the question is: Which robot pedagogical interaction style (PInS) should the robot express to facilitate the exploration and bring about the desired learning improvement? Two educative roles are distinguished: provide educational support as an *Expert* or as a *Facilitator*. An *Expert* provides support by knowledge sharing. A facilitator is designed to encourage critical thinking skills.

Note that each learning task needs guidance. The RR-task in our study needs only a minimal guidance for which we designed two (verbal) instruction strategies: Explanations and Evidence Descriptions. The *Expert* explains by explicit mention of cause and effect of the observed outcome of rolling the balls down, for example, "The lane without carpet gives less friction, therefore the roller is down faster". The facilitator provides descriptions of the observations highlighting the differences in variables that cause this outcome, for example, "The roller on the lane without carpet is down faster than the roller on the lane with carpet, you see?".

Furthermore, a robot shows a body language that can support or complement the instruction. In previous experiments, we designed and tested two body or non-verbal behaviour styles: 'Direct' and 'Friendly'. Direct behaviours are high in expansiveness (amongst other modulations are presented in Table 5.1) and therefore perceived as more dominant. They seem to be congruent with the *Expert* role. The friendly behaviours are amongst other modulations fast paced and therefore perceived as warm and affiliate. The friendly behaviours seem to be congruent with the *Facilitator* role.

We implemented these behaviours on a NAO robot, and evaluated the effect thereof on children's task execution and learning gain. For our study we formulated the following research questions:

Research Question 1 *What is the effect of pedagogical interaction style (*Expert *and* Facilitator*) shown by a NAO robot on children's a) learning approach, and b) learning gain?*

Research Question 2 *What is the role of non-verbal behaviour in expression of and effects by these pedagogical interaction styles shown by a NAO robot?*

We hypothesised that in an inquiry-based learning task children will demonstrate greater active involvement, more exploratory behaviour, and improved learning when interacting with a *Facilitator* robot (giving Evidence Descriptions and showing 'Friendly' motions) compared to an *Expert* robot (giving Explanations and showing 'Direct' motions) because the *Facilitator*style was stated by Grasha [1] to be suitable for self-discovery learning and Van Schijndel et al [2] found that children guided by parents using Evidence Descriptionss demonstrated a higher level of exploratory behaviour. Further we hypothesised that non-verbal behaviour plays a crucial role in the expression of these teaching styles and the effect thereof; non-verbal behaviour congruent to verbal strategy (i.e. stylised behaviours supportive to talk) is expected to yield better performances compared to non-stylised random motions because non-verbal behaviours play an important role in expression of style. We acknowledge possible mediating factors such as the child's individual learning capabilities and preferences, but study of these factors is outside the scope of this research.

During the present study we evaluated our approach repeatedly based on observations during the sessions. We adapted our set-up to draw children's attention towards the robot while it was speaking. Ultimately, we questioned whether the robot's presence influenced children's learning outcomes. We formulated a new research question and added a no robot feedback condition to investigate the effect of the educational robot's presence.

Research Question 3 *What is the effect of presence of the NAO robot giving feedback on the child's actions on children's a) learning approach, and b) learning gain?*

In the remainder of this paper, first we provide some necessary background on the subject, followed by the robot behaviour design (the stimuli used in the experiment) and the research materials used. Next, we present in two sections the method, results and discussion of the study iterations. We conclude this paper with an overall conclusion and learnings.

5.2. Background

While social robots emerge as collaboration partner in education, the urgency towards social intelligence is emphasised; they should be able to perceive and express social signals [3, 4]. More specifically, in educational settings they should adapt their role and style of interaction towards the learning activity and learner's characteristics [5]. Therefore, it is important to study what social expressions are appropriate for specific instructional situations.

Mubin et al. [5] categorised educational robots into three roles: tutor, tool, or peer, with the tutor robot functioning as an educator and therefore being of interest in the present study. However, robots as tutor have been given various descriptions and implementations. For example, Belpaeme et al. [6] designed a tutor that regarded the child as the more knowledgeable partner (putting the robot in the learners role), whereas Chen et al. [7] presented the robot as a knowledgeable playmate. Looije et al [8] defined an *educator* and a *motivator*. The motivator was designed to appear empathetic and trustworthy by showing reflective listening (affective reactions, asking questions), positive regards (giving complements), and attentiveness (looking, nodding). The educator showed behaviours similar to the motivator, and explained the correct answer whenever the child answer was incorrect. These roles were used by a single agent and selected based on the activity. Nourbakhsh et al. [9] defined an 'affective educator'; a robot showing happy and sad moods and providing information while giving a museum tour. These educational robots vary in level of (expressed) competence and affect, and were situated differently. These robots have been designed to express a certain role and study the effect thereof; however, validation of role expression in these studies is limited. An overview of validated expressions for robots and virtual agents is given in [10], but these studies do not investigate the effect of these expressions.

Teacher-student interactions in inter-human educational settings have a long standing research interest. Teaching styles are behaviour patterns that affect information presentation and interaction in classrooms [11], in recent years various theoretical models of teaching styles have been proposed. For example, Grasha [1] constructed five teaching styles to address differences in teachers' need for control and teacher-student relationships that result from differences in learners' capabilities (e.g., knowledge, motivation, responsibility). These styles are: Expert (transmitting information), Formal authority (providing feedback and establishing boundaries), Personal model (showing an example), Facilitator (encouraging critical thinking), and Delegator (available in the background during project work). Certain clusters of styles were linked to specific situations and strategies. For example, teacher-centred styles (i.e., expert and formal authority) were linked to teacher control and less capable students, while student-centred styles (i.e., facilitator and delegator) were linked to teachers willing to loosen control and more capable students. Similarly, in the Teacher Interpersonal Circle (TIC) Wubbels et al. [12] proposed eight teaching styles based on the two dimensions affiliation (opposition - cooperation) and control (dominance - submission).

Informal, self-regulated and inquiry-based, learning is promoted by exploration. Effectiveness of learning through investigation is promoted by guidance; guided inquiry resulted in a better learning process [13] and higher learning outcomes [14]. An effective instruction strategy in (informal) learning is *Scaffolding*, this strategy can be applied by both teachers and informal instructors such as parents to promote exploratory behaviour [15]. Scaffolding is based on the theory of the *Zone of Proximal Development* (ZPD) [16], predicting that experiences slightly advancing current abilities encourage and advance learning. As a consequence it requires knowledge about the learners individual developmental stage (e.g., prior knowledge) where the level of guidance should focus on [17]. This knowledge is often present for parents and classroom teachers but less so for museum educators and nurses providing healthcare education. Alternatives that do fit within the scaffolding principles but do require less extensive knowledge of the learner are *evidence descriptions* and *open-questions* [2]. Open-questions challenge the learner to think about his/her actions during explorations or may be used to hint an idea; the learning process is hereby directed [18]. Typically questions are predictive or about the cause of an observation. Open-questions promote critical thinking, but are invasive as they requires a response from the learner [19]. Evidence descriptions direct the attention of the learner to relevant variables and task elements and thereby structure the learning process. The instructor will highlight certain aspects by mentioning the observations, without explanation.

Effects of these instruction strategies have been investigated before in inter-human children's instructional coaching. Dependent on a specific museum exhibition, children showed more exploratory behaviour when they were guided with affective talk compared to when receiving explanations, when receiving explanations children directed their attention towards the exhibition but did not actively manipulate it [20]. In follow-up study children displaying the highest level of exploratory behaviour were mostly accompanied by parents giving feedback in the form of evidence descriptions [2]. The effect of increased exploratory behaviour on learning outcomes is not always established (see e.g., [18, 19, 21, 22]). Schulz [23] suggested that "children's exploratory play is affected by the quality of the evidence" while at the same time children lack skills to designing informative experiments (i.e., children fail to control relevant variables and tend to alter multiple variables simultaneously).

Inter-human coaching influences children's experimental behaviour, and proper guidance may assist them in designing informative experiments, thereby stimulating learning. As learning takes place more and more out of the classroom and on a individual or small group basis it is relevant to study the effect of coaching by a robot. Additionally, robots have the ability to select and perform actions and behaviours in a consistent and strategic way, making them especially useful as highly controllable social stimuli [24]. As shown in [25, 26] manipulation of non-verbal robot behaviour will lead to children having a different perception of the robots level of friendliness, competence and dominance. The current study aims to build on these insights and integrate them with the findings from research on human verbal instruction styles. Specifically interesting is the fact that we can control verbal and non-verbal behaviour separately, allowing us to distinguish the effect of both channels separately.

The design of robot pedagogical interaction style comprises the a) strategic action selection (i.e., what is the best strategy to support the learner given the context of use and user characteristics), and b) the interaction style (i.e., appropriate style of communication— and non-verbal expression thereof). In prior studies we have designed and validated synthesised expressions of warmth and competence [25] and of dominance [26, 27]. These factors contribute to teaching style expressions; which have been implemented in the PAL system and tested for their effect on children's system usage for their health-care education [28]. In that study the authors defined a 'Friendly' style to play (educational) games with a child and a 'Direct' style for activities such as reviewing learning goal progression. However, the data of this study was too limited to systematically study the effect of these validated styles on children's learning outcomes.

5.3. Robot Behaviour Model

To study the effect of robot pedagogical interaction stylewe created four specific scripts with verbal and non-verbal behaviours. The *Expert* style script is constructed based on the verbal strategy Explanations and the non-verbal style 'Direct'. The *Facilitator* style script is constructed based on the verbal strategy Evidence Descriptions and the non-verbal style 'Friendly'. To untangle the contribution of the non-verbal behaviour, we designed 'neutral' non-verbal behaviours as well. The two remaining scripts use the two verbal strategies each with neutral non-verbal behaviours.

We use the neutral, 'Direct' and 'Friendly' non-verbal style and according behaviours as created for the integrated PAL3.0 system and experiment [29]. The neutral behaviours are baseline behaviours that do not express a specific style but may be deictic or iconic. The 'Friendly' style expresses high friendliness and low competence and dominance, whereas the 'Direct' style expresses low friendliness and high competence and dominance. Note that the friendly style does not mean to present an character lacking knowledge (such as a peer), but restrains itself from the expression of possessed knowledge in order to promote the learner's sense of knowledge-ability. Motion parameters to display each factor (i.e., friendliness, competence and dominance) have been investigated in [10] and resulted in the model presented in Table 5.1 and Table 5.2.

We use the Explanations and Evidence Descriptions verbal behaviours as used before in human instruction experiments reported in [2, 19, 22]. The Explanations strategy mentions explicitly the cause and consequence of the outcome of scientific trial. The Evidence Descriptions strategy mentions the observation of the outcome emphasising differences in variables that influence this outcome. The scientific trials in this study involve rolling objects of varying weights down slopes with different surfaces (see Section 5.4 for details on the 'Rolling-Rolling' task [20]). We differentiate outcomes as valid or invalid, with invalid trials characterised by errors preventing meaningful conclusions such as the rolling objects hitting a wall slowing them down. Valid trials can be either using a control of variables strategy (CVS) meaning that only one (1) variable was systematically varied or none-CVS meaning that there is no valid conclusion based on varied variables such as when using only one object of varying both object weight and slope surface. All options and the final robot scripts are presented in Table 5.3. We were restricted to nine possible feedback options per verbal style due to technical limitations of the WOz set-up.

5

5.4. RESEARCH MATERIAL

We use a NAO robot developed by Softbank (`https://us.softbankrobotics.com/nao`). The NAO is approximately 60cm high, with 25 DOF but no facial expression. The behaviour was scripted with the robots-in-de-klas cloud service (`https://www.interactive-robotics.com/onderwijs/`) and using custom behaviours created in Choregraphe (`http://doc.aldebaran.com/1-14/software/choregraphe/`

Table 5.1
MOTION PARAMETERS AND VALUES FOR EACH NON-VERBAL STYLE

Parameter Style	Gaze Gaze	Gesture size	Gesture openness	Hand position horizontal	Hand position vertical	Head tilt	Gesture speed
Neutral	Fixed	Medium	Centred	Sideways	Centred	Straight	Medium
Friendly	Towards	Large	Centred	Peripheral	Up	Up	Fast
Direct	Towards	Medium	Open	Peripheral	Up	Straight	Medium

Table 5.2
NAO JOINT VALUES FOR THE STAND POSTURE FOR EACH NON-VERBAL STYLE. (RIGHT JOINTS ARE MIRRORED.)

Joint	(Parameter)	Default	Neutral	Direct		Friendly	
`HeadYaw`	(gaze)	-0.1	0.0	changing	(towards)	changing	(towards)
`HeadPitch`	(head tilt)	-9.6	0.0	-20	(straight)	-10	(up)
LShoulderPitch		80	80	70		60	
LShoulderroll		10.4	8.0	35		20	
LElbowYaw		-68.9	-70	-35		-100	
LElbowRoll		-23.9	-24	-80		-50	

Table 5.3
SENTENCES OF THE VERBAL INSTRUCTION STYLES FOR EACH POSSIBLE ROLLING TRIAL OUTCOME (TRANSLATED FROM DUTCH).

		Explaining	**Evidence description**
1	valid, CVS, roller	The weight of the roller does not matter for how fast the roller is down.	The light and heavy roller are down at the same moment, do you see!
2	valid, none-CVS	We don't know what causes the difference because you used multiple rollers. Use two rollers at the same time.	I see you used more than two rollers.
3	valid, CVS, lane	The high carpet has more friction than the other carpet, so the roller is down later.	The roller on the lane with high carpet is down later than the lane with the low carpet, you see.
4	valid, CVS, lane	The lane without carpet gives less friction, therefore the roller is down faster.	The roller on the lane without carpet is down faster than the roller on the lane with carpet, you see.
5	valid, none-CVS	You used only one roller. Now we cannot compare what makes a roller go faster or slower.	I see that you rolled down only one roller.
6	valid, non-CVS	You rolled down the same rollers off the same lane. Now we don't know why they are faster or slower. Use two different rollers or roll them of different lanes.	I see that you rolled the same roller of the same slope. There are different rollers and lanes with different materials, you see.
7	valid, none-CVS	We don't know what caused the difference in speed because you used different rollers on different lanes.	I see that you rolled down different rollers from different lanes.
9	invalid	One of the rollers went crooked. As a result, we do not know whether the test went well. Therefore it is good to repeat a test a number of times.	I saw that one roller went a bit crooked, while the other rolled straight down.
8	invalid	The rollers did not start at the same time. As a result, we do not know whether the test went well. Place the rollers on the wooden cover, then pull the lever so that the rollers start evenly.	I saw that one roller started rolling earlier than the other. With the lever you can let the rollers down, you see. You can put two rollers next to each other.

choregraphe_overview.html).

Further, we used the 'rolling-rolling' exhibit (from NEMO science museum, https://www.nemosciencemuseum.nl/nl/, and used in [20]). This version has three slopes with different materials (i.e., wood, high fibre carpet and low fibre carpet) and two different rollers (i.e., heavy and light) as shown in Figure 5.1 (see in use at https://youtube.com/shorts/4QDnrqWnZME). The exhibit allows for investigation of influence of surface material on rolling speed and the most common misconception: roller weight influences rolling speed [20, 30]. Variables not of interest were controlled by ensuring the rollers were identical in volume and size, differing only in colour for recognition. Rollers can be rolled down the slopes, each slope is wide enough to roll down two rollers simultaneously. Children could design experiments with independent variables: weight of the roller, and surface texture of the slope. Valid experiments require the child to release both rollers at exactly the same time. There is a handle at the top to ensure rollers are started simultaneously. Informative experiments require the child to vary only one (1) variable per experiment.

The NAO robot is collaborating with the child in experiments at the 'rolling-rolling' exhibit. Depending on the child's constructed trial (i.e., the used rollers, selected slopes and trial validity), the robot provides feedback and displays specific behaviours as presented in 5.3. The behaviours (both verbal and non-verbal) are pre-programmed, but selected in real-time in a Wizard of Oz setup.

Early in the experiment, after 20 children had participated (equally spread over the four conditions), we improved the experiment set-up slightly. We observed that the children did hardly pay attention to the robot during their 'rolling, rolling' activity. Instead of looking at the robot when it was giving feedback, the children would get the rollers and plan their next attempt. A second observation was that the rollers would skew, hit the

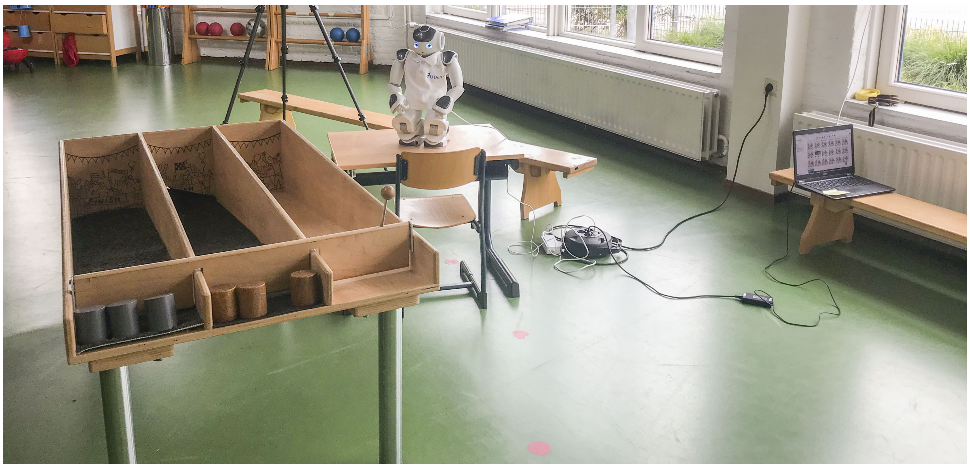


Figure 5.1: The 'rolling-rolling' exhibit in the present study set-up. This exhibit allows children to investigate the effect of roller weight and surface material on rolling speed by rolling two or more rollers down (starting simultaneous by using the lever) and observing rolling speed.

wall of the lane which made them slow down and ultimately lead to false expectations of a specific roller/lane being faster. Especially when rollers were rolled from the same slope and expected to roll at the same speed this could lead to erroneously feedback selection; it was hard to observe by the researcher in real time and to make sure the robot gave the appropriate response. To draw the children's attention toward the robot during feedback we added a spotlight on the 'rolling, rolling'-exhibit as a reminder to look at the robot while giving feedback. The spotlight would turn off when the rollers where down and the robot was speaking. The spotlight would turn on again after the robot was finished speaking, prompting the child to retrieve the rollers and begin the next attempt. Also five motivational sentences were added that directed attention towards the robot which could be selected at random via the WOz interface. Further, we changed from rollers to balls to minimise the slow-down effect from the skewness. Balls did vary in weight and colour but not in size, volume and material.

5.5. PART I: EFFECT OF VERBAL AND NON-VERBAL BEHAVIOUR

To evaluate the effect of robot pedagogical interaction style we conducted an experiment wherein we measured children's learning approach and learning gain in a inquiry-based learning activity while being guided by a robot displaying four pedagogical interaction styles. This study is inspired by a study with human participants by Van Schijndel et al. [20] who studied the effect of two verbal instruction strategies on children's behaviour at the 'rolling, rolling'-exhibit, but with a robotic instructor instead of a human and with additionally systematic variation in non-verbal style expression.

5.5.1. EXPERIMENTAL DESIGN

We designed an experiment with four conditions, two pedagogical interaction styles (i.e. *Expert* and *Facilitator*) that consist of congruent verbal and non-verbal behaviour (stylised and strategically selected), and for each verbal style a variation with non-stylised, neutral behaviours. Resulting in a total of four style conditions: 1) *Expert* consisting of verbal strategy Explanations and non-verbal style 'Direct'; 2) verbal strategy Explanations with neutral non-verbal behaviours; 3) *Facilitator* consisting of verbal strategy Evidence Descriptions and non-verbal style 'Friendly'; and 4) verbal strategy Evidence Descriptions with neutral non-verbal behaviours. Hereby we study replication of effects from instruction styles by robots, and investigate the role of non-verbal behaviour in style expression.

MEASUREMENTS

We measured learning gain, learning approach and perceived robot style. Perception of the robot was measured as a manipulation check, using the adjective rating as described in [25]. This instrument measures perceived dominance, affiliation, warmth, and competence by rating of 20 adjectives on a 3-point scale. These four factors form the impression of pedagogical interaction style such as expert and facilitator. Learning gain was measured with a pre- and post-test; a short questionnaire evaluating children's (prior) knowledge on the subject of the task. This test based on Siegler's Rule Assessment Methodology [31] and identical to the pre- and post-test used by Van Schijndel et al. in their human-human experiments [2, 21]. Before and after the 'rolling, rolling' activity, children were given a total of twelve questions. Each question presents a rolling-trial (i.e., used rollers and/or slopes) and three possible answers (i.e., roller a is fastest, roller b is fastest, the rollers have equal rolling speed). The researcher showed the rollers and held these above specific lanes. The child is asked to predict the correct answer (without seeing the actual experiment performed) to which roller is faster or rollers have equal speed. Each correctly answered question gives one point; for each test the child can earn a maximum score of twelve points, six (6) points for weight (rollers were evaluated) items and six (6) points for lane (slopes were evaluated) items. Learning gain is calculated by pre-test score minus post-test score. Learning approach is measured by behavioural observations of number of trials with usage of control of variables strategy (CVS), duration of play, and total number of trials. CVS implies only one (1) variable of interest is manipulated while all other variables are constant [32, 33]. The CVS scores the number of informative experiments. This scoring of behaviour is done offline, therefore video recordings were made of the interactions. Duration of play in minutes and number of experiments are included as indicators for motivation. At the start the child was instructed that he/she may quit at anytime whenever he/she feels the activity is done for whatever reason. After 10 minutes the end of the activity is initated by the robot; a total playtime of less then 10 minutes means that the child used the option to ask to quit.

Participants

A total of 53 children (after school care, $n = 32$, and primary school, $n = 21$) in the Netherlands, participated in our study. Children were 8-11 years of age ($M = 9.61, SD = 0.83$). Gender was near evenly distributed (male = 25, female = 28). All participants were naive to the research aim and had gotten acquainted with the NAO robot before. Children were almost equally spread over the four robot styles (*Expert* $n = 14$, Explanations-neutral $n = 14$, *Facilitator* $n = 13$, Evidence Descriptions-neutral $n = 12$).

Procedure

At the start children got acquainted with the robot in a class-session where the researcher gave a short demonstration and children were allowed to ask questions, feel the robot and/or make it move. Children participated in the experiment one by one. The duration of an individual session was about 30 minutes including briefing and questionnaire filling. The child entered the room, received the briefing and a minimal introduction by the researcher and was asked to complete the pre-test. Next, the child collaborated with the robot in the science experiments with the 'rolling, rolling'-exhibit where the robot would first introduce the task followed by feedback upon the child's actions after each rolling trial. The child could play a maximum of 10 minutes but was allowed to quit at any time. Finally, the child was asked to complete the post-test and the perception questionnaire. Children were sequentially assigned to one of the four robot conditions to ensure an even distribution. As a consequence groups were not controlled for age and gender.

5.5.2. Results

We explored the differences in children's perceptions of the robot and children's learning approach and learning gain.

PERCEPTION

To thoroughly analyse the differences in children's perception of the robot caused by the differences in robot pedagogical interaction style we performed analyses both between the four conditions as well as between verbal or non-verbal conditions only. The mean perception score for each factor within each of the four conditions is presented in Table 5.4.

For analysis of perceptions we performed a MANOVA. Using Pillai's trace, there was no significant effect of robot interaction style on how children perceived the robot, $V = 0.31, F(12,144) = 1.38, p = 0.184$. Separate univariate ANOVAs on the outcome variables revealed no effect for robot style on perceived affiliation $F(3,49) = 1.24, p = 0.307$, warmth, $F(3,49) = 0.82, p = 0.487$, competence, $F(3,49) = 2.13, p = 0.108$, and dominance, $F(3,49) = 1.14, p = 0.341$. In other words, robot pedagogical interaction style does not influence children's perception of the robots expression of warmth, affiliation, competence nor dominance.

To investigate the individual effects of verbal and non-verbal behaviour manipulations on children's perception of the robot we performed a MANOVA analysis with verbal and non-verbal behaviour as separated fixed factors. Using Pillai's trace, there was a significant effect of verbal behaviour on how children perceived the robot, $V = 0.23, F(4,46) = 3.43, p = 0.016$, there was no significant effect of non-verbal behaviour on how children perceived the robot, $V = 0.16, F(8,94) = 1.00, p = 0.440$. Separate univariate ANOVAs on the outcome variables for verbal behaviour revealed a significant effect on perceived competence, $F(1,49) = 5.62, p = 0.022$, and no significant effect for verbal behaviour on perceived affiliation, $F(1,49) = 3.34, p = 0.074$, warmth, $F(1,49) = 2.18, p = 0.146$, and dominance, $F(1,49) = 3.05, p = 0.087$. Separate univariate ANOVAs on the outcome variables for non-verbal behaviour revealed no significant effect on perceived competence, $F(2,49) = 0.66, p = 0.521$, affiliation $F(2,49) = 1.30, p = 0.283$, warmth, $F(2,49) = 1.02, p = 0.368$, and dominance, $F(2,49) = 0.98, p = 0.383$. In other words, the verbal behaviour manipulation had an effect on children's perception of the robot's level of competence in favour of the Explanations strategy. The non-verbal style manipulation did not generate differences in children's perception of the robot.

Table 5.4

MEAN SCORES FOR CHILDREN'S PERCEPTION OF THE ROBOT SPLIT BY DEPENDENT VARIABLES THE PERCEPTION FACTORS AND INPUT VARIABLES VERBAL AND NON-VERBAL BEHAVIOUR. VERTICALLY DEPICTED THE VERBAL STYLES EXPLAINING AND EVIDENCE DESCRIPTIONS. HORIZONTALLY THE NON-VERBAL NEUTRAL STYLE (N) OR A CONGRUENT (C) NON-VERBAL STYLE EITHER BEING 'DIRECT' OR 'FRIENDLY' DEPENDING ON THE VERBAL STYLE. IN ITALICS THE SCORES HYPOTHESISED TO BE HIGHEST.

	Dominance		Affiliation		Competence		Warmth	
	C	**N**	**C**	**N**	**C**	**N**	**C**	**N**
Explaining	*2.00*	2.51	16.89	18.18	*9.37*	10.20	14.97	15.78
Evidence Descriptions	2.01	1.79	*17.27*	16.55	8.64	8.07	*15.46*	14.38

Effect on Learning

For the analysis of the effect of robot pedagogical interaction style on learning approach and learning gain we performed separate MANOVA analyses. The analysis of learning approach included the dependent variables: 1) duration of play in minutes, 2) total number of experiments, and 3) total number of informative experiments. The analysis for learning gain included the dependent variables: 1) difference between pre- and post-test score for lane items, and 2) difference between pre- and post-test score for weight items. The mean scores for each variable per condition is presented in Table 5.5.

Using Pillai's trace, there was no significant effect of robot pedagogical interaction style on children's learning approach, $V = 0.13, F(9,144) = 0.67, p = 0.710$. Separate univariate ANOVAs on the outcome variables revealed no effect for robot style on any of the dependent measures: duration of play $F(3,48) = 0.69, p = 0.563$, number of experiments, $F(3,48) = 1.17, p = 0.329$ and number of CVS experiments, $F(3,48) = 0.24, p = 0.870$. Using Pillai's trace, there was no significant effect of robot pedagogical interaction style on childrens learning gain, $V = 0.63, F(6,98) = 0.53, p = 0.786$. Separate univariate ANOVAs on the outcome variables revealed no effect for robot style on difference in pre- and post-test scores for lane, $F(3,49) = 0.34, p = 0.801$, and weight, $F(3,49) = 0.76, p = 0.521$. In other words, variations in pedagogical interaction style did not influence children's how children execute the activity nor their knowledge gained from the activity.

The perception results indicate that children do perceive the robot differently based on its verbal behaviour. To investigate the individual effects of verbal behaviour manipulations on children's learning approach and learning gain we performed MANOVA analyses with verbal strategy (Evidence Descriptions or Explanations) as fixed factor. Using Pillai's trace, there was no significant effect of verbal behaviour on children learning approach, $V = 0.36, F(3,48) = 0.59, p = 0.624$. Separate univariate ANOVAs on the outcome variables revealed no effect for robot verbal behaviour on any of the dependent measures: duration of play, $F(1,50) = 1.33, p = 0.255$, number of experiments, $F(1,50) = 0.01, p = 0.928$ and number of CVS experiments, $F(1,50) = 0.14, p = 0.906$. Using Pillai's trace, there was no significant effect of verbal behaviour on children learning gain, $V = 0.03, F(2,50) = 0.76, p = 0.475$. Separate univariate ANOVAs on the outcome variables revealed no effect for robot verbal behaviour on difference in pre- and post-test scores for lane, $F(1,51) = 0.07, p = 0.796$, and weight, $F(1,51) = 1.50, p = 0.227$. In other words, we found no evidence for an effect of verbal behaviour of the robot on children's learning approach or learning gain.

Table 5.5

Mean effect scores split by factor and condition (SD in parentheses). Vertically depicted the verbal styles Explanations (Ex) and Evidence Descriptions (ED). Horizontally the non-verbal neutral style (N) or a congruent (C) non-verbal style either be 'Direct' or 'Friendly'

	Learning Approach						Learning Gain			
	Duration		experiments		CVS experiments		Lane items		Weight items	
	C	**N**	**C**	**N**	**C**	**N**	**C**	**N**	**C**	**N**
Ex	9.62 (0.96)	9.29 (1.82)	17.54 (3.26)	15.86 (3.66)	8.77 (2.56)	9.00 (4.42)	1.36 (2.21)	1.64 (3.30)	0.71 (1.98)	0.57 (1.83)
ED	8.69 (2.21)	9.17 (1.34)	15.31 (4.89)	17.92 (4.80)	8.46 (3.55)	9.58 (2.71)	1.23 (1.79)	2.17 (2.59)	0.38 (1.04)	-0.17 (1.12)

5.5.3. Discussion & Conclusion

The analysis shows no difference in learning approach or learning gain between the four pedagogical interaction styles. These results can be explained by the observation that the children did not pay visual attention to the robot —children would grab the rollers from the lower end of the 'rolling, rolling'-exhibit while the robot was speaking, thus not looking at the robot. Adjustments to the experiment set-up, as discussed in section 5.4, could not prevent this lack of attention towards the robot. This is in line with the perception measures not giving the results expected based on earlier studies (i.e., [25–27]); the non-verbal behaviour manipulations did not result in any difference in children's perception of the robot. While the children did not look at the robot they could still hear the robot. In the perceptions scores this is reflected in a small but significant difference in perceived competence between the two verbal strategies. Children listening to Explanations considered the robot to be more competent compared to a robot giving Evidence Descriptions. However, this difference in perception did not result in a difference in learning approach nor learning gain.

Based on the obtained results we cannot conclude whether or not robot pedagogical interaction styles influences children's learning outcomes. To influence children's behaviour it is required for them to observe and notice the robot behaviour. With this study we show that it is not evident that children will pay attention to the robot just because it is a novel stimulus. Children are too engaged in the 'rolling, rolling' activity and simple changes we made to the experimental set-up designed to draw children's attention towards the robot appeared unable to break that engagement. In our previous perception studies the robot was the centre of the activity. It remains unclear what is needed to accomplish children paying attention to the robot and possibly noticing robot behaviour in a inquiry-based learning activity with robot guidance.

5

5.6. Part II: No Robot Guidance

Obviously if children do not pay attention towards the robot and do not notice it's behaviour, they remain uninfluenced by the variations in the behaviour. To investigate the effect of robot presence, we added a fifth condition and recruited children in the same age-range. In this condition children did receive a task instruction from the robot but did not receive guidance from the robot after rolling trials.

5.6.1. Experimental Design

We designed a follow-up experiment with two conditions, with robot guidance (including all four style conditions and 53 participants from the previous experiment) and without robot guidance. Hereby we investigate the effect of robot guidance during the 'rolling, rolling' activity on learning approach and learning gain.

Measurements

Learning approach and learning gain were measured as mentioned in Section 5.5.1. Perception was not measured because there was no extended interaction with the robot and no variation in robot behaviour between participants.

Participants

An additional 20 children aged 8-11, were recruited at a primary school in the Netherlands. Combined with the 53 children from before, a total of 73 children, from after school care ($n = 32$) and a primary school ($n = 41$) in the Netherlands, participated in our study. Children aged 8-11 participated ($M = 9.62, SD = 0.78$) and slightly more boys that girls (male = 40, female = 33).

Procedure

The procedure was equal to Part I (see Section 5.5.1) except for the robot guidance during play. Children did receive instruction from the robot before play, but were not provided any feedback after each rolling trial. The robot remained present showing idle movements but did not speak. Further, the children did complete the pre- and post-test but not the perception questionnaire.

5.6.2. Results

5

To analyse the effect off robot guidance on learning approach and learning gain we performed separate MANOVA analyses. The analyses included the same dependent variables as in Part I (see Section 5.5.1). The mean scores for each variable per condition is presented in Table 5.6.

Using Pillai's trace, there was a significant effect of robot guidance on children's learning approach, $V = 0.43, F(3,69) = 17.40, p < 0.001$. Separate univariate ANOVAs on the outcome variables revealed a significant effect for robot guidance on duration of play, $F(1,71) = 10.81, p = 0.002$, and number of CVS experiments, $F(1,71) = 25.15, p < 0.001$. There was no significant effect for robot guidance on number of experiments, $F(1,71) = 0.39, p = 0.536$. Using Pillai's trace, there was no significant effect of robot guidance on childrens learning gain, $V = 0.06, F(2,70) = 2.14, p = 0.125$. Separate univariate ANOVAs on the outcome variables revealed no significant effect for robot guidance on difference in pre- and post-test scores for lane, $F(1,71) = 0.94, p = 0.335$, and weight, $F(1,71) = 3.73, p = 0.057$. In other words, the presence of a robot giving feedback after rolling trails did influence children's learning approach, but we found no evidence for an influence on the knowledge gained from the activity. Children who executed their rolling trials without robot played shorter and did less informative experiments while the total number of trials was the same between the children who played with and without robot.

Table 5.6

Mean effect scores split by factor and condition (SD in parentheses). Vertically depicted the verbal strategy Explanations (Ex) and Evidence Descriptions (ED). Horizontally the non-verbal neutral style (N) or a congruent (C) non-verbal style either be 'Direct' or 'Friendly', depending on the verbal strategy.

	Learning Approach			Learning Gain	
	Duration	Experiments	CVS experiments	Lane items	Weight items
No-robot	7.40 (3.02)	15.70 (9.03)	4.00 (4.17)	0.15 (1.39)	0.45 (1.32)
robot	9.21 (1.63)	16.66 (4.18)	8.81 (3.45)	0.57 (1.72)	1.58 (2.49)

5.6.3. DISCUSSION & CONCLUSION

The analyses shows a clear difference in learning approach between children who did their rolling trials with or without robot guidance. Children did the same number of trials but played longer and did more informative trials with robot guidance. The increase of informative trials can be explained by the observation that without robot guidance children were never corrected for invalid trials such as not using the lever to ensure rollers start synchronous. With robot guidance children received feedback (see Table 5.3) including hints on how to improve their trials. The fact that children who received robot guidance played longer, whilst doing the same total number of trials, possibly indicated that children make more deliberate choices of trials based on robot feedback. This is supported by the increased number of CVS experiments for children who did receive robot guidance. An alternative explanation for longer playtime is that the children are slowed down by the robot guidance in it self. However, this conflicts with the observation that children are not explicitly looking at, and thus waiting for, the robot while it gives the feedback. The decreased duration of play without robot guidance may also be explained by children's lack of understanding of the goal of rolling trials, insecurity about the next action to plan and thereby decreased engagement with the activity which we inferred from informal observations of children looking around and asking the researcher what to do leading up to asking to quit. These results do once more underline that the children are listening to the robot and are positively influenced by the robot's feedback, even though they are not explicitly paying at the robot.

5

The analysis showed no difference in learning gain but we do see some indications that robot feedback did improve cognitive outcomes. Further inspection of the data for learning approach showed that the difference in number of CVS experiments between children who did and didn't receive robot guidance is depends on the varied variable. This difference was significant when weight was varied (robot, $M = 4.06, SD = 2.82$, no-robot, $M = 1.20, SD = 1.36, F(1,71) = 18.78, p < 0.001$) but not when lane was varied (robot, $M = 4.13, SD = 2.84$, no-robot, $M = 2.80, SD = 3.64, F(1,71) = 2.73, p = 0.103$). This corresponds with the nearly significant learning gain results for weight, $F(1,71 = 3.73, p = 0.057$ and not significant results for lane, $F(1,71) = 0.94, p = 0.335$). The fact that the results for lane items are not significant for both learning approach and gain can be explained by the observation that children already understand the effect of surface material on rolling speed; during pre-test children often commented on choosing for a specific lane to be faster because the surface has less friction. This created a ceiling effect for lane items where children can hardly do better (pre-test, $M = 4.78, SD = 1.26$), especially when already receiving the maximum score of 6 points in the pre-test ($n = 27$). Further it influences their approach because children are less inclined to repeatedly execute trials for which the outcome is known to them. The 'rolling-rolling' exhibit was originally designed for and used with children in kindergarten (4-6 y.o.) in the inter-human study [20]. The participants in the present study where 7-11 so the task seems to have been a bit too easy for them. For weight items the pre-test scores were considerably lower ($M = 0.90, SD = 1.03$). While some learning did take place (post-test, $M = 2.18, SD = 2.10$) and children scored higher in the post-test with robot guidance ($M = 2.45, SD = 2.28$) then without ($M = 1.45, SD = 1.36$), the difference was not significant. This indicates that the robot guidance was helpful to facilitate doing more weight related trails but the guidance did not support children in increasing their understanding on —and thereby correct prediction of —the (lack of) influence of roller weight on rolling speed. Noteworthy, we changed from rollers to balls during the experiment because the rollers would skew, hit the wall and slow down quite often. This obviously influences the rolling trial outcome but was hard to observe in time for the experimenter leading to erroneous feedback from the robot (i.e., commenting that rollers rolled at equal speed while the child observed one to be sooner). The influence of this change on children's learning was beyond the scope of this study, but significant difference for lane post-test items between children (with robot guidance) using rollers ($M = 2.55, SD = 2.12$) and balls ($M = 2.39, SD = 2.34$) was not found ($F(1,51) = 0.06, p = 0.811$).

Based on the results, we cannot conclude that robot guidance has an effect on learning gain. An activity that allows for a wider range of learning opportunities may be favourable to increase the measurable effects in learning gain. Even though the activity was carefully crafted to facilitate learning within the zone of proximal development, the lane items appeared to not provide enough of a challenge to the children, while the weight items were not addressed appropriately by the robot feedback.

5

5.7. Overall Discussion

We highlighted the importance of personalised and situation-aware robot pedagogical interaction style. This study investigated the effect of different robot pedagogical interaction styles in one specific educational context. To place these results in a broader context it is necessary to study the presented styles with different activities as well; based on the work of Grasha [1] we expected the *Facilitator* style to yield more experimental behaviour in this inquiry-based learning task where the *Expert* style may be more effective for introducing new factual knowledge. A possible cause is that by giving explanations the instructor takes a knowledgeable role, and as a consequence the learner takes a more passive role. This may evoke a feeling of incompetence [34]. However, giving explanations does provide the learner with theoretically correct information which may have a positive effect on learning gain. On the contrary, children are suggested to be more engaged in exploration when provided with ambiguous evidence compared to unambiguous [35]. Thus we expect full explanations to hinder engagement in exploratory behaviour, whereas evidence descriptions provide focus for the child but leave investigation up to them. Further, prior research suggests a relation between exploratory behaviour and learning gain [18].

In the present study we studied the effect of verbal instruction styles and additional non-verbal behaviour in an inquiry-based learning task. We hypothesised increased learning gain for Evidence Descriptions and congruent non-verbal behaviour compared to neutral non-verbal behaviour and/or Explanations. We did not observe these hypothesised effects in the results. For the non-verbal behaviour part this is to be explained by the observation that children do not look at the robot while it gives feedback; children do not see the robot and do not report any differences in perception based on non-verbal behaviour. Children do hear the robot feedback, indicated by perception differences found in level of competence based in verbal behaviour and differences in learning approach between children who did and didn't receive robot feedback during rolling trials. However, there was no difference in learning approach or gain between the two types of verbal behaviour. Meaning that we did not find proof that children's learning is influenced by the type of verbal feedback.

5.8. CONCLUSION

5.8.1. RESEARCH QUESTIONS

Based on the results of the present study, we did not find any effect of robot pedagogical interaction style on children's learning approach or gain (RQ1). We could not study the role of non-verbal behaviour in the effect of robot pedagogical interaction style because no effect thereof was found (RQ2). Based on only verbal behaviour variations children did perceive the robot providing explanations as more competent than the robot giving evidence descriptions. This perception difference did not impact learning approach or gain in the present study. The presence of a robot giving feedback on children's rolling trials led children to play longer and do more informative experiments, however, this difference in learning approach did not impact learning gain from the rolling trials (RQ3).

5.8.2. LESSONS LEARNED

Based on results and observations of the present study we learned that the task selection is critical. With respect to the robot, it should be the primary focus; the task should not divert attention from it. With respect to the task, it should be challenging enough to show a learning effect (avoid ceiling effect). With respect to the feedback, it should be validated for its learning effect and optimised with human educators as models.

ACKNOWLEDGEMENT

We thank the children, parents and teachers of the participating daycare 'Zonnekinderen' and primary schools. We thank Nemo Science museum for lending us the rolling rolling exhibition. We thank iXperium for participant recruitment.

STATEMENTS AND DECLARATIONS

This work was partly funded by the EU Horizon 2020 PAL project (grant nr. 643783). The authors have no competing interests to declare that are relevant to the content of this article.

COMPLIANCE WITH ETHICAL STANDARDS

Selection and Participation of Children: We recruited 73 children (aged 8-11, M/F) from after-school care and primary schools in the Netherlands. An invitation to participate was offered to teachers who were contacted after they expressed their interest to iXperium centre of expertise (https://www.ixperium.nl/). Written informed consent was obtained from legal guardians. Verbal consent from the child was obtained prior to the experiment. The research was approved by the Human Research Ethics Committee of Delft University of Technology and subject to an approved Data Management Plan. The dataset generated and analysed during the current study is available in the 4TU repository, https://doi.org/10.4121/21803781.

REFERENCES

[1] A. F. Grasha, *A matter of style: The teacher as expert, formal authority, personal model, facilitator, and delegator,* College teaching **42**, 142 (1994).

[2] T. J. Van Schijndel and M. E. Raijmakers, *Parent explanation and preschoolers' exploratory behavior and learning in a shadow exhibition,* Science Education **100**, 153 (2016).

[3] T. Fong, I. Nourbakhsh, and K. Dautenhahn, *A survey of socially interactive robots,* Robot. and autonomous syst. **42**, 143 (2003).

[4] M. Pantic, R. Cowie, F. D'Errico, D. Heylen, M. Mehu, C. Pelachaud, I. Poggi, M. Schroeder, and A. Vinciarelli, *Social signal processing: the research agenda,* in *Visual analysis of humans* (Springer, 2011) pp. 511–538.

[5] O. Mubin, C. J. Stevens, S. Shahid, A. Al Mahmud, and J.-J. Dong, *A review of the applicability of robots in education,* Journal of Technology in Education and Learning **1**, 13 (2013).

[6] T. Belpaeme, J. Kennedy, A. Ramachandran, B. Scassellati, and F. Tanaka, *Social robots for education: A review,* Science robotics **3** (2018).

[7] H. Chen, H. W. Park, and C. Breazeal, *Teaching and learning with children: Impact of reciprocal peer learning with a social robot on children's learning and emotive engagement,* Computers & Education **150**, 103836 (2020).

[8] R. Looije, M. A. Neerincx, and V. d. Lange, *Children's responses and opinion on three bots that motivate, educate and play,* Journal of Physical Agents **2** (2008).

[9] I. R. Nourbakhsh, J. Bobenage, S. Grange, R. Lutz, R. Meyer, and A. Soto, *An affective mobile robot educator with a full-time job,* Artificial intelligence **114**, 95 (1999).

[10] R. Rijgersberg-Peters, J. Broekens, and M. A. Neerincx, *Parameter-based modulation of bodily shape and motion in robots andvirtual agents: A taxonomy and review,* (submitted).

[11] R. M. Felder, L. K. Silverman, *et al.*, *Learning and teaching styles in engineering education,* Engineering education **78**, 674 (1988).

[12] T. Wubbels, M. Brekelmans, P. D. Brok, J. Levy, T. Mainhard, and J. V. Tartwijk, *Let's Make Things Better –Developments in Research on Interpersonal Relationships in Education,* in *Interpersonal relationships in education* (Sense Publishing, Rotterdam, 2012) Chap. 14, pp. 225–249.

[13] A. W. Lazonder and R. Harmsen, *Meta-analysis of inquiry-based learning: Effects of guidance,* Review of educational research **86**, 681 (2016).

[14] L. Alfieri, P. J. Brooks, N. J. Aldrich, and H. R. Tenenbaum, *Does discovery-based instruction enhance learning?* Journal of educational psychology **103**, 1 (2011).

5

[15] W. Crain, *Theories of development: Concepts and applications: Concepts and applications* (Psychology Press, 2015).

[16] L. S. Vygotsky, *Mind in society: The development of higher psychological processes* (Harvard university press, 1980).

[17] P. van Geert and H. Steenbeek, *The dynamics of scaffolding,* New ideas in Psychology **23**, 115 (2005).

[18] M. van der Plaat, *De invloed van coaching stijl op explorerend gedrag en kennistoename bij kleuters in een museumsetting en de rol die nieuwsgierigheid hierbij speelt,* (2014).

[19] R. K. Franse, T. J. van Schijndel, T. I. Plankman, and M. E. Raijmakers, *Families' experiments and conversations at an open-ended exhibit in a science museum: Individual characteristics and the influence of minimal guidance strategies,* Science Education **105**, 707 (2021).

[20] T. J. Van Schijndel, R. K. Franse, and M. E. Raijmakers, *The exploratory behavior scale: Assessing young visitors' hands-on behavior in science museums,* Science Education **94**, 794 (2010).

[21] T. J. van Schijndel, I. Visser, B. M. van Bers, and M. E. Raijmakers, *Preschoolers perform more informative experiments after observing theory-violating evidence,* Journal of experimental child psychology **131**, 104 (2015).

[22] F. Wolbert, *Exploratief spel bij kleuters: Een onderzoek naar de invloed van verbale motivatie op de mate van explorerend gedrag,* (2014).

[23] L. Schulz and E. Bonawitz, *Serious fun: Preschoolers engage in more exploratory play when evidence is confounded,* Developmental psychology **43**, 1045 (2007).

[24] C. Breazeal, K. Dautenhahn, and T. Kanda, *Social robotics,* Springer handbook of robotics , 1935 (2016).

[25] R. Peters, J. Broekens, and M. A. Neerincx, *Robots educate in style: The effect of context and non-verbal behaviour on children's perceptions of warmth and competence,* in *Proc. 26th Int. Symp. on Robot and Human Interactive Commun.* (IEEE, Lisbon, Portugal, 2017) pp. 449–455.

[26] R. Peters, J. Broekens, K. Li, and M. A. Neerincx, *Robots expressing dominance: Effects of behaviours and modulation,* in *8th International Conference on Affective Computing and Intelligent Interaction (ACII)* (IEEE, Cambridge, UK, 2019) pp. 1–7.

[27] R. Peters, J. Broekens, K. Li, and M. A. Neerincx, *Robot dominance expression through parameter-based behaviour modulation,* in *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents (IVA)* (ACM, 2019) pp. 224–226.

[28] M. Neerincx, W. Vught, O. Blanson Henkemans, E. Oleari, J. Broekens, R. Peters, F. Kaptein, Y. Demiris, B. Kiefer, D. Fumagalli, *et al.*, *Socio-cognitive engineering of a robotic partner for child's diabetes self-management,* Frontiers in Robotics and AI **6**, 118 (2019).

[29] F. Kaptein, B. Kiefer, A. Cully, O. Celiktutan, B. Bierman, R. Rijgersberg-Peters, J. Broekens, W. Van Vught, M. Van Bekkum, Y. Demiris, *et al.*, *A cloud-based robot system for long-term interaction: Principles, implementation, lessons learned,* ACM Transactions on Human-Robot Interaction (THRI) **11**, 1 (2021).

[30] R. Allain, *Do heavier objects really fall faster?* https://www.wired.com/2013/10/do-heavier-objects-really-fall-faster/, accessed: 2023-01-01.

[31] R. S. Siegler, *The origins of scientific reasoning,* Children's thinking: What develops , 109 (1978).

[32] D. Klahr and M. Nigam, *The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning,* Psychological science **15**, 661 (2004).

[33] J. van der Graaf, E. Segers, and L. Verhoeven, *Scientific reasoning abilities in kindergarten: dynamic assessment of the control of variables strategy,* Instructional Science **43**, 381 (2015).

[34] J. P. Gutwill, *Challenging a common assumption of hands-on exhibits: How counter-intuitive phenomena can undermine inquiry,* Journal of Museum Education **33**, 187 (2008).

[35] C. Cook, N. D. Goodman, and L. E. Schulz, *Where science starts: Spontaneous experiments in preschoolers' exploratory play,* Cognition **120**, 341 (2011).

Part II

Personalised Content

In this part we will present three studies that investigate how to personalise learning content through the development of an ontology of learning objectives and the dashboards visualising these objectives to health care professionals and children. Within our scenario of the PAL project, we focus on the domain of diabetes self-management training for children aged 7–14. We ground our work in theories from ontology engineering and pedagogy which will be discussed below.

Background

Effective Learning and Learning Goals

Clear and concise learning goals play an important role in student motivation: they direct activities, guide assessment, and provide meaning. Knowing why things must be learned improves motivation and effectiveness of learning [1]. A learning goal must include a description of what the learner should be able to do, under what conditions and how well it should be done. These are encoded in the properties description and proficiency level (e.g., novice) and skill level [2].

Current learning goals must fit with the learner's prior knowledge to facilitate effective learning. One way of structuring past knowledge is by Bloom's skill levels, see below. Further, a distinction is made between declarative knowledge (facts, what a learner should know) and procedural knowledge (processes, what a learner should be able to do) [3]. The latter includes the following items: showing how it is done, doing by the learner, and internalisation. Procedural knowledge is often preceded by declarative knowledge.

Motivation (for learning) was suggested to benefit from goal setting and feedback on goal attainment (e.g., [4]), as well as tailoring of the educational path to the learner's prior knowledge. It is motivating for the learner to be aware of the goals [5] and to be involved in goal setting [6]. Self-regulated learning systems can exploit this by focusing on goal-oriented learning. A personalised guidance approach in ITSs provides learners with the most appropriate learning content accounting for individual learner characteristics such as current knowledge, interests and motivation [7]. This supports application of scaffolding, an approach where the individual learner is constantly challenged to perform slightly above their achieved level of proficiency. This approach is based on Vygotsky's [8] theory on the Zone of Proximal Development (ZPD). The ZPD indicates what a learner can do with guidance and lies between what a learner can do without help and what he/she can't do. Providing learners with experiences within their ZPD are expected to encourage and advance their individual learning [9, 10].

Ontology Engineering

An ontology is a representation of a domain through a formal specification of concepts and their relations and limitations [11, 12]. An ontology makes concepts explicit with specification of necessary attributes and relationships creating a shared common understanding of a domain [13]. In education this means a network of semantically related

Parts of this chapter have been published as **R. Rijgersberg-Peters, W. van Vught, J. Broekens, M.A. Neerincx**, *Goal Ontology for Personalised Learning and its Implementation in Child's Health Self-Management Support*, Published in IEEE Transactions on Learning Technologies.

learning objects within a specific instructional domain. For our applications, a hierarchy of concepts with a set of properties and relations is defined in the Web Ontology Language OWL.

II

OWL is a logic-based knowledge representation that is understandable for humans and interpretable by computer programmes, including:

- **individuals** instances of classes and properties. For example, individual *Emma* is instance of class `Child`.
- **classes** set of individuals with some common properties. For example, *Emma* and *Noah* are both children and thus belong to the class `Child`.
- **attributes** characteristics that individuals can have. For example, *Emma om:hasAge 8*, and *Emma hasSibling Noah*.
- **relations** inter-relations that classes and individuals can have. For example, `Child` is a *subClassOf* `Actor`. Subclasses allow for hierarchical arrangements from which the computer program can infer that *Emma* is an `Actor` (because she is a child and a child is an actor).

An ontology is defined by a set of so-called RDF triples where each triple contains a subject, object and predicate (i.e., a property or relationship that can be attributed to the entity or concept). The object of a property can be either an individual or a certain datatype. All classes, properties and individuals are uniquely identified with a URI. Ontologies can be constructed in Protégé (https://protege.stanford.edu/).

There are different types of ontologies. In education most common are the upper- and domain ontology. Further, ontologies exist for learner data, curriculum and services [11]. An upper ontology presents the formal semantics: the structure of a knowledge base. A domain ontology represents the factual information: the knowledge base that represents facts about the world that a machine can reason about and use rules and other forms of logic to deduce new facts.

Current Practises in Diabetes Education

Type 1 Diabetes Mellitus (T1DM) is a high-impact digestive disease which requires daily self-management and is diagnosed in a growing number of children. To improve well-being and avoid complications, long-term behaviour change is necessary [14]. Current training practices are informal, personalised and aimed at optimising child/patient autonomy and intrinsic motivation. Learning objectives are personal and change while ageing. Therefore, self-management education is highly personalised and directed by challenges faced in daily life and aimed at gradual development of attitudes, knowledge and skills needed for autonomous self-management. Formalisation of the learning process is limited to annual checklists such as the EADV[1] “weet- & doe-doelen” [knowledge- and skill goals].

[1] the Dutch organisation for diabetes nurses http://www.eadv.nl

Research Questions

The main research question for part II is: *How to personalise learning content based on personal learning objectives?*

We are interested to develop a domain independent structure for learning objectives and content, and a domain specific knowledge base as well as user interfaces to present them to various user groups. We derive three research questions, addressed in the chapters of part II.

Research question 1; Chapter 6
How to model learning objectives and implement them in an intelligent tutoring system?

Learner motivation benefits from collaborative goal setting, feedback on goal attainment (e.g., [4–6]), and tailoring of the educational path to the learner's prior knowledge (e.g., [9, 10]). Therefore, intelligent tutoring systems need a model of learning goals for personalisation of educational content, tailoring of the learning path, progress monitoring and adaptive feedback. For the formalisation of learning objectives we developed a domain independent upper ontology to structure learning objectives (goals, achievements and tasks), identified relevant knowledge and skills formalised in the domain ontology that specifies the knowledge base of diabetes self-management training. The developed ontology is machine readable and reasonable. For it to be usable by humans interfaces for different types of users are needed, as well as mechanisms within the system that uses the ontology capabilities such as insight in the learner's current knowledge and capabilities and a mechanism for objective and content selection. To this aim the ontology is implemented in a multi-modal tutoring system consisting of mobile educative games, a health diary, an embodied conversational agent, and a web application for authoring and monitoring. We developed both the interfaces for different user groups, and an algorithm using knowledge about the user for personalised content selection and adaptive interaction.

R. Rijgersberg-Peters, W. van Vught, J. Broekens, M.A. Neerincx, *Goal Ontology for Personalised Learning and its Implementation in Child's Health Self-Management Support*, published in IEEE Transactions on Learning Technologies.

Research question 2; Chapter 7
How to design a user interface for health care providers representing a domain specific knowledge base providing a structure for learning objectives to facilitate personal goal setting and progress monitoring?

Implementing personalised learning objectives needs an interface to present the objectives to the educator and a selection mechanism. In games, ability trees have proven useful and usable to structure and visualise skills (e.g., [15, 16]). Gamification of learning systems is widely applied, the usage of skill trees is an interesting approach in this line. We developed and evaluate a tree-based goal structure in a monitor-and-control dashboard for healthcare providers. We propose guidelines for the interface of collaborative goal setting using a skill tree.

R. Peters, J. Broekens, and M.A. Neerincx, *Guidelines for Tree-based Collaborative Goal Setting*, Proceedings of the 22nd International Conference on Intelligent User Interfaces, (2017).

Research question 3; Chapter 8
How to design a user interface for children that integrates educational objectives with gamification features to advance the interests and engagement of children?

Implementing personalised learning objectives needs an interface to present the objectives and related learning content to the learner as well as provide feedback on progression. By linking in-app activities to relevant educational objectives we aim to increase intrinsic motivation and thereby usage of the learning system. We developed and evaluate a mobile app for children delivering relevant learning content based on personal learning objectives with feedback provided through both a dashboard and an socially interactive agent.

R. Peters, E. Oleari, F. Sardu, and M.A. Neerincx, *Usability of the PAL Objectives Dashboard for Children's Diabetes Self-Management Education*, Proceedings of the 5th International Conference on e-Society, e-Learning and e-Technologies, (2019).

II

REFERENCES

[1] S. Ebbens and S. Ettekoven, *Effectief leren* (Noordhof uitgevers BV, 2005).

[2] R. F. Mager, *Preparing instructional objectives. 2d. rev. ed,* (1984).

[3] R. J. Marzano, D. Pickering, D. E. Arredondo, G. J. Blackburn, R. S. Brandt, C. A. Moffett, *et al.*, *Dimensions of learning* (Association for Supervision and Curriculum Development Alexandria, VA, 1992).

[4] J. F. Bryan and E. A. Locke, *Goal setting as a means of increasing motivation.* Journal of applied psychology **51**, 274 (1967).

[5] P. Black and D. Wiliam, *Assessment and Classroom Learning,* Assessment in Education: Principles, Policy & Practice **5**, 7 (1998).

[6] R. Kleinrahm, F. Keller, K. Lutz, M. Kölch, and J. Fegert, *Assessing change in the behavior of children and adolescents in youth welfare institutions using goal attainment scaling,* Child and Adolescent Psychiatry and Mental Health **7** (2013).

[7] Brusilovsky, *Domain modeling for personalized guidance,* in *Design Recommendations for Intelligent Tutoring Systems: Volume 4 Domain Modeling,* edited by R. A. Sottilare, A. C. Graesser, X. Hu, A. Olney, B. Nye, and A. M. Sinatra (US Army Research Laboratory, Orlando, Florida, 2016) Chap. 15, pp. 165–2184.

[8] L. S. Vygotsky, *Mind in society: The development of higher psychological processes* (Harvard university press, 1980).

[9] L. Berk and A. Winsler, *Vygotsky: His life and works" and" vygotsky's approach to development,* Scaffolding children's learning: Vygotsky and early childhood learning , 25 (1995).

[10] B. R. Schadenberg, M. A. Neerincx, F. Cnossen, and R. Looije, *Personalising game difficulty to keep children motivated to play with a social robot: A bayesian approach,* Cognitive systems research **43**, 222 (2017).

[11] K. Stancin, P. Poscic, and D. Jaksic, *Ontologies in education–state of the art,* Education and Information Technologies **25**, 5301 (2020).

[12] S. Wang, *Ontology of learning objects repository for pedagogical knowledge sharing,* Interdisciplinary Journal of E-Learning and Learning Objects **4**, 1 (2008).

[13] S. Staab and R. Studer, *Handbook on ontologies* (Springer Science & Business Media, 2010).

[14] D. Freeborn, T. Dyches, S. O. Roper, and B. Mandleco, *Identifying challenges of living with type 1 diabetes: child and youth perspectives,* Journal of clinical nursing **22**, 1890 (2013).

[15] Y. Atif, R. Benlamri, and J. Berri, *Learning Objects Based Framework for Self-Adaptive Learning,* Education and Information Technologies **8**, 345 (2003).

[16] A. N. Viet and D. H. Si, *ACGs: Adaptive Course Generation System - An Efficient Approach to Build E-learning Course,* in *Proc. CIT'06* (2006) pp. 259–265.

II

6

Goal Ontology for Personalised Learning and its Implementation in Child's Health Self-Management Support

Intelligent tutoring systems need a model of learning goals for personalisation of educational content, tailoring of the learning path, progress monitoring and adaptive feedback. This paper presents such a model and corresponding interaction designs for the coaches and learners (respectively, a monitor-and-control dashboard and mobile app with supportive communications through a virtual agent), all deployed and tested in a system for child diabetes self-management training. We developed a domain independent upper ontology to structure learning goals and related concepts (such as achievements and tasks), and a domain ontology that specifies the knowledge-base (for in our case, diabetes self-management training). With this approach we relate knowledge elements (e.g., skill) to educational tasks and to learners' knowledge development (e.g., achievements). The ontology was implemented in a multi-modal tutoring system consisting of mobile educative games, a health diary, an embodied conversational agent (ECA), and a web application for authoring and monitoring. We show that our model provides a coherent and concise foundation for: a) formalisation of learning in the diabetes self-management domain, but also for other domains such as mathematics; b) personal goal-setting and thereby personalisation of the educational process including ECA's guidance; and c) creating awareness of progress on the personal educational path. We found that a motivational tutoring system requires a rich set of learning activities and accompanying materials of which a subset

This chapter has been published in IEEE Transaction on Learning Technologies, (2023) [1]

is offered to the learner based on personal relevance. The implemented model proved to accommodate the personal, agent-guided learning paths of children with diabetes, under different treatments from hospitals in Italy and the Netherlands.

6.1. Introduction

New media technologies such as Intelligent Tutoring Systems (ITSs) offer opportunities to advance educational processes and guidance by direct and personalised instructions and feedback. These systems support student independence (of classroom and teacher) and improve learning gain. Specifically extracurricular, self-regulated learning may benefit from this, for example, in eHealth applications designed to support behaviour change into a healthy life-style by increasing knowledge and control over ones well-being. Success rates depend on the user's (intrinsic) motivation for system usage and adherence to the educational program [2]. Embodied conversational agents (i.e., virtual or physcial robots) may stimulate the motivation. Here, the challenge is to overcome the so-called *novelty effect* of educational robots: Humans initially show high motivation with general interest in the new interactive technology, but this interest can wear off quickly after repeated interactions [3].

Motivation (for learning) was suggested to benefit from goal setting and feedback on goal attainment (e.g., [4]), as well as tailoring of the educational path to the learners prior knowledge. It is motivating for the learners to be aware of the goals [5] and to be involved in goal-setting [6]. Self-regulated learning systems can exploit this by focusing on goal-oriented learning. A personalised guidance approach in ITSs provides learners with the most appropriate learning content accounting for individual learner characteristics such as current knowledge, interests and motivation [7]. This supports application of scaffolding, an approach where the individual learner is constantly challenged to perform slightly above their achieved level of proficiency. This approach is based on Vygotsky's [8] theory on the Zone of Proximal Development (ZPD). The ZPD indicates what a learner can do with guidance and lies between what a learner can do without help and what he/she can't do. Providing learners with experiences within their ZPD are expected to encourage and advance their individual learning [9, 10].

The goal of this paper is to advance learner's self-education, supporting and augmenting educators' guidance by construction of reusable learning objectives and improvement of automated selection thereof. To realise this we needed: 1) the formalisation of learning objectives and 2) the representation of this knowledge in a way software technology (agents) can reason about objectives. For the first challenge we need to identify relevant items of knowledge and skills within a specified domain. Further we need insight in the learner's current knowledge and capabilities and a mechanism for personal goal- and task-selection. Hence we define the following research questions:

Research Question 1: *How to model learning objectives?*

Research Question 2: *How to deploy the model in a learning environment?*

We use domain modelling to support personalised guidance by formalising the domain knowledge elements, the knowledge content, and the user's current knowledge level. Further we developed interfaces for various user groups to author and/or visualise this information, and we formalised reasoning about this information; we implemented algorithms using this information for personalised content selection and adaptive interaction. In the remainder of this paper we will discuss these elements and user evaluation thereof. We conclude with recommendations and future work.

6.1.1. PAL CASE STUDY

The PAL project[1] developed a multi-modal application supporting children (7-14 y.o.) acquiring knowledge, skills, and attitudes required for diabetes self-management. The PAL system is composed of an embodied conversational agent (robot and its avatar), an extendable set of (mobile) health applications and dashboards for caregivers which all connect to a common knowledge-base and reasoning mechanism. Within the project we constructed an integrated ontology, *PALO*, the PAL Ontology[2] serving as a common language to support flexible, normative behaviour, establish mutual understanding in (child-agent) interactions, and integrate and use existing knowledge from various domains. The PALO is assembled from several independently developed, but inter-related, ontologies expressed in OWL [11]. Neerincx et al. provide an overview of the PAL development and evaluation [12]. This paper presents the *PAL Objectives Model* ontology, a sub-ontology of *PALO* that defines the knowledge, skills and attitudes required for diabetes self-management as well as it's implementation in the personalised learning support of the PAL system.

6

6.2. BACKGROUND

We aim to construct clear, consistent learning objectives that both require and facilitate shared common understanding of the meaning of these objectives. Therefore we look into how learning goals are formulated in pedagogy and ontologies for education that formalise learning objectives. An ontology allows for making concepts explicit with specification of necessary attributes and relationships [13]. Lastly, we present some necessary background on the domain of diabetes self-management learning.

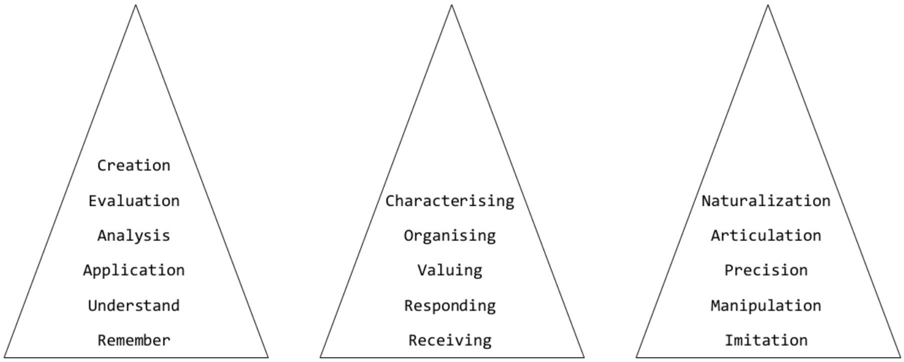
6.2.1. EFFECTIVE LEARNING AND LEARNING GOALS

Clear and concise learning goals play an important role in student motivation; it directs activities and guides assessment, and gives meaning. Knowing why things must be learned improves motivation and effectiveness of learning [14]. A learning goal must include a description of what the learner should be able to do, under what conditions and how well it should be done. These are encoded in the properties description and proficiency level (i.e. novice etc.) and skill level [15].

Current learning goals must fit with the learner's prior knowledge to facilitate effective learning. One way of structuring past knowledge is by Bloom's skill levels, see be-

[1] Personal Assistant for a healthy Lifestyle, http://www.pal4u.eu

[2] Publicly available upon request, https://confluence.ewi.tudelft.nl/display/PALsCE/PAL+ONTOLOGY



(a) cognitive [20] (b) affective [23] (c) psychomotor[21]

Figure 6.1: Taxonomies for learning goal skill levels.

low. Further, a distinction is made between declarative knowledge (facts, what a learner should know) and procedural knowledge (processes, what a learner should be able to do) [16]. The latter includes the following items: showing how its done, doing by the learner, and internalisation. Procedural knowledge is often preceded by declarative knowledge.

Bloom's Taxonomy of learning goals [17] is the predominant model structuring educational learning objectives into levels of complexity and specificity. There are three submodels covering learning objectives in cognitive (knowledge), affective (emotional) and psychomotor (action) domains (see Figure 6.1). Each domain is composed of categories with a sequential, hierarchical link; learning of the lower levels enables the learning of skills in higher levels. This scaffolding of higher-level skills from lower-level skills is an application of constructivism [18]. There are different versions of the categories (see [19] for an overview). We choose to adhere for the cognitive domain to the revised taxonomy by Krathwohl [20] in favour of the original version by Bloom, Krathwohl et al. [17]. In the psychomotor domain we listed the categorisation by Dave [21] in favour of the version by Simpson [22] because it presents more concrete actions and is thereby more applicable to differentiate levels of skill goals.

6.2.2. Ontologies

Advancements of semantic web for education allow for developments towards collaborative, personalised and self-organised learning in intelligent systems. Where prior hyperlinking is limited to passive access to materials, the semantic web supports automated reasoning about it. Further an ontology facilitates shared common understanding of a domain [13].

Ontology Engineering

An ontology is a representation of a domain through a formal specification of concepts and their relations and limitations [24, 25]. In education this means a network of semantically related learning objects within a specific instructional domain, for our applications, a hierarchy of concepts with a set of properties and relations defined in the Web Ontology Language OWL.

OWL is a logic-based knowledge representation that is understandable for humans and interpretable by computer programmes, including:

individuals instances of classes and properties. For example, individual *Emma* is instance of class `Child`.

classes set of individuals with some common properties. For example, *Emma* and *Noah* are both children and thus belong to the class `Child`.

attributes characteristics that individuals can have. For example, *Emma om:hasAge 8*, and *Emma hasSibling Noah*.

relations inter-relations classes and individuals can have. For example, `Child` is-a *subClassOf* `Actor`. Sub-classes allow for hierarchical arrangements from which the computer programme can infer that *Emma* is an `Actor` (because she is a child and a child is an actor).

An ontology is defined by a set of RDF triples where each triple contains a subject, object and predicate. The object of a property can be either an individual or a certain datatype. All classes, properties and individuals are uniquely identified with a URI. Ontologies can be described in Protégé (https://protege.stanford.edu/).

There are different types of ontologies. In education most common are the upper- and domain ontology, further, ontologies exist for learner data, curriculum and services [24]. An upper ontology presents the formal semantics; the representation of a knowledge-base. A domain ontology represents the factual information; the knowledge-base that represents facts about the world that a machine can reason about and use rules and other forms of logic to deduce new facts.

ONTOLOGIES FOR EDUCATION

In recent years, the popularity of ontology engineering for the development of e-learning systems increased. In this section we discuss relevant literature on the use of ontologies for learning, focusing on ontologies that include learning goals[3].

In education, domain-dependent and domain-independent ontologies have been developed focusing on, respectively, (a) the definition and (hierachical) structure of knowledge concepts (e.g., for IT-education, and English language and mathematics [26, 27]), and (b) the learning goals.

Ontologies are proposed for various usages. For example, to explain core concepts to instructors and to generate lesson plans [26], or define essential knowledge for an online tutoring program by automated scraping of learning goals from instruction materials [27]. Furthermore, formalisation of existing learning goals may be used for course,

[3] Learning goals are often referenced as 'learning objectives' and learning -materials and -activities as 'learning objects'. For clarity we adhere to the terms goal, task and material as used through out this paper. We use the term learning objective as super-class of the terms learning -goal, -task and achievement.

6

curriculum, and syllabus modelling, as well as automated calculation of learning goal attainment [28], for course alignment of goals, tasks and assessment [29, 30], to optimise discovery of learning materials [25, 31], or to form collaborative learning groups based on goals, triggers, materials and scenarios [32]. We found limited usage of ontologies for personalised learning and/or goal-setting. In one study, a hierarchy between learning goals was used to suggest materials that fit the learner's ability and requirements including personal learning goals [31], whereas David *et al.* [29] developed a 'digital twin' guiding learners based on current learner profile and status. Ontologies are used in recommender systems for learning content based on user interest or needs. These ontologies typically represent knowledge about the learner and the learning content (objects) but do not specify learning objectives. Thereby personalising the content but not necessarily the learning path. For example, [33] facilitates online course selection, and [34] identifies suitable online resources for concepts based on MI profiles. In [35] assessment based content adaptation does support personalising of the learning path by recommending content based on performance. We expect a rise of use of ontologies for personalised learning in the future. Rashid et al. [27] explicitly stated their intention to implement their ontology in a smart tutoring framework.

There appears not to be one single best practise for learning goal modelling. For example, Chung et al. [28] define an identifier, description, cognitive level, attitude level, and skill level and form relations between activities and defined goals, whereas Ng [31] defines title, description, sequencing rules, related tasks, proficiency level, performance indicator, cognitive state, and relations with knowledge reference in another ontology. However, the common denominator seems to be a formalisation of hierarchical relations between learning goals and inclusion of a human-understandable description of the goal and an indicator of depth of knowledge. The latter is referencing often to Bloom skill levels (i.e., in [25, 26, 28, 31], but may also embed 21st century skills (i.e., [30]), or SOLO (i.e., [29]). Further, relations between goals and tasks occur (e.g., [28, 30, 31]), in some cases referencing Bigg's triangle for effective learning (i.e., [29]).

6.2.3. Current Practises in Diabetes Education & PAL

Type 1 Diabetes Mellitus (T1DM) is a high impact digestion disease which requires daily self-management and is diagnosed in a growing number of children. To improve well-being and avoid complications, long-term behaviour change is necessary [36]. Current training practices are informal, personalised and aimed at optimising child/patient autonomy and intrinsic motivation. Learning objectives are personal and change while ageing, therefore, self-management education is highly personalised and directed by challenges faced in daily life and aimed at gradual development of attitudes, knowledge and skills needed for autonomous self-management [37]. Formalisation of the learning process is limited to annual checklists such as the EADV[4] "weet- & doe-doelen" [knowledge- and skill goals].

There is a need for an ontology that underpins the personalised learning process of children via educative games and an embodied conversational agent. Current ontologies show the benefits of a hierarchical structure and the inclusion of core elements of

[4] the Dutch organisation for diabetes nurses http://www.eadv.nl

6

Bloom's hierarchy, but do not provide a complete ontological model yet. The next section provides the development of such a model, elaborating on the research presented.

6.3. (PAL) OBJECTIVES MODEL

We created an ontology formalising learning objectives for self-regulated learning using Protégé (see Section 6.2.2). The ontology is developed in an iterative process. The initial goal ontology was loosely based on the task ontology by Welie [38] and then further developed with experts from education, data science and paediatric healthcare, taking into consideration the usability requirements from the PAL project. Resulting are two separate owl-files for the upper ontology and the domain ontology[5]. The upper ontology (*Objectives Model*) presents the formal semantics; a data-model representation for a knowledge-base. The domain ontology (*PAL Objectives Model*) is the knowledge-base presenting factual information about diabetes self-management education.

The upper ontology, *Objectives Model* (*OM*), is domain independent and describes relevant concepts and relations between them. The *OM* contains three sub-models: A) domain model, B) user model, and c) content model (see Figure 6.2). The three sub-models store different types of data, as was also suggested for models for adaptive educational systems [7]. The domain model (this is *not* the same as the domain ontology) is a domain independent structure of domain knowledge elements, it structures the global knowledge-base consisting of three main concepts `Achievement`, `Learning Goal`, and `Task`. These concepts all are sub-classes of `Objective`. The user model is an over-lay model structuring the storage of personal user data; child specific (progress) data on these concepts and it models the `Child` itself. The content model structures the learning `Material` related to tasks.

The domain ontology, *PAL Objectives Model* (*POM*), instantiates the upper ontology. Domain specific instances are created accordingly to the upper ontology data-model. In the remainder of this section we present the domain independent upper ontology and illustrate this with examples from the domain ontology created for the PAL-project in the domain of diabetes self-management training for children aged 7-14 and starting at diabetes on-set.

6.3.1. DOMAIN MODEL

The domain model structures the formal representation of knowledge about learning objectives in a specific domain, as formalisation of a "concept map" in OWL, with a class structure, machine readable and with automated reasoning possibilities, and consists of three main concepts and three supporting classes as presented below and in Figure 6.3. Learning goal is the primary concept of the knowledge-base structure. Achievements group goals in relevant challenges. Tasks prescribe activities to train new knowledge defined in goals. Supporting constructs are present that provide additional information about mainly the learning goals.

The three main classes (i.e., `LearningGoalType`, `AchievementType`, and `TaskType`) each have two data properties being *`om:label`* and *`om:description`*. The latter is a full sentence description of the objective. For example, "I remember how much carbs stan-

[5] Available upon request at the first author.

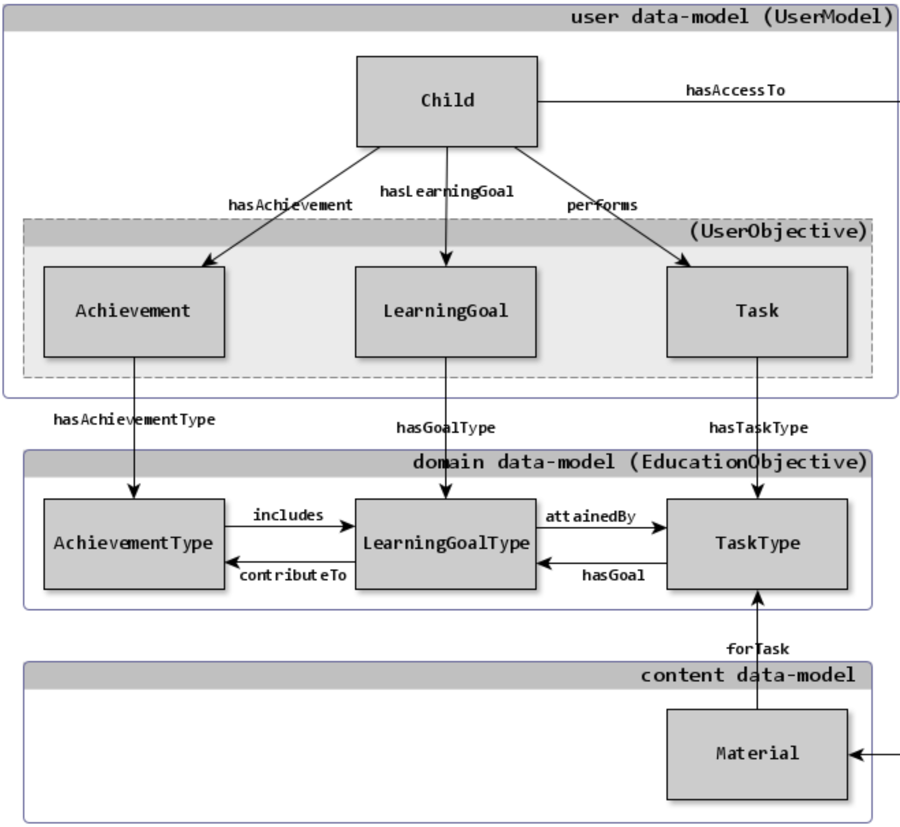


Figure 6.2: Overview of the three sub-models in the upper ontology, the objectives model, with their most important classes and relations.

dard foods contain." The former is a short-hand description in a few words, intended for small GUI elements. For example, 'Sleepover' and 'Amount of CHO'. See for further examples Listing 6.1.

```
%Learning goal
pom:carbs2 om:label "Amount of CHO (standard)"^^xsd:string
pom:carbs2 om:description "I remember how much carbs standard foods
    contain."^^xsd:string
%Achievement
pom:sleepover om:label "Sleepover"^^xsd:string
pom:sleepover om:description "I can go to a sleepover."^^xsd:string
%Task
pom:memoryAH om:label "Memory CHO (hard)"^^xsd:string
pom:memoryAH om:description "Play the 1 player memory game and remember
     the amount of carbs per food. You need a 'hard' score to complete
    this task."^^xsd:string
```

Listing 6.1: Example for objective data-properties label and description

Learning Goal defines the desired end level of knowledge. Individuals from the class `LearningGoalType` form a structured set of domain knowledge elements (DKEs) that collectively span the total domain or a desired sub-part. Each individual present precisely one (1) knowledge element such as 'Amount of CHO' from Listing 6.1. A goal may directly follow-up on one or multiple other (easier/less advanced) goals, and/or precede one or multiple advanced goals. For example, learning to count carbohydrates is preceded by knowing the amount of carbohydrates in various food items. Relations between goals are formalised in object properties *om:requires* and its inverse *om:requiredFor* as follows:

```
pom:carbs2 om:requiredFor pom:carbs3
pom:carbs3 om:requires pom:carbs2
pom:carbs3 om:requiredFor pom:countCarbs3
pom:carbs3 om:requiredFor pom:countCarbsK4
pom:countCarbs3 om:requires pom:carbs3
pom:countCarbsK4 om:requires pom:carbs3
```

Listing 6.2: Example for goal relational properties requires and requiredFor

Goals also hold relationships with other objective types; a goal may be *om:attainedBy* one or more tasks, and *om:contributeTo* one or more achievements. For example, 'Amount of CHO' is trained playing a memory game (*AH* about the *A*mount of carbohydrates in food products on *H*ard difficulty) and two sorting games (*CEH* about the amount of *C*arbohydrates in food and *NEH* about *N*utrition in general in food products, both on *E*asy mode but requiring a *High* score), and belongs to six achievements amongst whom to go to a sleepover and to eat at a restaurant:

6

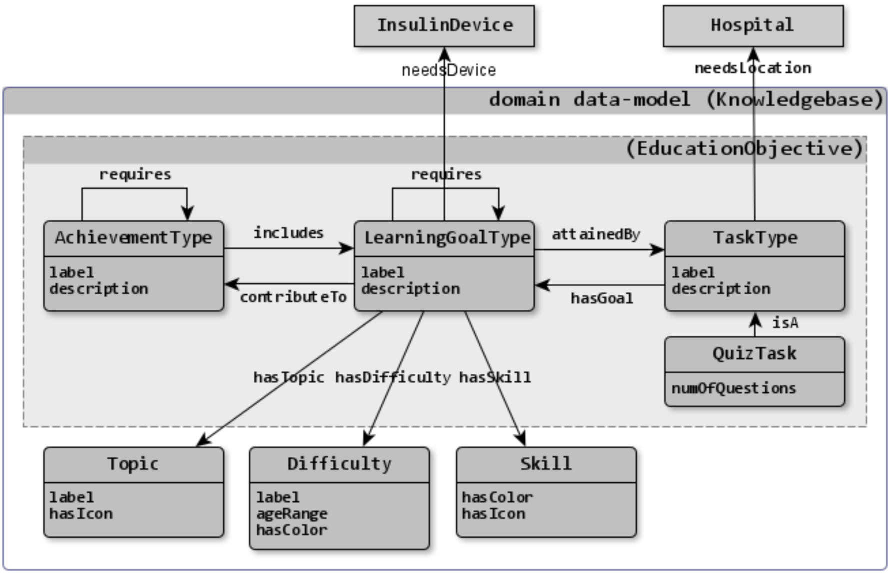


Figure 6.3: Objectives model domain model classes, properties and relations (extended version).

```
pom:carbs2 om:attainedBy pom:memoryAH
pom:carbs2 om:attainedBy pom:sortCEH
pom:carbs2 om:attainedBy pom:sortNEH
pom:carbs2 om:contributeTo pom:sleepover
pom:carbs2 om:contributeTo pom:restaurant
```

Listing 6.3: Example for goal relational properties attainedBy and contributeTo

Lastly, goals have object properties providing information about the subject and educational levels. A generic description of the subject(s) in given by *om:hasTopic*, such as 'Nutrition' and 'Carbs'. The level is given by *om:hasDifficulty* and *om:hasSkill*. Where skill indicates the desired end level of factual knowledge, motor skill or attitude. For example, 'Amount of CHO' has a desired end state of the child being able to appoint the amount of carbohydrate in everyday food products, thus remembering of factual knowledge. Whereas the goal to count carbohydrates over combined food products requires the child to apply knowledge in new situations. Difficulty indicates the proficiency level on a scale [1-6]. Remembering the amount of carbohydrates in various foods is not typically the first thing learned after diabetes onset and is not expected at early age (<8 y.o.), therefore, it has difficulty level 2 (advanced beginner).

```
pom:carbs2 om:hasTopic pom:carbs
pom:carbs2 om:hasTopic pom:nutrition
pom:carbs2 om:hasDifficulty om:difficulty2
pom:carbs2 om:hasSkill bloom:remember
pom:countCarbs3 om:hasSkill bloom:apply
```

6

Listing 6.4: Example for goal object-properties

Additionally, a goal may hold a requirement for usage of a certain physical device. In the pal-case an insulin device either being a pen or pump. For example, a goal related to learning to bolus (a correction of the amount of insulin administer via an insulin pump) is only relevant for children who use a pump for their insulin intake:

```
pom:bolus1 om:needsDevice pom:pump
```

Listing 6.5: Example for goal object-properties needsDevice

Achievement Individuals of the class `AchievementType` form a set of domain knowledge components (KC). Each KC represents a fragment of knowledge within a specific domain (i.e., a relevant challenge). Achievements do *not* define new knowledge, but form a set of related goals on a similar difficulty level. Achievements should be phrased as a main accomplishment that the child is eager to achieve to promote understanding and relevance of the achievements and optimise motivation for learning. Each achievement may, similar to goals, precede or succeed one or multiple other achievement(s). For example, the achievement 'Sleepover' is preceded by an achievement related to basic diabetes knowledge.

```
pom:sleepover om:requires pom:basicDiabetes
pom:basicdiabetes om:requiredFor pom:sleepover
```

Listing 6.6: Example for achievement relational properties

Further, goal relations are formalised by the property `om:includes` (this is the inverse of goal property `om:contributeTo`). For example, the achievement 'Sleepover' at difficulty 2 groups five goals:

```
pom:sleepover om:includes pom:carbs2
pom:sleepover om:includes pom:askHelpKnow2
pom:sleepover om:includes pom:FunStressKnow2
pom:sleepover om:includes pom:InsulinForCarbsKnow2
pom:sleepover om:includes pom:RegSnackKnow2
```

Listing 6.7: Example for achievement relational properties

Achievements do not have other properties; topic(s) and difficulty level(s) are inherited from included goals. For example, 'Sleepover' takes from 'carbs2' the topics 'Nutrition' and 'Carbs' and 'Insulin' from 'insulinForCarbs2'. An achievement (goal set) is created on each difficulty present in included goals. For example, 'Sleepover' includes goals on difficulty level 1-4, subsets are created on each level and the child is required to complete 'Sleepover1' (i.e., attain all included goals) before he/she is able to train the goals included in 'Sleepover2' and so on.

Task Individuals of the class `TaskType` represent a structured set of activities to train goals based on the ITS application(s). Each tasks is related to a specific learning goal, formalised in the object-property `om:hasGoal` (inverse of `om:attainedBy`). Goals can be trained by more than one task, but each task is designed to train one (1) goal.

6

```
pom:memoryAH om:hasGoal pom:carbs2
pom:sortCEH om:hasGoal pom:carbs2
pom:sortNEH om:hasGoal pom:carbs2
```

Listing 6.8: Example for task relational property

There are no requirements such as precessors. Tasks can be location specific (`om:needsLocation`). For example, a video introducing diabetes is directed at Dutch children, while a video on the gluca pen is in Italian:

```
pom:watchVDDB om:needsLocation pom:hospital_2
pom:watchVDDB om:needsLocation pom:hospital_3
pom:watchVGP om:needsLocation pom:hospital_4
```

Listing 6.9: Example for task object-property

Additionally, sub-classes for specific activities within the ITS may have extra properties. In our case, `QuizTask` has a property `om:numOfQuestion` to store the number of questions fitting a specific task.

```
pom:answerMBG3 om:numOfQuestion "12"^^xsd:int
```

Listing 6.10: Example for task sub-class property

Object-properties carry additional information about a learning objective. Information presentation can be more rich compared to data-properties because the object properties can have data-properties themselves. For example, setting an icon for a topic presenting the topic in the user interfaces. Learning goal type has four object-properties: topic, difficulty and skill, and insulin device. Task type has one object-property: hospital. These objects are discussed below.

1. Topic is the most straightforward property indicating the general subject of a learning goal. For example, nutrition, carbohydrates, or insulin (The topics of an achievement are derived from the set of topics of included goals). It carries a label, selected from a pre-defined set, and a reference to a stored icon representing the topic.

```
pom:carbs2 om:hasTopic pom:nutrition
pom:nutrition om:label "Nutrition"^^xsd:string
pom:nutrition om:hasIcon "images/food.png"^^xsd:anyURI
```

Listing 6.11: object-property topic example

2. Difficulty level indicates the general level of expertise of a learning goal on a scale [1-6]. The concept is based on skill-trees used in games and provides a directed hierarchical (from easy to more advanced per subject) data-structure. In the case of learning goals it prescribes the assumed proficiency level of the child on a given topic allowing for personalised scaffolding; objectives are selected through assessment and authoring, constructing an individual learning path. Each goal is assigned a difficulty level from a fixed set (i.e., rdf individuals: novice, advanced beginner, competent, proficient, expert, master). Each difficulty has a label and colour for presentation in the user interfaces, and a suggested age range (see Listing 6.12). The age range indicates the approximate age group of children for whom a specified goal is likely suitable. For example, a young child is expected to have goals on the first or second level rather on the final, depending on their diabetes onset, and developmental level. Important is that difficulty is not an age restriction, meaning that a child can progress on different difficulty levels depending on topic.


```
pom:carbs2 om:hasDifficulty om:difficulty2
om:difficulty2 om:label "Advanced beginner"^^xsd:string
om:difficulty2 om:hasColor "ff7f50"^^xsd:string
om:difficulty2 om:ageRange "7-9"^^xsd:string
```

Listing 6.12: object-property difficulty example

3. Skill level specifies the complexity and specificity of each learning goal based on Bloom's taxonomy of learning objectives (see Section 6.2.1). Skill is used to organise preceding and succeeding goals (within or over difficulty levels); a goal aimed to apply a specific piece of knowledge is typically learned only after respectively understanding and remembering this knowledge. Each goal is assigned a skill level from a fixed set of 16 levels (over three classes, i.e., knowledge, attitude and psychomotor; see for an example Listing 6.13, line 1-3) defined in the bloom-ontology. OM imports this ontology of skill levels based on the taxonomy of Bloom [17], where `om:Skill ≡ bloom:LearningObjective`. The current (diabetes) domain knowledge base includes mainly cognitive goals to remember or understand, and goals to execute a physical action requiring application of cognitive skills. The chosen implementation allows for extension and makes this limitation transparent.

```
1 pom:carbs2 om:hasSkill bloom:remember
2 bloom:remember rdf:type bloom:Remembering
3 bloom:Remembering rdf:subClassOf bloom:Knowledge
```

```
bloom:remember bloom:hasColor "2d3092"^^xsd:string
bloom:remember bloom:hasIcon "images/brain-blue.png"^^xsd:anyURI
```

Listing 6.13: object-property skill example

4. Device specifies any technological artefact involved in a learning goal. In the PAL case an insulin device required to complete either an injection or bolus (see Listing 6.14, line 1-2), where `pom:InsulinDevice` is a subclass of `om:Device`, and insulin device is used to personalise learning goals by relevance. `Device` has data-properties *om:label* and *om:hasIcon* for interface presentation.

```
1 pom:bolus1 om:needsDevice pom:pump
2 pom:injectInsulinHelp2 om:needsDevice pom:pen
3 pom:pump om:label "Pump"^^xsd:string
4 pom:pump om:hasIcon "f566"^^xsd:string
5 pom:pen om:label "Pen"^^xsd:string
6 pom:pen om:hasIcon "<i class='fas fa-syringe'></i>"^^xsd:string
```

Listing 6.14: object-property device example

5. Location specifies a physical place where the child is located at or related to. In the PAL case the location refers to the hospital where the child is a patient (see Listing 6.15), where `pom:Hospital` is a subclass of `om:Location`, and the hospital location is used to personalise learning content in language and facilitate treatment plan differences between hospitals. For example, whereas the participating Italian hospital preferred dietary restrictions, the Dutch hospitals prefer insulin corrections. `Location` has a *om:label* and *om:address*.

```
pom:watchVDDB om:needsLocation pom:hospital_2
pom:hospital_2 om:address "Ede, Netherland"^^xsd:string
pom:hospital_2 om:label "Gelderse Valei"^^xsd:string
```

Listing 6.15: object-property device example

6.3.2. USER MODEL

The user model is an overlay-model representing the current state of domain knowledge of a particular learner. In other words, the user's individual objectives with for each objective the progress value and a reference to an domain specific knowledge element (i.e., an *Type, the relation between goals and achievements, and goals and tasks is inherited from this type and not part of the child's objectives). The key principle of the overlay model is that for each objective in the domain model, an individual user knowledge model stores the progress of the user on this objective. This model functions as an internal memory for the system to know what tasks the user did or didn't do and what knowledge is or isn't attained.

The user model (Figure 6.4) includes the `Child` with some datatype properties that describe the child (e.g., name, birth date and gender), and some object properties useful for personalised goal-setting. Further, there are three object properties relating to three objectives classes within the user model (i.e., `LearningGoal`, `Achievement`, and `LearningTask`).

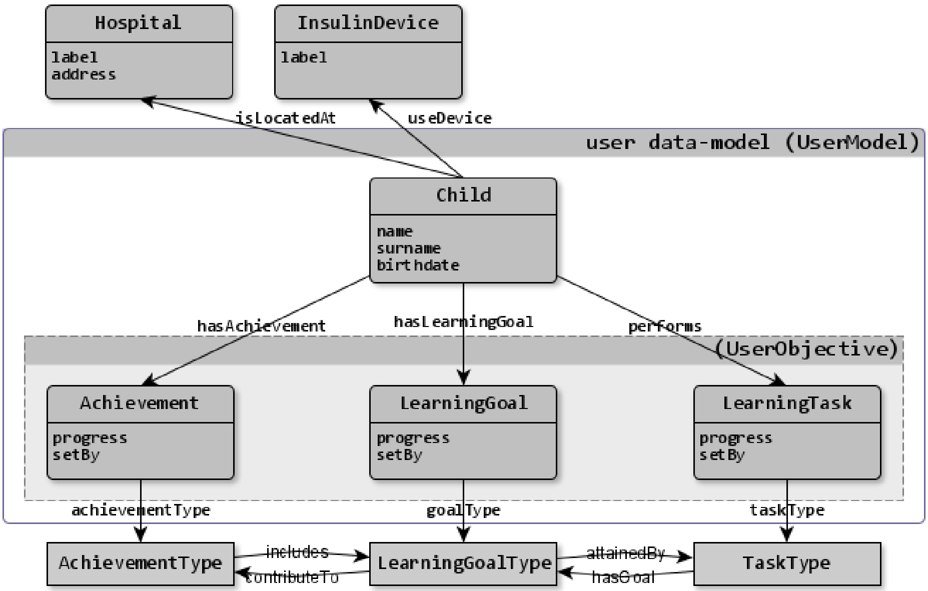


Figure 6.4: Objectives model upper user model classes, properties and relations.

Child defines a specific child. An individual of the class `Child` presents the characteristics of the child used by the system, including progress on each of the objectives. `Child` is defined in the PALO domain ontology (dom), individuals are given the prefix pal. *OM* uses datatype properties that are crucial for personalisation of the educational path such as age (derived from birth date) to select a starting point for suggested skill level and, for the diabetes domain relevant property, gender to in- or exclude specific goals such as those related to period. Further, *OM* uses object properties *`dom:useDevice`* and *`dom:isLocatedAt`* for personalisation. See Listing 6.16. This data is added via the PALControl application in consultation with the healthcare professional. Other characteristics defined in the PALO domain ontology such as relatives and preferences for colour are omitted because they are irrelevant for personalised learning.

```
pal:child1 rdf:type dom:Child
pal:child1 dom:hasName "Henk"^^xsd:string
pal:child1 dom:hasSurname "Jansen"^^xsd:string
pal:child1 dom:birthdate "2006-04-12"^^xsd:date
pal:child1 dom:hasGender dom:male
pal:child1 dom:useDevice dom:pen
pal:child1 dom:isLocatedAt pal:hospital_2
```

Listing 6.16: data and object properties that define a child and may be used for personalisation example

The *OM* user model extends the class `Child` from PALO with three objects properties (*`om:hasLearningGoal`*, *`om:performs`*, and *`om:hasAchievement`*) to store child specific progress data on learning goals, -tasks, and achievements (Listing 6.17).

```
pal:child1 om:hasLearningGoal pom:henksgoal1
pal:child1 om:hasLearningGoal pom:henksgoal2
pal:child1 om:hasAchievement pom:achievement283
pal:child1 om:performs pom:answerI1henk
```

Listing 6.17: child data and object properties example

User Objectives are included in the user model as well (i.e., `Achievement`, `LearningGoal`, and `LearningTask`). They are related to the child by the object properties presented above. A child can have this personal objectives for each of the DKE's in the knowledge base. Each of the three classes has the data properties *om:progress* and *om:setBy* which respectively store the child's percentage of attainment of the objective and who last updated the progress value (see Listing 6.18). The latter can have values either 'system' or 'user', where system means the child-system interaction resulted in progress update and user means the value was last manually update via an authoring tool.

```
pom:henksgoal1 om:hasProgress "0.8"^^xsd:float
pom:henksgoal1 om:setBy "system"^^xsd:string
pom:insulingoalhenk om:hasProgress "0.0"^^xsd:float
pom:insulingoalhenk om:setBy "user"^^xsd:string
```

Listing 6.18: data properties of objectives example

Further, each class has there own object property relating them to the specific knowledge elements, individuals in the domain ontology (see Listing 6.19).

```
pom:henksgoal1 om:goalType pom:carbs2
pom:answerI1henk om:taskType pom:answerI1
pom:achievement283 om:achievementType pom:sleepover
```

Listing 6.19: object properties of objectives example

6.3.3. CONTENT MODEL

The content model includes a concept-based organisation of educational materials such as the actual videos and questions for quizzes. Individuals of the class `Material` list all relevant materials and connect specific content with one or more learning goals via tasks with object property *om:forTask* (see Listing 6.20).

```
pom:video1 rdf:type om:Material
pom:video1 om:forTask pom:WatchVVDB
pom:WatchVVDB om:hasGoal pom:BasicPrinciplesKnow2
```

Listing 6.20: materials related to learning goals example

`Material` has two data properties: *om:medium* and *om:link*. The first defines the sort of material and can be any of 'web', 'video' or 'app', with app meaning that the material is included in the ITS system and web or video are linking to any source on the world wide web. Dependent on the value for *om:medium*, *om:link* is either an external hyperlink or in-app location as shown in Listing 6.21.

```
pom:video1 om:medium "video"^^xsd:string
pom:video1 om:link "https://www.youtube.com/watch?v=K4UsrTNz37I"^^xsd:
    anyURI
```

```
pom:answerI1 om:medium "app"^^xsd:string
pom:answerI1 om:link "QUIZ$I1"^^xsd:anyURI
pom:BS_PlayNEH om:medium "app"^^xsd:string
pom:BS_PlayNEH om:link pom:"GAME$BREAKSORT$EASY$NUTRIENTS"^^xsd:anyURI
```

Listing 6.21: material datatype properties example

6.4. CASE: Diabetes Self-management Education

The *Objectives Model* of the previous Section 6.3, provides the foundation of the personalised self-regulated learning support for children with diabetes that has been developed in the PAL-project. This project has followed a human-centred iterative development process, in which the upper ontology was instantiated in a *PAL Objectives Model* (see Listings in Section 6.3) for diabetes self-management in the Netherlands and Italy. Both *POM* and *OM* have been integrated in the so-called PALO ontology — a set of ontologies specific to the PAL system. In the remainder of this section we describe the domain knowledge acquisition & modelling, the domain ontology development, and the deployment of this ontology within the PAL system.

6.4.1. Content Creation

Acquiring the domain knowledge and deriving a coherent and concise set of learning objectives out of it, consisted of complementary activities. First, we formalised diabetes self-management training goals based on an existing list of guidelines of knowledge and skills. Secondly, we organised expert sessions to introduce the conceptualisation of the knowledge structure in the upper ontology and to collaboratively construct individuals for achievements and tasks. Then, we developed and provided an authoring tool to facilitate experts in creating their own goals, achievements and tasks for diabetes self-management education, in a format that matches with the overall ontology structure.

Formalisation of existing guidelines

To support personalised learning of diabetes self-management skills we needed to know the relevant domain knowledge elements, the education paths towards mastering these, and a structured way to included the DKEs in an ITS. We constructed a first set of DKEs from a topic-based list of knowledge and skill goals (Weet- & doe-doelen, see Section 6.2.3). We formalised the goals in *PAL Objectives Model* as `LearningGoalType`, added the *`om:description`* and the Bloom level (*`om:hasSkill`*) amongst the other properties listed in 6.3. Further, we made explicit the underlying hierarchical structure in *`om:requires`*. We added achievements in `AchievementType` as substantial goals that can only be attained by first attaining a set of smaller goals. A first version of the 'goal-tree' was implemented in OWL/RDF and integrated in the PALO ontology which was connected to the PAL system [37, 39, 40]

Expert sessions

To improve the content of the domain model (*PAL Objectives Model*) we consulted domain experts. The aim of these session was to get them acquainted with the model and to elicit content. Paediatricians specialised in diabetes type 1 were introduced the upper ontology (*Objectives Model*). In this phase a doctor and nurses from two Dutch hospitals were involved. In brainstorm sessions domain specific achievements, goals and tasks were gathered. The health care professionals that participated in these sessions were already involved in previous iterations of the PAL project and had some experience on describing achievements, goals and tasks on diabetes self-management for children.

The process of these sessions was as follows: The researcher would suggest a topic, such as nutrition. The health care professionals were asked to think of as many learning goals as possible. Next, these goals would be grouped by age-range (the age a child is likely to be able to comprehend the knowledge or skill). Within each age range clusters were formed and given a description; an achievement was defined. Then experts were asked to think of possible activities to train each knowledge element. Figure 6.5b shows the resulting objectives structure for the topic of nutrition.

(a) session in progress

(b) result for one topic

Figure 6.5: Expert session for domain content creation.

COLLABORATIVE CONSTRUCTION OF ONTOLOGIES

Subsequently we had several intensive sessions with experts from the two Dutch hospitals as well as the Dutch diabetes association, followed by sessions with experts from an Italian hospital to check and refine the model and documentation. The PAL project used the Wiki Socio-Cognitive Engineering (WiSCE) tool [6] for the project team to collaborate and share knowledge efficiently. It is built and maintained in Atlassian Confluence [7], providing guidance and structure for the different stakeholders involved in the Research & Development process. To make the knowledge of health care professionals (HCPs) explicit and to facilitate authoring of the domain model by HCPs, WiSCE provided shared templates in which they could add and edit the content directly (see Figure 6.6). The templates facilitated joined content creation and refinement, assuring compliance with the ontology structure (consistency and completeness).

[6] https://scetool.ewi.tudelft.nl

[7] https://www.atlassian.com/software/confluence

Goal: Interpret a blood glucose measurement

<table>
<tr><td>ID</td><td>IM1.1, IM2.1, IM2.1, IM3.1, IM3.2, IM4.1, IM4.2</td></tr>
<tr><td>Description</td><td>The child needs to learn how to interpret the blood glucose measurement.

The child needs to know what acceptable values are (skill level 1), need to understand the effect of a value (skill level 2), the child needs to learn how to interpret the values with help (skill level 3) and eventually needs to learn how to interpret the values on its own (skill level 4).

I can interpret a blood glucose measurement with help - I can interpret a blood glucose measurement myself</td></tr>
<tr><td>Skill</td><td>
<table>
<tr><th>Development level</th><th>Skill level</th><th>Tasks</th></tr>
<tr><td>1</td><td>I know

I know which glycemic values are fine for me.</td><td>QUIZ (QIM1)

B&S (BIM1)</td></tr>
<tr><td>2</td><td>I know

I know the which action to perform for each glycemic value.</td><td>QUIZ (QIM2)

MEMORY(MIM2)</td></tr>
<tr><td>2</td><td>I can do myself

I can detect when an action is needed based on my glycemic value myself.</td><td>REAL WORLD TASK (RIM2)</td></tr>
<tr><td>3</td><td>I understand

I understand which action action to perform for each glycemic value.</td><td>QUIZ (QIM3)

MEMORY (MIM3)</td></tr>
<tr><td>3</td><td>I can do with help

I can perform the correct action to improve my glycemic value with help.</td><td>TIMELINE TASK (TLG3)</td></tr>
<tr><td>4</td><td>I know

I know what to take into account while reasoning about my glycemic value.</td><td>QUIZ (QIM4)</td></tr>
<tr><td>4</td><td>I can do myself

I can perform the correct action to improve my glycemic value myself.</td><td>REAL WORLD TASK (RIM4)</td></tr>
</table>
</td></tr>
<tr><td>Achievements</td><td>Swimming, Sleepover, Go on Holiday, School camp, Being ill, Sports game, Sports training, Cycle to school, Go on Holiday, Dinner in restaurant, Roller coaster, Lazy weekend</td></tr>
<tr><td>Topic(s)</td><td>Glucose measurement</td></tr>
<tr><td>Dependencies</td><td>The skills in this goal are only dependent on each other in terms of development level (i.e. one should posses all skill on a lower development level before it is able to achieve a skill on a higher development level)</td></tr>
</table>

Figure 6.6: WiSCE's Confluence template example filled for a specific goal.

The information gathering process was as follows: The HCPs were asked to think about achievements that are relevant in the daily lives of children with type 1 diabetes. After that, we went through each achievement asking the HCPs which knowledge and skills a child should have to attain this achievement and attach a corresponding development level per skill. Gathered content was recorded in the Confluence templates and used in later sessions for refinement and definition of new content. When no new achievements and goals were raised, the sessions were devoted on selecting appropriate (learning) tasks for each goal per skill level. These were also documenting in Confluence. The resulting content was reviewed by other health care professionals.

While linking goals and achievements, we experienced that most achievements require a similar subset of goals representing basic principles of diabetes self-management. Therefore, a new achievement type was created holding the basic principles for each development level. Children are required to complete these basic principle achievements as a prerequisite for other achievements on that developmental level. This way an appropriate foundation on disease self-management knowledge can be assured. Furthermore, the construction of achievements centred around daily life challenges resulted in multiple achievements with similar type of subjects (e.g., cycle to school and play outside, both addressing physical activity) and therefore having shared goals. As a result, a child can attain a goal that results in progress on multiple achievements; a skill that can be applied in different contexts.

6.4.2. PAL SYSTEM

This section highlights the usage of the *PAL Objectives Model* as implemented in the PAL system, an diabetes education framework aimed to provide goal-based personalised guidance in self-management training. The PAL system is a mHealth application to support children of 7-14 y.o. with diabetes mellitus type 1 in their self-management training. Kaptein et al. [40] provide a comprehensive overview of the system components; in this section we will elaborate on the role of *POM* in goal-setting and adaptation of system interaction.

The *POM* is integrated in the core of the PAL system and three user interfaces. The interfaces visualise the relations within *POM*, facilitate selection of personal goals and tasks, and show feedback on learning progress to three different user groups (i.e., children, informal care givers, and health care professionals). The PAL core —the 'brain' of the system— calculates progress based on task performance, generates verbal feedback on progress, and adapts the non-verbal behaviour of the robotic PAL based on task selection.

AUTHORING

For authoring of personal learning goals a web-based application, *palControl*, was developed (see Figure 6.7). The user interface visualised the goal hierarchy and relations with achievements encoded in respectively *om:requires*, *om:includes*, and *om:contributeTo*. The logic within the *POM* ontology supported goal selection. Besides, the tool was used to add new children and provide information to the user model such as the `InsulinDevice` and `Hospital` of the child and the child's personal `LearningGoal` and `Achievement` selection. An earlier version of the tool was described and evaluated in [37]. Next we will describe briefly the process of goal selection.

Visualisation of achievements and goal hierarchy The interface shows all the achievements and an option to (un)fold specific or all achievements to show/hide the goals within each achievement. The achievements are ordered from left to right based on their `Difficulty`. Both achievements and goals are colour-coded indicating their status (derived from the value in *om:progress*): grey depicts an inactive objective, yellow an active one and green an attained one. Goals may hold icons for who last updated the status (system or user; *om:setBy*) and to indicate their relevance for either pen or pump users (*om:needsDevice*; derived from the associated `LearningGoalType`). A detailed description of each goal is given upon mouse-over.

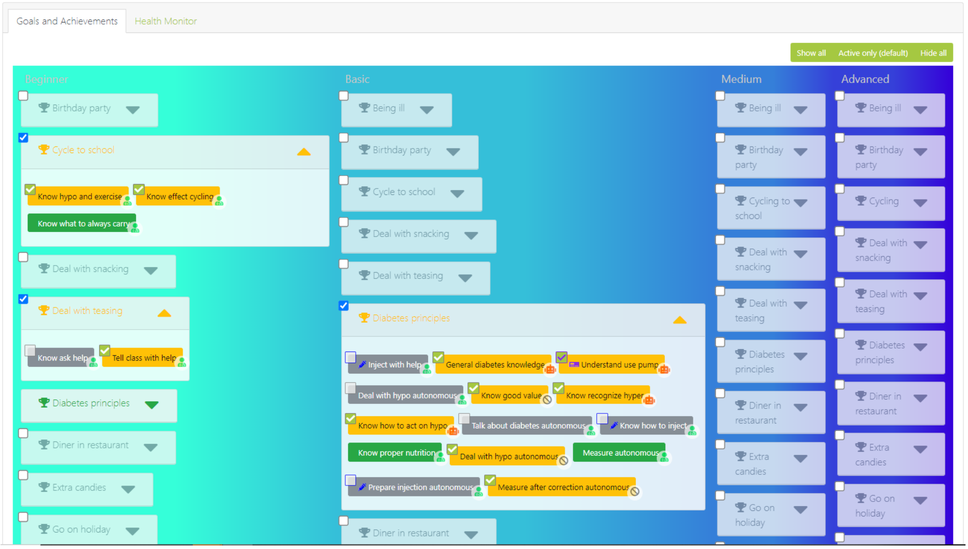


Figure 6.7: Visualisation and goal selection in *palControl&Inform*.

Collaborative goal setting In a face to face meeting the HCP and child will discuss challenges the child faces in his/her daily life. Based on this they will manually select a relevant achievement (`AchievementType`) or individual goals (`LearningGoalType`). A future development would be to select the first goals based on the outcome of an assessment upon intake of the child. For each selected objective individuals (`Achievement`, `LearningGoal`) that track the *om:progress* of the child are created in accordance.

Automated selection Based on the manual selection, an automated selection mechanism will ensure all required achievements, goals and tasks are selected based on relations in the *POM* ontology. The selection mechanism activates an achievement whenever one or more goals within are activated (*om:contributeTo*), and it activates all goals within an achievement whenever an achievement is selected (*om:includes*). When activating goals it will take into consideration relevance of this goals for the specific child based on the child's `Device` (i.e., pen or pump). And for each activated goal the related tasked are activated (*om:attainedBy*) taking into consideration their relevance based on the `Location` (i.e., the Netherlands or Italy) of the child. Further, for each items the hierarchy is traced, selecting prerequisite objectives as well (*om:requires*).

LEARNING ACTIVITIES

Children can develop their knowledge and skills by performing activities in the mobile *myPAL app*. The application contains various games, a medical diary, and the *myPAL Objectives Dashboard* (*cPOD*, Figure 6.8). The *cPOD* is described and evaluated before in [41]. Here we will present the important functionality of the *POM* in the *myPAL app*, specifically the *cPOD*.

Awareness of objectives The first aim was to visualise the objectives to the children and create awareness of their personal learning path. The interface shows all active achievements, the goals within the achievement, and tasks to attain each goal. Inactive goals (`LearningGoal` with *`om:progress`* 0.0) are shown as well so the child knows what more is to come to complete the achievement in the near future. Also, the child can see more advanced `Difficulty` levels of the achievement that can be made available later on in live. For each achievement and goal a clear progress indicator is shown so the child knows how much activity is required for attainment. Thus the interface combines information from both the domain model and the user model.

Task selection From the *cPOD* the child can directly start the learning activity required to progress towards goal attainment. Next to each goal is a button opening a new screen that displays all related and relevant (based on `Device`) tasks. Whenever a goal —and thus related tasks— is selected on multiple difficulty levels (i.e., prerequisites are not yet met), the system will offer the tasks on the lowest, active level.

Stylized behaviors Based on selection of a specific task the behaviour of the PAL actor (robot or virtual) is adapted to act more as an instructor or more peer-like behaviour. The robot roles suitable for each activity are defined in the PiNS ontology and a behaviour manager is responsible for selection of according non-verbal behaviours from a predefined and annotated set of behaviours.

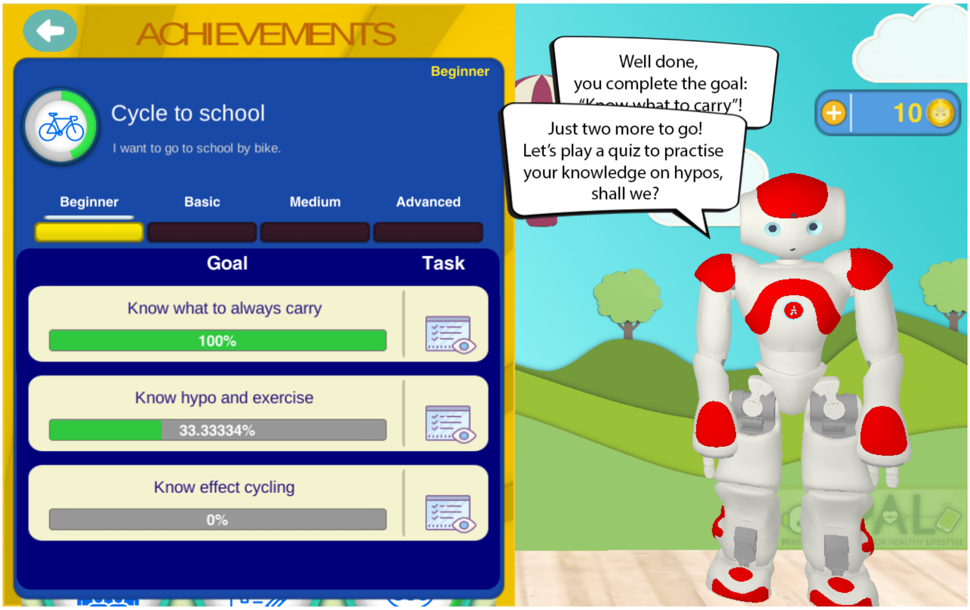


Figure 6.8: Progress tracking and task selection for children in *myPAL Objectives Dashboard*.

Feedback on performance The PAL actor will also provide feedback on goal attainment and progress to motivate the child. An episodic memory module and ontology are connected to the *POM*, they use the progress data to make remarks about recent attained goals and ones that showed no progress for some time.

Attainment visualisation

An important function of usage of *POM* is to give insight in to progress of the child towards autonomous self-management to the child itself (as described above), to the informal care giver (in PALInform) and to health care professionals (in PALControl).

Goal calculation First, progress on each objective is calculated based on the relations between objectives defined in the ontology and a child's task performance. Upon activation of objectives an individual is created in the child model with `om:hasProgress` 0.0. Whenever a task is completed successfully, the progress is updated to 1.0. Also the progress for the related goal is recomputed; a percentage of completion is calculated based on the total number of *available* tasks for this goal and the number of these tasks that are completed. Finally, the achievement progress is updated for all active achievements in which the goal is included. Here the percentage of completion is calculated based on the total number of *relevant* goals in a specific achievement and the number of them that are attained. (Whenever a previous inactive achievement is activated the progress will be set dependent on included goals already attained.)

Visualisation Health care professionals can view the child's progress in *palControl* (Figure 6.7). Green nodes indicate completed achievements and goals (`om:progress` is 1.0). Upon mouse-over on the setter-icon the completion date is shown. This view gives a full overview of the status of all achievements. The HCPs can also access *palInform* (Figure 6.9) to view a timeline indicating dates on which the child attained a new goal or achievement. The view can focus on weekly, monthly or quarterly overview. *palInform* is also accessible to authorised informal care giver such as parents. The goal-tree is visible for parents as well but opposed to HCPs they cannot edit goal status.

6.4.3. Evaluation Results & Discussion

We evaluated the model with three user groups: 1) children with type 1 diabetes mellitus, 2) their parents as informal caregiver, and 3) health care professionals (i.e., paediatric diabetes nurses and/or doctors). With each user group we evaluated the *OM* data-model (i.e., the class definitions: concepts and relations), the domain specific content defined in *POM* (i.e., the individuals that are concrete achievements, goals, tasks and accompanying learning materials), the user interfaces (resp. the *myPAL Objectives Dashboard* for children (*cPOD*) and *palControl&Inform*), and the functionality in the *PAL system* depending on the *POM* (i.e., how it supported motivation, personalisation and adaptation of the system). These evaluations were part of the integrated PAL 3.x evaluation, divided in two cycles (3.0 and 3.5) running May 2018-January 2019. In the remainder of this section we present the results per user group and briefly discuss these results.

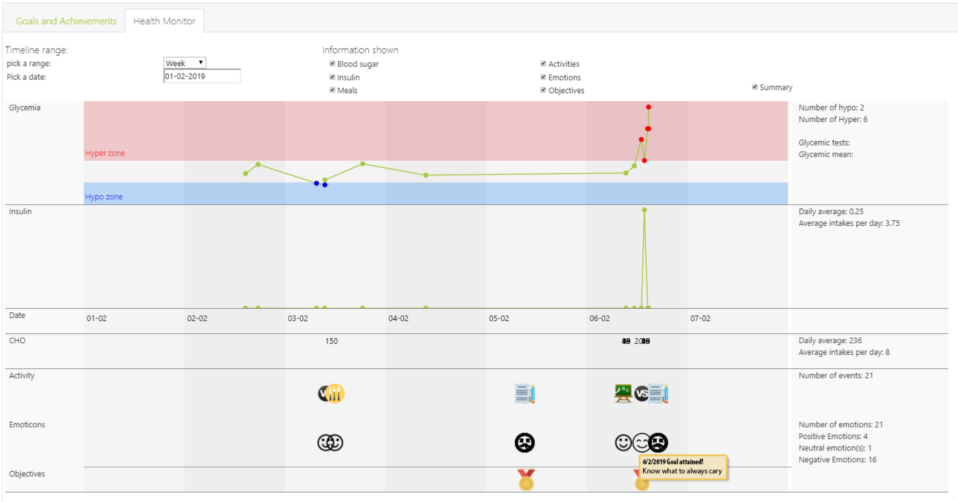


Figure 6.9: Progress tracking for (informal) care givers in *palControl&Inform*.

CHILDREN

The *POM* data-model and content, as well as the *cPOD* were evaluated with children both structured in a usability study and "in the wild" as part of the integrated PAL 3.x experiment cycles.

The usability study (N=12) has been reported extensively in earlier work [41]. From this study we learned that the *cPOD* was insufficiently clarifying the different concepts and relations in the *POM* (i.e., the achievement, learning goal, and task as well as properties such as difficulty levels and the difference between global and personal data). Therefore, we were unable to evaluate whether the *POM* data-model and content were fitting user needs. This user interface was improved upon for the integrated experiments.

During the PAL 3.x experiment 24 Dutch children used the *myPAL app* for several weeks. In the 3.0 cycle children (n=14) did have personal learning objectives, but generic activity content, in the 3.5 cycle (n=10) activity content was adapted towards personal objectives. Afterwards, children filled a questionnaire about the *myPAL app* with a number of questions dedicated to the *cPOD*. The first seven questions were statements to be rated on a five-point scale: either fully agree, partially agree, don't know, partially disagree or fully disagree. This was followed by four open-ended questions.

The majority of children (3.0 n=11, 3.5 n=9) agreed (partially) that it is important to have personal goals for disease management, and that the goals were helpful in disease management (3.0 n=7, 3.5 n=9). Further, children believed tasks to support goal attainment (3.0 n=10, 3.5 n=5). Self reported goal attainment and task completion varied between children: In the 3.0 cycle 6 children (partially) agreed to have attained all their goals, opposed to 5 children that (partially) disagreed. In the 3.5 cycle 3 children reported to have attained their goals and 4 reported not to have done so. In the 3.0 cycle 5 children (partially) agreed to have competed all their tasks, opposed to 6 that (partially) disagreed. In the 3.5 cycle 4 children reported to have competed their tasks and

3 (partially) disagreed to have done so. Children (partially) agreed that the activities in the *myPAL app* were supporting training (3.0 n=12, 3.5 n=8) and attaining (3.0 n=10, 3.5 n=7) of their goals.

For the open questions, eight children were excluded because they either did not answer (n=6), or answered about the *myPAL app* in general and not specifically the *myPAL Objectives Dashboard* (*cPOD*, n=2). Children liked about the *cPOD* that they could set their personal goals, that they could undertake activities and attain a goal/achievement, being able to see when they attained a goal/achievement, seeing their progressing towards attainment, being rewarded with points (allowing them to customise the game background and robot dance). Children reported less dislikes; five children disliked nothing. Incidental dislikes were: not being able to set new goals, not attaining a goal for each activity done, and “mean feedback” given by the robot. Seven children self-reported to have understood everything about the *cPOD*. Three children did not understand everything but did not specify what was unclear. Specific responses were that is was unclear where to find the achievements, how progress was reached, and why some achievements were not attained. None of the children provided suggestion to improve the *cPOD*, only for additional rewards such as new games and robot clothing.

It is doubtful whether children consciously used the *myPAL Objectives Dashboard*, and as a consequence experienced increased motivation to attain their goals and achievements. Three children admittedly never looked at the *cPOD*. Researcher observation suggests that young children never used the *cPOD*, whereas some older children did look at and consciously attempted to attain their objectives, however, this was not reflected in the questionnaire responses that did on average not vary between the older and younger children. Nonetheless, children did play games and reached goal attainment (unconscious). In cycle 3.5 the *POM* facilitated personalised content (opposed to cycle 3.0 where generic content was available to all children). For example, children were presented only those quiz questions related to and contributing towards attainment of selected goals. This personalisation of content was noted by some of the children. However, it did not yield significant differences in questionnaire responses (although, children proved to learn more with PAL-system, in general [12]). It should be noted that a relatively small number of children participated in this evaluation. The results give rise to better integrate the dashboard in the overall PAL-usage (including the provision of guiding avatar dialogues) and, subsequently, test the effectiveness of the *myPAL app* in a controlled experiment with and without the *cPOD*.

Parents

A total of 10 parents, who participated in the PAL3.x study, filled a questionnaire and four of them participated in a semi-structured interview about *palControl&Inform*. The questionnaire included questions about there usage and satisfaction of the *palControl&Inform*, the interview further elaborated on this and asked them about future expectations.

None of the parents ever used *palControl&Inform* due to various reasons: parents forgot its existence (n=2), or were unaware of its existence (n=3), checked progress in the *myPAL app* (n=1), or reported their child never to have used the *myPAL app* (n=2). Two parents did not provide a reason for never using *palControl&Inform*. Due to lack of usage of *palControl&Inform* we could not collect information about the parents understanding of the *PAL Objectives Model* and implementation thereof in the user interface.

Albeit never using the app, parents did report to expect added value of objectives in the child's self-management development (3 partial agree, 4 fully agree), agreed that insight in their child's progress is valuable (5 fully, 2 partial), and hold a strong preference for being able to author their child's objectives (5 fully agree, 2 partial agree). One parent commented that goal-setting via consultation with the HCP (as is the designed practise) was the favourable process.

Upon further questioning, while looking at *palControl&Inform*, parents elaborated further on their opinions. Parents expressed personal preferences such as updates via e-mail, goals including personal values. One of the four interviewed parents reported their child to purposely pursue goal attainment, two other parents believed that the personalised game content facilitated relevant learner experience and, consequently, contributed towards self-management learning.

Even though parents agree that insight into their child's self-management goals is important, none of them ever used the tool offered. This indicated that the tool did not meet their needs. The fact that suggestions were given to send an e-mail about it indicated that parents are too distracted to remember to check the *palControl&Inform*. A follow-up study should acquire further requirements to improve the strategies and user interfaces to support this user group for their role in child's self-management learning.

HEALTH CARE PROFESSIONAL

Health care professionals were heavily involved in the development of the *OM* and *POM*, therefore, we did not further evaluate the model with these professionals. To check where the model was understandable for others as well we showed the *palControl&Inform* to a paediatric diabetes nurse that was involved in the PAL project before but not familiar with this version of the ontology. This session resulted in the following remarks: The achievements and learning goals are formulated and visualised clearly. Also the difficulty levels were clear and fitting the objectives. Goal-setting works quicker compared to the earlier version; the HCP was able to select goals for a child within a few minutes. At first, the HCP did not notice the automated selection of prerequisites; this was due to some lag in visualisation of updated goal/achievement status. This lag also resulted in the HCP multiple times clicking the same node causing it to deactivate unintended. The domain specific goals were evaluated as complete, clear, relevant and fitting yet challenging for children aged 7-14. Formulation of achievements as daily live challenged was appreciated. The HCP did mention similarity between goals, indicating that it was unclear that goals could be 'repeated' on a different difficulty level.

6.5. DISCUSSION

To support personalising of a learning path in self-regulated learning we created a knowledge base of domain objects and objectives. This ontology is different from existing works such as [34, 35] who personalise content but do not take into account learning objectives and the learners real-time state thereof, while educational theories state the importance of the learning objectives as starting point for personalisation of education (e.g., [14]).

We have created a domain independent *Objectives Model* and individuals specific to the diabetes self-management domain. The upper ontology provided the foundation of a comprehensive and coherent set of achievements, goals and tasks, as well as the storage of child data necessary for personalised goal-setting. The model was successfully integrated in the PAL learning support system for children (aged 7-12) with diabetes. The inclusion of the *POM* showed valuable in creating awareness of personal goals, monitoring of progress, offering of relevant tasks and tailored feedback. In the evaluation of the overall system [12], the children proved to learn more with the PAL-system compare to care-as-usual. However, the specific contribution of the *PAL Objectives Model* on the observed learning gain, needs to be assessed in future studies.

The current domain specific individuals have been created in collaboration with diabetes-care experts from both the Netherlands and Italy (which have a different view on the diabetes regime and self-management support), resulting in an ontology that allows for location-based customisation of tasks. To establish the goal-set and learning content, was a demanding (multi-disciplinary) job. Up-scaling of the PAL-system to other hospitals and countries might require further extensions of the domain specific models. For this, the availability of a knowledge-base to chose objectives and an authoring tool to make some adjustments, would be very helpful.

More specifically, the role of the 'topic' description and assignment of 'difficulty' has been subject of discussion throughout the development of *POM*. Each goal has a difficulty level and one or more topics. Achievements are "empty" shells to structure these goals; they do not define knowledge but structure DKEs. Achievements are important concepts for the HCP to summarise goals and have a overview of the knowledge base. This raises the question whether achievements should get a property of topic and difficulty that is filled with the information inherited from included goals. This functionality is not supported in RDF, thus would require an external script scraping this information. Alternatively, an API function can query for this information and show it in the interface without it being written into the knowledge base. This functionality would support topic-based achievement searching.

For the specifications of a specific goal, it is important to first define the desired end knowledge state, and then trace back on the required knowledge. The preceding knowledge levels proved to span across the three different Bloom-pyramids (i.e., cognitive, affective and psychomotor skills) and, consequently, we had to apply the Bloom taxonomy in a flexible way. For example, to know when to administer insulin the child should first understand the effect of insulin. And to be able to administer insulin (skill), a child first needs to know where injections can be given (factual knowledge). More generally, the learning goals that the PAL-system can accommodate will be constrained by the technological capabilities of the interactive agent, the content of the educative games and the monitoring possibilities. The learning with the PAL system is intended to *supplement* or *augment* the teaching of the caregivers and parents, who supervise the process and complement where needed.

The *Objectives Model* provides the foundation of an interactive learning environment that supports teachers and learners in a comprehensive and coherent way with the personalised goal-setting & task planning, and the progress monitoring & agent-based feedback. The ontology was developed for this purpose and, consequently, differs from other ontologies that have been developed for different purposes and applications. Corresponding to the ontology development for automated scenario-based training [42], our *PAL Objectives Model* distinguishes an upper- and lower-ontology supporting the learning of "frames" or "mental models" with less emphasis on the modelling of the physical environment and game dynamics. Similar to the ontology for serious games [43], our *POM* assesses the individual performance for personalisation (however with less emphasis on learner style). Further, our model bares resemblance to the domain- and user models in the domain of adaptive hypermedia (AH) (see for an overview [7]) with less emphasis on the modality of the content and knowledge types. We started to construct a concise model, focusing on the learning objectives and tasks for personalisation and human-agent communication. The ontologies for scenario-based training, serious games and adaptive hypermedia could provide elements to refine the user and domain models of the *Objectives Model*.

6.6. CONCLUSION

6.6.1. SUMMARY

To answer the first research question ("How to model learning objectives.") we looked into how learning goals are formulated in pedagogy and ontologies for education. Effective learning goals must attune to the appropriate level, a way to structure this is Bloom's taxonomy [17, 20]. Further a learning goal must have attributes presenting relations and descriptions (e.g., [14, 15, 28, 31]). Formalisation in an ontology supports shared understanding and automated reasoning [13]. We modelled educational objectives (i.e., achievements, learning goals and accompanying tasks). The upper ontology structured the classes and relations and defined domain independent constructs (i.e., levels and topic). The domain model specified diabetes self-management training objectives for young children. These objectives were defined with experts (paediatric doctors and diabetes nurses) and considered relevant to and covering the diabetes domain to a considerable extend. From this we conclude that our upper model adequately supported the formalisation of implicit knowledge of health care professionals on diabetes self-management training. To answer the second research question ("How to deploy the model in a learning environment.") we took a more practical approach by implementing the model successfully in an multi-modal (robot, tablet, and web applications) system providing personalised learning content. A field study with children with type 1 diabetes in the Netherlands and Italy showed our model to contribute in support of children's basic needs (autonomy, competence, and relatedness) [12]. However, there is room for improvement in user interface of and accompanying materials for progress towards goal attainment and a more advanced personalised guidance selection mechanism. In the remainder of this section we will highlight important lessons learned in the form of guidelines for the formalisation of learning objectives and future challenges and opportunities for this line of work.

6.6.2. DESIGN GUIDELINES

Recommendations for formalisation of domain specific learning goals, achievements, tasks and materials in the *OM* knowledge structure, and integration thereof in a multi-modal ITS.

- Consult domain- and pedagogic experts to make an inventory of important learning goals and define learning activities (tasks and accompanying materials) to train goals.

- Formulate achievements from logical learning units (e.g., daily challenges) that require a subset of the knowledge and skills encapsulated in the goals to improve relevance.
- Formulate achievements and goals from the perspective of the child. This facilitates a sense of ownership and increases experienced relevance.
- Provide instruction and explanation to the child on how achievements, goals and task are selected and can be attained (i.e., that progress on a goal is gained by task completion, and benefits earned by this).
- Define user characteristics and relate them to goals and/or tasks to indicate personal relevance.
- Embed the objectives in the ITS application by making easily accessible to the (child) user and integrate them in other system functionality such as feedback provided by an pedagogical agent.

6.6.3. Challenges and Opportunities

Suggestions for future work in terms of considerate additions to and usage of the *Objectives Model*.

- Children showed difficulty understanding the concepts (achievement, goal, task) and the way these items were intended to guide their learning path. It's worthwhile to further explore and understand these difficulties in order to update the model and improve the instructions and explanations for better understanding.
- The acquisition and formalisation of objectives is a time consuming effort which might be supported by better tooling (the WiSCE Confluence page appeared still rather 'technical' for experts to directly create the domain model content).
- An increased number of learning activities (i.e., tasks and accompanying materials such as questions and carddecks) is necessary to maintain interest and engagement. In the current implementation often the number of possible tasks for selected goals were minimal and contributed to motivation loss. Recommender systems for online resources can be connected to find more content (see [34]).
- Especially for children who do not consciously use the *cPOD*, it is important to have an intelligent task selection mechanism in place that guides the child towards completion of those tasks that contribute to goal attainment.
- Add to the user model an estimated knowledge level based on age and/or pre-test results (and in the PAL case diabetes onset) to lessen the burden of initial goal-setting. Or add a probabilistic model to estimate the child's knowledge based on other criteria such as interests and knowledge on other topics. This kind of model can be used for more fine-grained recommendations of the most appropriate next

activity, for example, to watch an explainer video if estimated knowledge is low and play a quiz when estimated level of proficiency is good.

- Expand the leaner model with domain independent knowledge measures such as a learning aptitude test (see [35]).
- Expand the diabetes knowledge base with cognitive goals on higher level's within Bloom's taxonomy and formalise the attitude goals stated to be important in diabetes self-management training but not yet modelled.
- Add to the user model a buggy model (i.e., user's structural mistakes) to indicate possible misconceptions. Use this data to re-activate specific learning goals for re-training.

REFERENCES

[1] R. Rijgersberg-Peters, W. van Vught, J. Broekens, and M. A. Neerincx, *Goal ontology for personalized learning and its implementation in child's health self-management support,* IEEE Transactions on Learning Technologies **17**, 903 (2024).

[2] D. H. Schunk and B. J. Zimmerman, *Motivation and Self-Regulated Learning: Theory, Research, and Applications* (Routledge, 2008).

[3] C. Breazeal, K. Dautenhahn, and T. Kanda, *Social robotics,* Springer handbook of robotics , 1935 (2016).

[4] J. F. Bryan and E. A. Locke, *Goal setting as a means of increasing motivation.* Journal of applied psychology **51**, 274 (1967).

[5] P. Black and D. Wiliam, *Assessment and Classroom Learning,* Assessment in Education: Principles, Policy & Practice **5**, 7 (1998).

[6] R. Kleinrahm, F. Keller, K. Lutz, M. Kölch, and J. Fegert, *Assessing change in the behavior of children and adolescents in youth welfare institutions using goal attainment scaling,* Child and Adolescent Psychiatry and Mental Health **7** (2013).

[7] Brusilovsky, *Domain modeling for personalized guidance,* in *Design Recommendations for Intelligent Tutoring Systems: Volume 4 Domain Modeling,* edited by R. A. Sottilare, A. C. Graesser, X. Hu, A. Olney, B. Nye, and A. M. Sinatra (US Army Research Laboratory, Orlando, Florida, 2016) Chap. 15, pp. 165–2184.

[8] L. S. Vygotsky, *Mind in society: The development of higher psychological processes* (Harvard university press, 1980).

[9] L. Berk and A. Winsler, *Vygotsky: His life and works" and" vygotsky's approach to development,* Scaffolding children's learning: Vygotsky and early childhood learning , 25 (1995).

[10] B. R. Schadenberg, M. A. Neerincx, F. Cnossen, and R. Looije, *Personalising game difficulty to keep children motivated to play with a social robot: A bayesian approach,* Cognitive systems research **43**, 222 (2017).

[11] D. L. McGuinness, F. Van Harmelen, *et al.*, *Owl web ontology language overview,* W3C recommendation **10**, 2004 (2004).

[12] M. Neerincx, W. Vught, O. Blanson Henkemans, E. Oleari, J. Broekens, R. Peters, F. Kaptein, Y. Demiris, B. Kiefer, D. Fumagalli, *et al.*, *Socio-cognitive engineering of a robotic partner for child's diabetes self-management,* Frontiers in Robotics and AI **6**, 118 (2019).

[13] S. Staab and R. Studer, *Handbook on ontologies* (Springer Science & Business Media, 2010).

[14] S. Ebbens and S. Ettekoven, *Effectief leren* (Noordhof uitgevers BV, 2005).

[15] R. F. Mager, *Preparing instructional objectives. 2d. rev. ed,* (1984).

[16] R. J. Marzano, D. Pickering, D. E. Arredondo, G. J. Blackburn, R. S. Brandt, C. A. Moffett, *et al.*, *Dimensions of learning* (Association for Supervision and Curriculum Development Alexandria, VA, 1992).

[17] B. Bloom, *Taxonomy of educational objectives. Vol. 1: Cognitive domain,* (1956).

[18] L. S. Vygotsky, *Mind in society: The development of higher psychological processes,* (Harvard university press, 1980) Chap. Chapter 6: Interaction between learning and development, pp. 79–91.

[19] M. E. Hoque, *Three domains of learning: Cognitive, affective and psychomotor,* The Journal of EFL Education and Research (JEFLER) **2**, 45 (2016).

[20] D. R. Krathwohl, *A revision of bloom's taxonomy: An overview,* Theory into practice **41**, 212 (2002).

[21] R. Dave, *Psychomotor levels in developing and writing objectives,* (1970).

[22] E. J. Simpson, *The classification of educational objectives, psychomotor domain.* (1966).

[23] D. R. Krathwohl, B. S. Bloom, and B. B. Masia, *Taxonomy of Educational Objectives, Handbook II: Affective Domain* (David McKay Company, Inc., 1956).

[24] K. Stancin, P. Poscic, and D. Jaksic, *Ontologies in education–state of the art,* Education and Information Technologies **25**, 5301 (2020).

[25] S. Wang, *Ontology of learning objects repository for pedagogical knowledge sharing,* Interdisciplinary Journal of E-Learning and Learning Objects **4**, 1 (2008).

[26] T. Kasai, H. Yamaguchi, K. Nagano, and R. Mizoguchi, *Building an ontology of it education goals,* International Journal of Continuing Engineering Education and Life Long Learning **16**, 1 (2006).

[27] S. M. Rashid and D. L. McGuinness, *Creating and using an education standards ontology to improve education.* in *SW4SG@ ISWC* (2018).

[28] H. Chung and J. Kim, *An ontological approach for semantic modeling of curriculum and syllabus in higher education,* International Journal of Information and Education Technology **6**, 365 (2016).

[29] J. David, A. Lobov, and M. Lanz, *Attaining learning objectives by ontological reasoning using digital twins,* Procedia Manufacturing **31**, 349 (2019).

[30] E. L. Baker, G. Chung, and J. Herman, *Ontology-based educational design: Seeing is believing,* Los Angeles, CA: CRESST (2009).

[31] L. Ng, *Ontological model for representation of learning objectives,* Ph.D. thesis, Interactive Arts and Technology-Simon Fraser University (2005).

[32] A. Inaba, T. Supnithi, M. Ikeda, R. Mizoguchi, and J. Toyoda, *An overview of learning goal ontology,* in *Proc. of ECAI2000 Workshop on Analysis and Modelling of Collaborative Learning Interactions, Pages* (2000) pp. 23–30.

[33] S. Assami, N. Daoudi, and R. Ajhoun, *A semantic recommendation system for learning personalization in massive open online courses,* International Journal of Recent Contributions from Engineering, Science & IT (iJES) **8**, 71 (2020).

[34] P. Wongthongtham, K. Y. Chan, V. Potdar, B. Abu-Salih, S. Gaikwad, and P. Jain, *State-of-the-art ontology annotation for personalised teaching and learning and prospects for smart learning recommender based on multiple intelligence and fuzzy ontology,* International Journal of Fuzzy Systems **20**, 1357 (2018).

[35] S. Sarwar, Z. U. Qayyum, R. García-Castro, M. Safyan, and R. F. Munir, *Ontology based e-learning framework: A personalized, adaptive and context aware model,* Multimedia Tools and Applications **78**, 34745 (2019).

[36] D. Freeborn, T. Dyches, S. O. Roper, and B. Mandleco, *Identifying challenges of living with type 1 diabetes: child and youth perspectives,* Journal of clinical nursing **22**, 1890 (2013).

[37] R. Peters, J. Broekens, and M. A. Neerincx, *Guidelines for tree-based learning goal structuring,* in *Proceedings of the 22Nd International Conference on Intelligent User Interfaces, IUI*, Vol. 17 (ACM, Limassol, Cyprus, 2017) pp. 401–405.

[38] M. V. Welie, *Task-based user interface design,* Evaluation , 217 (2001).

[39] H.-U. Krieger, R. Peters, B. Kiefer, M. A. Van Bekkum, F. Kaptein, and M. A. Neerincx, *The federated ontology of the pal project interfacing ontologies and integrating time-dependent data,* (2016).

[40] F. Kaptein, B. Kiefer, A. Cully, O. Celiktutan, B. Bierman, R. Rijgersberg-Peters, J. Broekens, W. Van Vught, M. Van Bekkum, Y. Demiris, *et al.*, *A cloud-based robot system for long-term interaction: Principles, implementation, lessons learned,* ACM Transactions on Human-Robot Interaction (THRI) **11**, 1 (2021).

[41] R. Peters, E. Oleari, F. Sardu, and M. A. Neerincx, *Usability of the pal objectives dashboard for children's diabetes self-management education,* in *Proceedings of the 2019 the 5th International Conference on e-Society, e-Learning and e-Technologies* (ACM, 2019) pp. 22–28.

[42] M. M. Peeters, K. V. d. Bosch, M. A. Neerincx, and J.-J. C. Meyer, *An ontology for automated scenario–based training,* International Journal of Technology Enhanced Learning **6**, 195 (2014).

[43] B. Said, L. Cheniti-Belcadhi, and G. El Khayat, *An ontology for personalization in serious games for assessment,* in *2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)* (IEEE, 2019) pp. 148–154.

7

Objectives Dashboard for Healthcare Professionals

Educational technology needs a model of learning goals to support motivation, learning gain, tailoring of the learning process, and sharing of the personal goals between different types of users (i.e., learner and educator) and the system. This paper proposes a tree-based learning goal structuring to facilitate personal goal setting to shape and monitor the learning process.

We developed a goal ontology and created a user interface representing this knowledge-base for the self-management education for children with Type 1 Diabetes Mellitus. Subsequently, a co-operative evaluation was conducted with healthcare professionals to refine and validate the ontology and its representation. Presentation of a concrete prototype proved to support professionals' contribution to the design process. The resulting tree-based goal structure enables three important tasks: ability assessment, goal setting and progress monitoring. Visualization should be clarified by icon placement and clustering of goals with the same difficulty and topic. Bloom's taxonomy for learning objectives should be applied to improve completeness and clarity of goal content.

This chapter has been published in proceedings of the 22nd ACM International Conference on Intelligent User Interfaces (IUI), (2017) [1].

7.1. Introduction

Advancements in media technologies provide new opportunities for education. For example, Intelligent Tutoring Systems (ITSs) provide immediate tailored instructions or feedback to a learner to facilitate effective learning while lessening the students dependency on a teacher. Also, consider eHealth applications that have been designed to increase a person's knowledge and control over health and well-being. Especially in self-regulated learning motivation is highly important to optimize adherence to the education program [2]. Research suggests that goal setting and feedback on goal attainment enhance motivation (e.g., [3–5]) and learning gain (e.g., [6, 7]). Moreover, personal goal setting allows for tailoring of the learning process, this is applicable to personalization of educational technology (e.g., [8]).

Incorporating personal learning goals in educational technologies requires knowledge of learning goals relevant to the domain, a mechanism to set personal learning goals and to share this information between different types of users (e.g., doctor and patient, teacher and student) and with the system, and means to monitor learning progress. Ability-trees are used in games to structure and visualize skills that allow the player to tailor character development and game-play. Gamification has been applied to educational technologies. For example, task completion is rewarded with points or achievements. Using an ability-tree for learning goals is an interesting approach to provide a solution for goal structuring, setting, and monitoring.

In this paper we propose guidelines for a tree-based learning goal model and user interface to support collaborative goal setting. In a case study, on self-management education for children with Type 1 Diabetes Mellitus, we explore requirements for a tool to set personal learning goals. We developed a knowledge-base (ontology) formalizing learning goals and tasks based on medical protocols. We created a user interface presenting these learning goals in a tree-based graph, and enabling personal goal setting. Based on a co-operative evaluation we formulated guidelines to improve the design.

7.2. Background

7.2.1. Learning Goals

Effective learning requires commitment, adherence and motivation, which can be increased by learning goals [3, 7]. Goals enhance motivation independent of their source (i.e., assigned, self- or collaborative set), if relevance is provided [9]. However, performance is lower for unexplained, assigned goals than self- or collaboratively set goals [10]. Contrary, Kleinrahm et al. [4] found that cooperative goal setting and reflection increased motivation.

Black and Wiliam [6] concluded that awareness of goals and goal attainment improves learning gain. Similarly, goal-setting theorist believe that feedback results in setting higher goals [10]. Which in turn, leads to better performance [9]. This fits Vygotsky's [11] theory on the zone of proximal development (ZPD), predicting that experiences slightly advancing current abilities encourage and advance learning [12].

In instructional classroom learning goals are often strategically chosen to align with institutional or national standards. However, in self-paced learning the education process can and should be tailored by the learner. Self-regulated learning (SRL) increases learning gain, but is more demanding in terms of effort and thus motivation [2]. SRL theorists believe that strategies such as goal setting, self-monitoring and self-evaluation are vital for effective learning [13]. Motivation comes from goal orientation, self-efficacy believes, task value, and outcome expectations [2, 14]. SRL benefits from a mastery goal orientation (i.e., focus on learning of the ability), while a performance orientation (i.e., focus on demonstration of abilities) declines performance [14–16].

7.2.2. Ability Trees

In graph theory a tree is defined as an undirected graph in which two nodes are connected by exactly one edge. In a directed graph edges have an associated direction, and in a rooted tree one node is designated as the root. In computer science a tree is a non-linear, hierarchical data structure represented by a root node and linked children or sub-trees [17]. The direction can be from (out-tree) or to (in-tree) the root [18]. In the remainder of this paper we use *tree* as data structure.

Graph theory based models are used in e-learning environments guiding the self-paced learning process, and enabling personalization thereof (e.g., [19, 20]). Learning object graphs formalize the structure of a course, representing mandatory and recommended learning objects—including objectives but mainly content—and relations between them. Through assessment and authoring, objectives are selected, constructing individual learning paths that align with the learner's profile.

Tree-based data structures are used in games allowing players to customize their game experience. For example, in the strategy game Civilization players can choose to develop skills in alphabet or mathematics, but only after having achieved the writing skill. In role-playing games such as Diablo (I and II), players develop their character using points to gain magical powers. These so called ability-trees are visual, hierarchical representation of possible sequences of developments. Abilities are displayed in branching paths and open up after completing required prerequisites. These structures, based on abilities opposed to levelling, allow players to excel in some areas while progressing more slowly, or not at all, in others. Expected is that ability-trees are familiar, and therefore understandable, to children.

7.3. Case: Diabetes Self-Management Education

7.3.1. Current Practises in Diabetes Education

Type 1 Diabetes Mellitus (T1DM), diagnosed for a growing number of children, is a high impact digestion disease which requires daily self-management. Thus, to improve well-being and avoid complications, long-term behaviour change is necessary [21]. Learning objectives are personal and change while ageing. Therefore, self-management education is highly personalized. It is directed by challenges faced in daily life and aimed at gradual development of attitudes, knowledge and skills needed for autonomous self-management.

In the Netherlands, formalization of learning goals is limited to (annual) check lists arranged by topic and age such as the *weet & doe-doelen* (knowledge & skill goals) composed by the Dutch organization for diabetes nurses EADV (`http://www.eadv.nl`). Hierarchical relations between goals are implied; a goal can be prerequisite for one or more others (e.g., injecting insulin requires knowledge of appropriate body parts). Further, goals become increasingly complex for older children. As a result, goals may cover multiple topics and thus precede or succeed goals on different topics. Moreover, specific goals are irrelevant to some children (e.g., pump users do not necessarily need to learn injecting themselves).

Active involvement of patients in the disease management and education process is essential [22]. Objectives should be defined collaboratively between patient and caregivers. This is in line with the Motivational Interviewing (MI) guiding style adopted in healthcare counselling. The principle of MI [5] is to explore and resolve a patient's ambivalent feelings towards change, opposed to coercing or persuading. MI is believed to increase the patients commitment to developing self-management abilities.

7.3.2. Diabetes Education Framework

The PAL project[1] develops mHealth technology providing educational support to children with T1DM. The aim is to gradually increase children's self-management abilities and responsibilities. The envisioned system includes an embodied conversational agent (robot and avatar), extra-curricular educational child-agent activities, and an authoring tool.

Authoring Tool

The authoring tool is a web-based application for healthcare professionals (HCPs) designed to support goal setting, progress monitoring, and attainment registration. The current state is a functional, but minimal, prototype presenting diabetes self-management learning goals, and providing an interface to set personal goals or register attainment together with a child.

7

Diabetes Learning Goal Structure

We formalized the learning goals, as proposed by the EAVD, in an ontology (Figure 7.1). Learning goals are classified by type (i.e., knowledge or skill) and values are given for difficulty and topic. Additionally, restrictions are added for prerequisite goals. (The progress and state properties are specific to a child and values are given at a later time.) Further, achievements are added for each topic and difficulty combination, and tasks (e.g., 'win a quiz on insulin') are linked to learning goals (e.g., 'know locations for insulin injection'). The restrictions and relations allow the system to provide personalized content, and calculate and update goal progress automatically.

We created a tree-based visualization of the goals and achievements (Figure 7.2) because the merging structure of an in-tree fits the diabetes learning goals; from leafs with a single focus topics (e.g., 'Nutrition'or 'Insulin') to multilevel topics (e.g., 'Nutrition in social context') and ultimately the root node 'Self-management'. Further, tree presentations have been applied successfully to structure abilities in games.

Nodes represent a learning goal (i.e., knowledge or skill) or achievement. Edges depict connections to prerequisite nodes. Attaining all learning goals on one topic at one difficulty grants the achievement and unlocks the possibility to advance on this topic.

[1] Personal Assistant for a healthy Lifestyle: `http://www.pal4u.eu/`

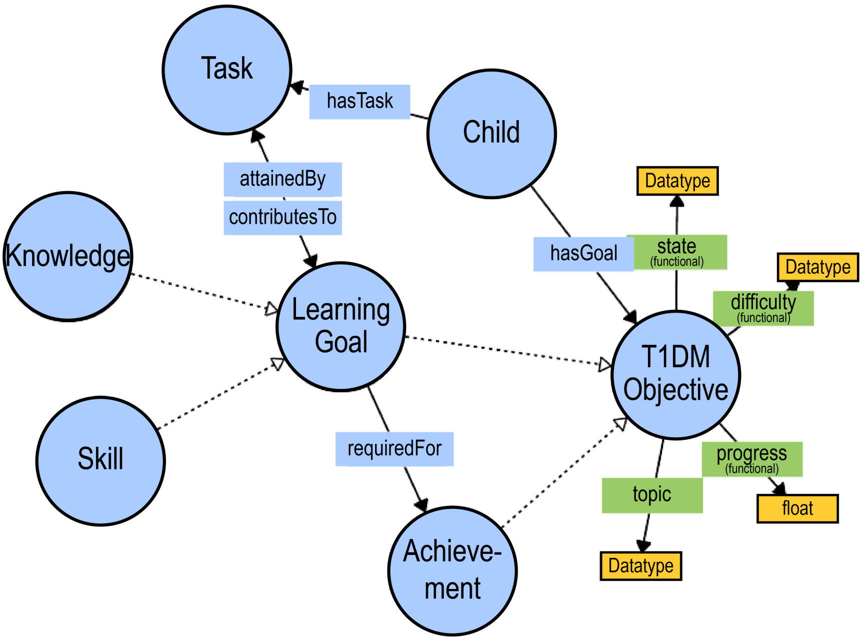


Figure 7.1: The Diabetes Education ontology. Nodes depict the objective types (classes). Arrows depict object or data properties, dotted arrows depict subtype relations. Self-management learning goals are instantiated in the knowledge or skill class.

For example, a child who attained all difficulty 1 goals on glucose (i.e., knows why measurement is needed, how to correct a hypo, and understands the measurement value) is granted the achievement 'Novice Glucose', and may advance to the next level. Edges connecting a node to multiple, represent nodes prerequisite to more than one others. For example, knowledge of the correct response to a glycemic value is required for both 'Basic Glucose' and 'Independent Measurer'. The proposed model facilitates tailoring to a child's situation and development by selection of personal learning goals, while enforcing to have obtained prerequisite abilities.

COLLABORATIVE GOAL SETTING INTERFACE

A minimal graphical interface was created to support goal setting and progress monitoring (Figure 7.2). It displays the goal-tree and provides mechanisms to switch between goal states (active, inactive and attained). Active goals, pursued in the near future, are yellow. Attained goals are green. Inactive goals are greyed-out but visible to raise awareness, allowing children to ask about them. The goal state is changed by clicking the node (register attainment) or the selection box at the top-left corner (activate). Upon activation, all prerequisite goals activate automatically, the user can inactivate (irrelevant) goals clicking the selection box. The goal-tree is presented top-down, first displaying goals for new learners, to avoid frequent scrolling.

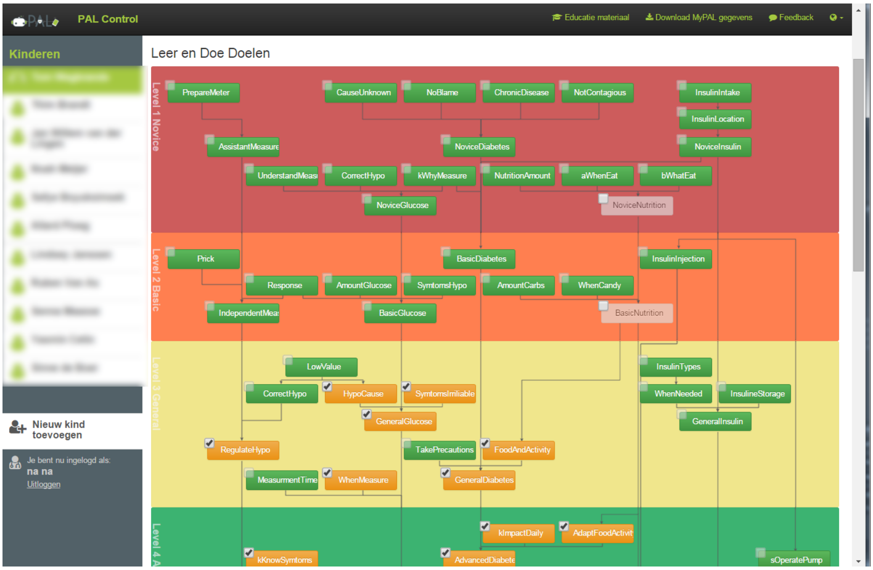


Figure 7.2: The authoring tool (PAL Control), displaying the diabetes learning goals (i.e., knowledge and skills to attain to progress towards self-management) and achievements with current progress for a 10 year old boy. Attained objectives are green, yellow ones are active. Coloured horizontal bars depict difficulty levels. Topics are arranged vertically.

7

7.4. Methodology

To gain insight about diabetes education protocols and elicit user requirements, interviews were conducted with HCPs at an early stage. Implicit knowledge and experience appeared fundamental to forming of the, highly personalized, educational process. Besides, HCPs are not used to thinking from a technological design perspective. Hence, development of a goal ontology and user interface were no trivial tasks. Therefore, we selected a co-operative, formative evaluation method providing a minimal example and collaboratively composing guidelines for further development.

Evaluations were conducted with 7 HCPs (6 Dutch nurses, 1 Italian doctor), and 35 children (aged 7-12 M/F) and their parents, in 3 hospital (2 Dutch, 1 Italian) in May-June 2016. Each child visited the hospital two times, once at the start and end of a three week period. In between children played educational activities at home. The first consultation covered personal goal setting using the authoring tool. In the second meeting progress was discussed. An observer was present and audio, keystrokes, and clicks were recorded during at least each nurse's earliest consultation. Additionally, in the Netherlands, training sessions of approximate half an hour were carried out prior to consultations using a think-out-loud protocol whilst preparing goals for the first child. Semi-structured interviews were conducted posterior with all professionals.

7.5. RESULTS AND DISCUSSION

A total of 11 goal setting consultations have been observed and analysed. A typical consultation was attended by a paediatric diabetes nurse, child, and one parent and lasted for about 10 minutes. All meetings included assessment of the child's current abilities, goal suggestions by the professional, conformation by the child and/or parent, and registration of goal state (active, inactive or attained) in the authoring tool. These steps were repeated for individual goals by a top-down walk through of the goal-tree.

Assessment of the child's current abilities was mostly straight forward: the nurse asked whether or not the child knows or can do x, where x is a specified goal. In six sessions answers were given by the child and parent in collaboration or turn, in three sessions the child responded alone, in two sessions only the parent was involved (both cases a 7-year old child). In four occasions assessment was done more implicitly by 'small-talk' (e.g., "Your horseback riding right? Do you have any difficulties with your diabetes then?"). If agreed on goal attainment, it was registered as such. A goal was set active if no agreement was reached. Assessment was suspended for a topic if goals were set active.

Although the selection mechanism was easily understood and goal setting was done effectively, the *topic clusters* were not clear. For example, Nurse 1 was looking at glucose goals in search for a goal on nutrition. Further, the *achievement concept* was hard to grasp. For example, instead of activating goals and achievement, Nurse 1 explained that she did not select the achievement because the related goals were not yet attained, and Nurse 3 marked achievements attained while related goals were still active. In addition, *handling irrelevant goals* was troublesome for two nurses. For example, Nurse 4 registered 'Insulin Injection' attained because the child, as a pump user, did not need to learn how to inject insulin by pen (opposed to leaving it inactive). Moreover, questions were raised about the *intent or meaning for specific goals*. For example, Nurse 3 doubted whether to activate 'Insulin-type Needed', because the child did know the facts but was not yet able to apply this knowledge in daily situations. Further, nurses were unable to select goals they had in mind for a child because they were not present in the goal-tree. For all but one, this were attitude goals such as feeling more secure. In posterior interviews with HCPs, the following five issues on the goal ontology and user interface were discussed.

First, the *complexity of the goal-tree* was too high. For the nurses, the meaning of 'achievement' (i.e., *not* concerning new knowledge or skill) was not clear. Furthermore, the user interface was not clear: Icons to clarify the topic were proposed, and all HCPs suggested or favoured visualization of achievements by displaying related goals in nested nodes. One nurse suggested showing only the active difficulty level, hiding others. To support tailored learning goals, this must be done per topic or a collapse-on-select mechanism could be considered: when selecting an achievement it unfolds and enables (de-) activating containing goals.

Second, the *clarity and completeness of the goal content* showed shortcomings. The ontology did not include attitude goals and lacked distinction between factual knowledge and the ability to apply this knowledge. Application of Bloom's Taxonomy [23] for educational objectives and gradual development of more complex skills might solve these shortcomings, distinguishing cognitive (knowledge), psycho-motor (skill) and affective (attitude) processes. The cognitive dimension includes multiple levels such as remembering, understanding and applying [24]. Explicitly formulating goals on these levels bridges knowledge development, practice and assessment.

Third, *feedback on goal attainment for progress monitoring* was missing (i.e., whether a goal was registered attained manually or by the system). When the start and end meeting were attended by different nurses, they were unsure about newly attained goals. So, information about goal-state changes should be provided for nurse shiftes and time periods (i.e., simple visualization of goals attained since the last meeting by placement of an icon).

Fourth, *goal setting and assessment of current abilities* were partly supported. Although goals can be selected at any difficulty, while the recursive mechanism ensures that prerequisites are activated, nurses started at the top of the tree and worked their way down assessing each goal. As a result, an overview of current abilities was created, and for each topic unattained goals were set active. This is different from the current practise where a single focus is chosen based on the child's experiences. Nurses had different preferences, either favouring several goals allowing the child varied experiences in the learning framework or favouring a mechanism selecting a focus topic from active goals. Nurses agreed that the top-down goal assessment provided a valuable overview of the current state of abilities of a child. It was time consuming, but nonetheless considered more usable than current check-lists. Three nurses suggested to let the child play a game (e.g., with the robot) to assess abilities and serve as input for goal setting.

Fifth, *collaboration with the child during goal setting* was partly supported. According to the nurses, the authoring tool eased interaction between them and the child. However, goal setting was less collaborative than desired. Children's involvement was limited to (dis)agreement; they did not proactively discuss specific goals, while active involvement is key to motivation for behaviour change [5]. Children's involvement can be improved by allowing them to select their personal focus from active goals. Moreover, HCPs may benefit from training in collaborating with a child using the authoring tool. Nonetheless, goal setting was believed helpful making the education process more interesting and transparent to the children.

The present study has some limitations such as a lack of quantitative data proposed by the number of participants and method. We do not report on usability statistics because they do not provide novel information for research and development of intelligent user interfaces. We plan to expand our user base and methods to evaluate alternative interfaces and investigate the effect of goal setting on learning outcomes. Although, other structures might be feasible as well, we have chosen a tree-structure because this fits our domain. Further research, presenting alternative structures, is needed to make any conclusions on structure preferences. Suggestions provided in this paper are applicable to tree-based structures.

7.6. Concluding Guidelines

The main challenges addressed in the present work are the development of an ontology for diabetes self-management education and an interface to support collaborative personal goal setting and monitoring. An authoring tool was created for this purpose and co-evaluated with healthcare professionals.

Guideline 1: The authoring tool should provide *clear, visual feedback* on goal *structure*, and *active state and progress*. For example, by usage of icons depicting topic and state changes.

Guideline 2: The *different concepts* of the model (e.g., goal and achievement) should consistently have a *different representation* (e.g., shape) in the user interface.

Guideline 3: The domain should be fully covered in goal content. In our case *differentiation* between *affective and cognitive*, and *factual and application* objectives should be embedded (e.g., Bloom's taxonomy).

Guideline 4: The authoring tool should support, next to goal setting, progress monitoring and attainment registration, *assessment of current abilities*.

REFERENCES

[1] R. Peters, J. Broekens, and M. A. Neerincx, *Guidelines for tree-based learning goal structuring,* in *Proceedings of the 22Nd International Conference on Intelligent User Interfaces, IUI*, Vol. 17 (ACM, Limassol, Cyprus, 2017) pp. 401–405.

[2] D. H. Schunk and B. J. Zimmerman, *Motivation and Self-Regulated Learning: Theory, Research, and Applications* (Routledge, 2008).

[3] J. F. Bryan and E. A. Locke, *Goal setting as a means of increasing motivation.* Journal of applied psychology **51**, 274 (1967).

[4] R. Kleinrahm, F. Keller, K. Lutz, M. Kölch, and J. Fegert, *Assessing change in the behavior of children and adolescents in youth welfare institutions using goal attainment scaling,* Child and Adolescent Psychiatry and Mental Health **7** (2013).

[5] S. Rollnick, W. R. Miller, and C. C. Butler, *Motivational Interviewing in Health Care: Helping Patients Change Behavior* (The Guilford Press, New York, 2008).

[6] P. Black and D. Wiliam, *Assessment and Classroom Learning,* Assessment in Education: Principles, Policy & Practice **5**, 7 (1998).

[7] A. M. Grant, *An integrated model of goal-focused coaching : An evidence-based framework for teaching and practice,* International Coaching Psychology Review **7**, 146 (2012).

[8] 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 child's diabetes self-management,* (2016).

[9] E. Locke and G. Latham, *New Directions in Goal-Setting Theory,* Current Directions in Psychological Science **15**, 265 (2006).

[10] E. A. Locke and G. P. Latham, *Building a practically useful theory of goal setting and task motivation: A 35-year odyssey.* Am. Psychol. **57**, 705 (2002).

[11] L. S. Vygotsky, *Mind in society: The development of higher psychological processes* (Harvard university press, 1980).

[12] B. R. Schadenberg, M. A. Neerincx, F. Cnossen, and R. Looije, *Personalising game difficulty to keep children motivated to play with a social robot: A Bayesian approach,* Cognitive Systems Research (In Press).

[13] B. J. Zimmerman, *Becoming a self-regulated learner: Which are the key subprocesses?* Contemporary Educational Psychology **11**, 307 (1986).

[14] P. R. Pintrich, *The role of motivation in promoting and sustaining self-regulated learning,* Educational Research **31**, 459 (1999).

[15] C. Ames, *Classrooms: Goals, Structures, and Student Motivation,* Educational Psychology **84**, 261 (1992).

[16] C. A. Wolters, L. Y. Shirley, and P. R. Pintrich, *The relation between goal orientation and students' motivational beliefs and self-regulated learning,* Learning and Individual Differences **8**, 211 (1996).

[17] T. A. Sudkamp, *Mathematical preliminairies,* in *Languages and machines: an introduction to the theory of computer science* (Addison Wesley, 2005).

[18] K. Mehlhorn and P. Sanders, *Algorithms and Data Structures: The Basic Toolbox* (Springer Science & Business Media, 2008).

[19] Y. Atif, R. Benlamri, and J. Berri, *Learning Objects Based Framework for Self-Adaptive Learning,* Education and Information Technologies **8**, 345 (2003).

[20] A. N. Viet and D. H. Si, *ACGs: Adaptive Course Generation System - An Efficient Approach to Build E-learning Course,* in *Proc. CIT'06* (2006) pp. 259–265.

[21] D. Freeborn, T. T. Dyches, S. O. Roper, and B. Mandleco, *Identifying challenges of living with type 1 diabetes: Child and youth perspectives,* Clinical Nursing **22**, 1890 (2013).

[22] American Diabetes Association, *Diagnosis and classification of diabetes mellitus,* Diabetes Care **36**, 67 (2013).

[23] B. Bloom, *Taxonomy of educational objectives. Vol. 1: Cognitive domain,* (1956).

[24] L. Anderson, D. R. Krathwohl, and B. Bloom, *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (Allyn & Bacon, 2001).

8

Objectives Dashboard for Children's Diabetes Self-management Education

Children will only benefit from educational technologies and e-coaches when they understand the long-term consequences and are (intrinsically) motivated to use these support systems. This paper presents an Objective Dashboard that integrates educational achievements, goals and tasks with gamification features (such as challenges, scores and rewards) to advance the interests and engagements of children with type 1 diabetes when using the Personal Assistant for a healthy Lifestyle (PAL) system. By linking in-app activities (e.g., play a quiz or keep a diary) to relevant educational achievements, and to skills and knowledge required in daily life, we aim to increase intrinsic motivation and thereby usage.

We designed a dashboard displaying personalised achievements, learning goals and tasks in the domain of diabetes self-management education. We used common user interface design patterns such as layering, colouring, and iconic presentation to organise complex information and reinforce the relations between concepts. Subsequently, we conducted a usability evaluation with twelve children, on the basis of which we refined our design. We found that, colouring and layering were to some extent effective, however, iconic representations were insufficient. Therefore, we recommend to provide short, descriptive labels at any time.

This chapter has been published in proceedings of the 5th ACM International Conference on e-Society, e-Learning and e-Technologies (ICSLT), (2019) [1].

8.1. Introduction

eLearning systems in general, and more specifically Intelligent Tutor Systems (ITS), offer great possibilities; current state-of-the-art media technologies facilitate learning anywhere anytime, independent of a trainer and institute. However, the success of this *self-regulated* learning is fully dependent on adherence and thus requires high (intrinsic) motivation [2]. Further, these technological advancements support tailoring of learning process and material (e.g., [3–6]) leveraging on personalization, which is believed to increase relevance and thereby motivation. However, provision of long-term interesting interactions remains a challenge. For example, Leite et al. [7] concluded that, when using robots in education and health care, novelty wears off quickly and interaction diminishes. But also eLearning courses in higher education show high drop-out rates [8]. The implementation of game features in eLearning applications (i.e., *gamification*) is a popular means to increase and maintain interest [9]. However, gamification benefits risk depletion once rewards lose their value. A more traditional technique improving task performance is goal-setting [10]. In classroom teaching, goals influence motivation [11] and learning outcomes [12]. Goal-setting also has been applied to ITS (e.g., [13, 14]).

One such context, where self-regulated learning is crucial, is diabetes self-management education. For day-to-day diabetes management children need to acquire, from early on, specific knowledge, skills and attitudes. This progress towards self-management is not straightforward and the path should be personalised [15]. Drawing from both gamification and educational sciences, educational objectives have been added to the *PAL system*[1], a mHealth application providing personalised, diabetes self-management education. The *PAL system* includes web-based dashboards for caretakers, an app with a medical diary, quiz and several mini games for children, and a (virtual) robot 'pal'. To enhance motivation for and usage of the app, and ultimately learning gain, educational achievements, goals, and the path to attain them, must be clear to the child. Therefore, a *PAL Objectives* system has been developed as an integrated module of the *PAL system*. The *PAL Objectives* consists of the *PAL Objectives Model* and various dashboards presenting this model for different user roles. In this paper, we present the design and usability evaluation of the *PAL Objectives Dashboard* for children embedded in the *myPAL app*. We conclude proposing design guidelines for graphical user interfaces (GUI) visualising personalised achievements and learning goals to children.

8.2. Background

It is well known that motivation plays a crucial role in the learning process, especially for children [16]. It represents the driver to go behind the merely "doing something" and it is the counterbalance to keep on trying when frustration coming from failures and negative experiences appears, possibly hampering confidence and self-esteem. Motivation is classified as *extrinsic* or *intrinsic* [17]. The former comes as an effect of outside forces, (e.g., prizes or incentives), while the latter comes from within the person who is learning because of personal interests or fun.

Gamification is the usage of game elements in non-game contexts [18] and this has been applied to educational system to circumvent limitations and increase user action.

[1] Personal Assistant for a healthy Lifestyle: http://www.pal4u.eu

Well known examples of gamification are points, badges and leader-boards, which showed effective in increasing user productivity and effort [19]. These reward strategies are believed to provoke mainly *extrinsic* motivation. However, gamification features can also motivate the user by providing goals, progress, and encouragement [20]. Visualisation of progress may induce the feeling of competence [18], which is an important aspect of *intrinsic* motivation according to the *self-determination theory* (SDT) [21].

Complementary, from educational sciences, it is known that goal-setting, progress feedback and goal attainment enhance motivation [22, 23] and learning gain [12]. While some argue that ownership of goals is important [24], other claim that relevance of the goals predicts motivation [25]. Further, according to Vygotsky's theory on the *Zone of Proximal Development* (ZPD), goals slightly advancing the learners established abilities encourage and advance learning [26].

8.3. PAL OBJECTIVES DESIGN

The *PAL Objectives* system has been developed to motivate children to complete educational tasks and thereby improve self-management skills by providing relevant, challenging educational objectives and facilitate monitoring of progress. The *PAL Objectives* consists of 1) *PAL Objectives Model* (*POM*) a knowledge-base of domain specific achievements, learning goals and tasks [27, 28], 2) *palControl&Inform* a web-application for collaborative goal-setting and monitoring [28], and 3) *PAL Objectives Dashboard* for children (*cPOD*) displaying personal objectives (i.e., achievements, goals, and tasks) in the *myPAL app*.

8.3.1. PAL OBJECTIVES MODEL

The dashboards visualise personalised achievements, learning goals and tasks from the *POM*, which consists of a global layer defining domain specific achievements, goals and tasks and relations between them, and a user specific overlay storing progress.

8

Achievements, learning goals and tasks represent educational objectives on different conceptual levels. Achievements represent challenging, relevant daily live events, such as having a sleepover or attending a birthday party, that require certain knowledge or skills. Achievements may group multiple learning goals presenting a set of skills, knowledge and/or attitudes needed to engage in the event. Learning goals represent the desired end state of specific knowledge, skill or attitude to be attained by a child (e.g., count carbohydrates). Tasks represent actions a child can undertake in the *myPAL app* to progress on learning goals (e.g., play a carbohydrates sorting game or filling the nutritional diary). Further are specified the relations between objectives such as prerequisites, and other properties such as level and topic. The separation on levels and topics facilitates non-linear learning; a child can progress faster or slower on certain topics which meets the nature of self-management skill acquisition.

Personal achievements and learning goals may be selected from the global knowledge-base via the *palControl&Inform* —contributing tasks are automatically selected based on selected goals. Further, goal-setting is adapted towards the user model. For example, goals are selected matching the device a child used for insulin intake (i.e., pen or pump). Personal progress on each objective (i.e., inactive, attained or completed) is stored on a user specific overlay of the global layer.

The structure of *POM* is domain independent and loosely based on an existing task ontology [29, 30]. Domain specific content has been created in collaboration with diabetes domain experts. The knowledge-base is formalised in an ontology as part of *PALO*, an integrated PAL ontology [27].

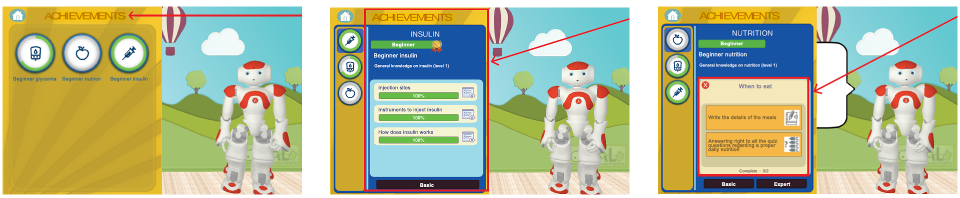
8.3.2. PAL Objectives Dashboard

The *cPOD* was designed to display educational contents, and related acquisition path, to the child in an attractive and challenging way. It should provide a clear overview of attained knowledge and skills and activities to undertake to advance progress. However, the complex, abstract and interconnected *POM* information showed challenging to present [28]. Therefore, we aimed to organise the information in the *cPOD* as understandable chunks of information and easy navigation between layers of information.

To focus attention to key information, the start screen (Figure 8.1a) displays only relevant, high-level information [31, p. 67], being the personal achievements with a progress indicator. A green indicator was chosen for the colour's symbolic meaning of success [32, pp. 48]. Active achievements are listed first, followed by those attained. Progress changes may be notified by the avatar.

Layering [32, pp. 146] was applied organising information as to manage the complexity and reinforce relationships between items. More fine-grained information such as included learning goals, and progress values are shown in the achievements pane (Figure 8.1b) that is opened by clicking on the achievement. The salient difficulty level is adapted to the child's current level of knowledge, adhering to the ZPD. Previously attained levels are displayed at the bottom to facilitate reviewing past progress. All active achievement are shown in a sidebar to ease accessibility, with the selected achievement placed on top and marked because of saliency and to minimise scrolling [33].

Tasks to attain a learning goal are shown in the goal pane (Figure 8.1c), which is accessible by clicking on the goal from the achievement pane. For each task, an icon visualises the status (i.e., active or done) and type of task (e.g., game, diary). The icons match those of the activities in the *myPAL app* to support recognition and reduce cognitive load [31, p. 81], [32, pp. 132]. The child can choose to start the activity by clicking the task. If so, confirmation [32, pp. 54] is asked by the avatar to avoid accidental, unwanted navigation outside the *cPOD* [31, p. 19-20].



(a) start screen (b) achievement pane (c) learning goal pane

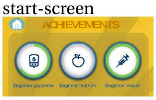
Figure 8.1: Examples of the *myPAL Objectives Dashboard* screens.

8.4. Usability Evaluation

Usability is a quality attribute assessing, from users' perspective, the ease of use of an interface [34] and also refers to qualitative and quantitative evaluation methods to identify any problem while using it [35]. The underlying aim is to remove those interaction problems (e.g., layout, contents, symbols, images) which cause any frustration [36]. However, when dealing with interfaces for children, particular attention has to be paid not only to the design, but also to the usability evaluation methods to be chosen [37]. In fact, children are not simply small adults, but individuals with their own perceptions and communication rules [38].

Method Following [39], to define the evaluation metrics for the *cPOD*, the *Goal-Question-Metric* (GQM) approach [40] was applied to ISO 9241-11 (1998/2018) usability characteristics of effectiveness, first-time learnability, understandability, and operability. The GQM-model is a paradigm to define and evaluate operational *design goals*, translating them into *questions*, which are quantifiable through both quantitative and qualitative *metrics* [41]. In the present study, we linked the selected usability characteristics (e.g., learnability) to specific *design goals* (e.g., cognitive load) and related *questions* (e.g., "To what extent does a child understand the status of his/her achievements?") and, lastly, pointed out the corresponding performance *metrics* (e.g., number of errors). A set of task-scenarios were created covering all possible usability aspects highlighted through the GQM. The usability test was structured as *task-oriented*; researchers followed a script guiding the user through the proposed activities, resembling real-life usage of and experience with the *cPOD*. Table 8.1 depicts a sample of a task and script[2].

Table 8.1
EXAMPLE OF A TASK SCRIPT AND ORIGINATING GQM

Characteristic	Goal	Question	Metric	*cPOD* location	Task	Script
First time learnability	Cognitive load	To what extend do children understand the display of status of achievements?	Success rate Free comments	start-screen 	**Task 3:** The user must be able to mention the status of each achievement based on the progress displayed by the circle surrounding each icon.	Looking at the three achievements, can you tell me which is: a) completed, b) partial, c) not completed? Can you tell me how you know?

Measures We took into account both objective and subjective measures: *the error rate*, defined as the number of participants failing usability-task execution, and *participants' verbal responses and researcher's notes* during task execution. These measures were marked on an usability checklist (see Table 8.2) by a researcher during the test. The checklist included all task-scripts with supposed user actions and predicted errors. Moreover, all screen interactions were recorded with a screen reader.

[2] To conserve space we provide only one example, the interested reader can find the full task scripts on: http://www.pal4u.eu/wp-content/uploads/Usability-test_tasks-details-and-results.pdf.

Table 8.2
EXTRACT OF THE USABILITY CHECKLIST FOR TASK 3

***cPOD* location**	**Script**	**Correct**	**Error**	**Case**	**Notes**
start-screen 	Looking at the three achievements, can you tell me which is: a) completed, b) partial, c) not completed? Can you tell me how you know this?			1. The child is unable to distinguish between the achievements’ status.	
	Do you notice any difference between the three achievements?			2. The child does not notice any difference between the circles surrounding the three achievements.	

Participants A total of 12 children with type 1 diabetes, 8 to 10 y.o., participated in the usability evaluation. Of which six were Italian (mean age=8.57, std.dev=0.79) and six were Dutch (mean age=9.5, std.dev=0.55), all participating at autumn camps (October 2017) organised in the context of the PAL project by two diabetes patients’ associations, which are part of the project consortium. As suggested by [42], a minimum number of eight participants was chosen to reliably reveal possible usability issues.

Materials In each session we used a Lenovo Tab2 tablet (www.lenovo.com) on which was installed the *myPAL app* 2.0 version, embedding the proposed *cPOD* prototype, and the AZ-screen-recorder app.

Procedure The activities took place at the camp premises. All 12 children participated in the usability evaluation once, individually, and for a duration of maximum 30 minutes. After that the session was ended, regardless of task-script and question completion. During each session a researcher (the *facilitator*) was responsible for all aspects of the study protocol, including initial greetings, guidance through the selected tasks, data collection, assisting and debriefing the participant as shown in Figure 8.2. At the start of each session the participant received a brief introduction regarding the aim of the study. Then, the child was given a tablet with the *myPAL app* with the *PAL Objectives Dashboard* displaying three predefined example achievements. These achievements were selected to visualise the different possible states: fully attained, active with partial progress, and active with no progress yet (see Figure 8.1a). Then, the child was asked to go through the *cPOD* and perform and comment on the different usability-tasks.

8

8.5. RESULTS AND DISCUSSION

We counted the error rates from the usability checklists and organised children’s subjective comments (if any), as well as facilitator’s notes or elaborations, through affinity diagrams clustered by Nielsen’s usability heuristics [43].

Consistency and standard issues: “users should not have to wonder whether different words, images, or actions mean the same thing.” Some graphic representations hindered ease of use leading children falsely thinking they were interactive. For example, the green-coloured progress bars in the achievement pane - task 6.6.

Figure 8.2: Position of child, tablet and researcher.

Match between system and the real world: "the system should speak the user's language and follow real-world conventions." Eight out of 12 participants did not understand that the *cPOD* was intended to reflect personal achievements activated in collaboration with a health care professional - task 2.1. This might, however, result from using pre-selected achievements for the evaluation rather then actual personal achievements. Further, only half of the children (6/12) understood the detailed information shown in the achievement pane - task 5.1. For example, the achievement label was evaluated as too generic by six participants, misleading them to refer to the overall achievement as to one of the linked goals. Lastly, the corresponding difficulty level, expressed through predefined labels (e.g., beginner) were evaluated as unclear by 5 children, who were unable to attribute meaning to it - tasks 5.4/6.

Visibility of system status issues: "the system should always keep users informed about what is going on." The achievement progress indicator (i.e., a circle surrounding the achievement icon progressively colouring green depending on progress (see Figure 8.1a) was correctly understood by all participants except one - task 3. Contrary, when exploring the goals for a specific achievement, the highlight indicating in the sidebar the achievement under investigation (i.e., a blue square surrounding the achievement icon; see Figure 8.1b) was ineffective. In fact, an 83% error rate (10/12 children) was scored in the corresponding task 5.2.

Documentation issues: "it may be necessary to provide help or documentation in the system." Less then half of the participants were able to explain or point out the list of tasks to do in the *myPAL app* in order to attain a learning goal - tasks 6.3. Further, none of the children understood 'their' current active difficulty level - task 6.4. And only 1/3 understood the next level to proceed - task 6.5.

8.5.1. RE-DESIGN OBJECTIVES DASHBOARD

The findings presented above informed recommendations improving the *cPOD* design. These have been implemented in the PAL3.0 System, released May 2018 and currently under evaluation.

Consistency and standard issues: Each goal element is now interactive, the user can click anywhere (description, bar, icon) to navigate to the goal detail-pane - task 6.6.

Match between system and the real world: Once entered in the achievement pane, the section header now reads "My Achievements" to clarify that these are personal achievements to be attained - task 2.1. At the top of the achievement pane is now clearly indicated the current-looked-at-achievement by display of icon, label (short name) and description (full sentence clearly stating the purpose of the objective) - task 5.1. The corresponding difficulty level labels have been rephrased to be more self-explanatory - tasks 5.6. And all possible difficulty levels are presented on top of the achievement pane, with the active level coloured yellow and attained levels green - task 5.4/6.4/6.5.

Visibility of system status issues: When clicking on an achievement to access its details (i.e., view description and goals), now only the achievement pane is displayed, without the list of the other achievements in the sidebar - task 5.2.

Documentation issues: To clarify tasks contributing to goal progress, the task-icon behind each goal is now labelled "tasks". Clicking anywhere on the goal item will navigate to the goal pane showing the details of the corresponding tasks - task 6.3. Additionally, some refinements were made to compensate for side effect of other changes in the design. To improve the readability of the information architecture contained in it, the start-screen shows two tabs. The primary tab displays the active achievements, attained achievements are accessible in the second tab. This is meant to ease the navigation in the case of various achievements activated, as to avoid for the children the drawback of having too much scrolling in the page [33]. Further, the achievement pane shows the text "Goals for this achievement" to clarify the meaning of goal-elements depicted right after this text. For each goal-element a full sentence description is given and the progress bar displays the text "*status of completion*" followed by the percentage.

8.5.2. DISCUSSION

Although the *PAL Objectives* have been developed to optimise children's adherence to the *myPAL app*, the present evaluation was limited towards usability. It is necessary to assure that possible ineffectiveness does not result from usability issues. In fact, if children have difficulties understanding their goals and achievements these are unlikely to motivate them to engage in *myPAL app* activities. To investigate the effect of the *PAL Objectives* on children's motivation using the *myPAL app*, we collect data on perceived user experience and actual system usage. We compare between an ongoing study including the refined *cPOD* and and an earlier version of the *myPAL app* presenting objectives in an unsorted list. Preliminary usage statistics, for Dutch participants only, show that average duration of using the *cPOD* increased in the ongoing study (~ 4 min.) compared to the previous version (~ 1 min.). This may indicate that the *cPOD* provided more interesting information and interactions then a simple list of objectives. Comparison with an earlier version of the *myPAL app*, without the sophisticated *cPOD*, may provide some insights towards the usefulness of the *cPOD*. However, possible differences are not necessarily caused by the *cPOD* because many other features of the app were changed in parallel. A controlled study comparing motivation for and usage of an app between a group with and without the *cPOD* is needed to base any conclusion on the effectiveness of the *PAL Objectives* system.

8.6. CONCLUDING GUIDELINES

The main challenge addressed in the present work is the design of an dashboard to display personalised diabetes self-management educational achievements and learning goals to children in order to increase their usage of a mHealth app. To this aim, the *myPAL Objectives Dashboard* was designed following common interface design patterns and evaluated with diabetic children. The colouring indicating status was well understood. Children were able to navigate between layers of information, but did not necessarily understood the layered information. Children experienced difficulties interpreting the meaning conveyed in iconic presentations. Based on reported usability issues, we present additional guidelines for the design of a dashboard for children (see Table 8.3).

Table 8.3
USABILITY ISSUES AND SOLUTIONS FOR THE *cPOD*, AND DERIVED GUIDELINES FOR THE DESIGN OF A OBJECTIVES DASHBOARD.

Identified Usability Issue	**Design Refinement**	**Guideline**
(false) Expectation of interactivity of the progress bar	Each goal-element is fully interactive, navigating to goal details	3. Ease navigation between layers
Unaware that personal objectives are displayed	The *cPOD* heading includes the word "my"	1. Provide descriptive labels
Lack of understanding of achievement information	The achievement pane header shows the icon, label and description The goal are preceded by "goals for this achievement"	2. Connect elements accordingly 1. Provide descriptive labels
Lack of understanding of the level of viewed achievement Lack of understanding of the active difficulty level Lack of understanding of future difficulty levels	All possible difficulty levels for the achievement are listed and coloured according to their status (i.e., green for attained, yellow for active)	Provide all information necessary to interpret meaning
Lack of understanding of the difficulty levels	Revised labels and added colouring	1. Provide descriptive labels
Unable recognise achievements in the sidebar	Removed the sidebar, limit navigation between achievements via the start-screen	3. Ease navigation
Lack of understanding of task information	The task-icon is accompanied with the label "tasks"	1. Provide descriptive labels

Guideline 1: provide descriptive labels. Even though children may tend to not read texts [33], visual elements may not speak enough for themselves . Complementary, clear, short labels can help a child to understand elements. For example, an icon navigating to task details should be clarified with the label "tasks". Specifically for complex information structures, such as the *POM*, relations between items must be clarified. For example, by placing descriptive labels before displaying information such as "your current level:" and "goals for this achievement:".

Guideline 2: connect elements accordingly. When elements are separated their connection is not self-evident. Placing visual elements in close proximity of and in the same 'box' as descriptive labels strengthens their connection. For example, placing the icon of the selected achievement inside the achievement pane supports understanding that details of this achievement are displayed. For children, the understandably of the connection between elements (i.e., the icon and details of a single achievement) prevails the need for quick and easy access to other elements (in this case other achievements).

Guideline 3: ease navigation between layers. To avoid cluttering and cognitive overload information needs to be layered and more detailed information should only be given on request. Requesting, however, must be facilitated easily, intuitively, and at the right moment. Moreover, requests that lead to the same information should be given by a single action. In the *cPOD* case, from the achievement pane, for each goal, all sub-elements (i.e., description, progress, task icon, and the box surrounding them) should be interactive and navigate to the goal pane. Further, unclear, cluttering elements such as menu bars should be avoided.

REFERENCES

[1] R. Peters, E. Oleari, F. Sardu, and M. A. Neerincx, *Usability of the pal objectives dashboard for children's diabetes self-management education,* in *Proceedings of the 2019 the 5th International Conference on e-Society, e-Learning and e-Technologies* (ACM, 2019) pp. 22–28.

[2] D. H. Schunk and B. J. Zimmerman, *Motivation and Self-Regulated Learning: Theory, Research, and Applications* (Routledge, 2008).

[3] A. Jones, K. Dennis, C. A. Basedow, P. Alves-oliveira, S. Serholt, H. Hastie, L. J. Corrigan, W. Barendregt, A. Kappas, A. Paiva, and G. Castellano, *Empathic Robotic Tutors for Personalised Learning : A Multidisciplinary Approach,* in *Soc. Robot.*, Vol. 1 - 9388 (Springer, Paris, 2015) pp. 285—-295.

[4] D. Leyzberg, S. Spaulding, and B. Scassellati, *Personalizing robot tutors to individuals' learning differences,* Proc. 2014 ACM/IEEE Int. Conf. Human-robot Interact. - HRI '14 , 423 (2014).

[5] O. A. B. 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 **92**, 174 (2013).

[6] B. R. Schadenberg, M. A. Neerincx, F. Cnossen, and R. Looije, *Personalising game difficulty to keep children motivated to play with a social robot: A bayesian approach,* Cognitive systems research **43**, 222 (2017).

[7] I. Leite, C. Martinho, and A. Paiva, *Social robots for long-term interaction: a survey,* International Journal of Social Robotics **5**, 291 (2013).

[8] J. K. Njenga and L. C. H. Fourie, *The myths about e-learning in higher education,* British journal of educational technology **41**, 199 (2010).

[9] L. De-Marcos, A. Domínguez, J. Saenz-de Navarrete, and C. Pagés, *An empirical study comparing gamification and social networking on e-learning,* Computers & Education **75**, 82 (2014).

[10] E. A. Locke and G. P. Latham, *Building a practically useful theory of goal setting and task motivation: A 35-year odyssey.* Am. Psychol. **57**, 705 (2002).

[11] C. Ames, *Classrooms: Goals, structures, and student motivation.* (1992).

[12] P. Black and D. Wiliam, *Assessment*, Vol. 5 (2006) pp. 37–41, arXiv:0402594v3 [arXiv:cond-mat] .

[13] M. Cabrita, H. op den Akker, R. Achterkamp, H. J. Hermens, and M. M. Vollenbroek-Hutten, *Automated personalized goal-setting in an activity coaching application.* in *SENSORNETS* (2014) pp. 389–396.

8

[14] K. M. Law, V. C. Lee, and Y.-T. Yu, *Learning motivation in e-learning facilitated computer programming courses,* Computers & Education **55**, 218 (2010).

[15] S. Rollnick, W. R. Miller, and C. C. Butler, *Motivational Interviewing in Health Care: Helping Patients Change Behavior* (The Guilford Press, New York, 2008).

[16] E. L. Deci, R. J. Vallerand, L. G. Pelletier, and R. M. Ryan, *Motivation and education: The self-determination perspective,* Educational Psychologist **26**, 325 (1991), https://doi.org/10.1080/00461520.1991.9653137 .

[17] R. M. Ryan and E. L. Deci, *Intrinsic and extrinsic motivations: Classic definitions and new directions,* Contemporary Educational Psychology **25**, 54 (2000).

[18] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, *From game design elements to gamefulness: defining gamification,* in *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (ACM, 2011) pp. 9–15.

[19] J. Hamari, J. Koivisto, and H. Sarsa, *Does gamification work?–a literature review of empirical studies on gamification,* in *2014 47th Hawaii international conference on system sciences (HICSS)* (IEEE, 2014) pp. 3025–3034.

[20] J. Koivisto and J. Hamari, *Demographic differences in perceived benefits from gamification,* Computers in Human Behavior **35**, 179 (2014).

[21] R. M. Ryan and E. L. Deci, *Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being.* American psychologist **55**, 68 (2000).

[22] J. F. Bryan and E. A. Locke, *Goal setting as a means of increasing motivation.* Journal of applied psychology **51**, 274 (1967).

[23] M. Koo and A. Fishbach, *Dynamics of self-regulation: How (un) accomplished goal actions affect motivation.* Journal of personality and social psychology **94**, 183 (2008).

[24] R. Kleinrahm, F. Keller, K. Lutz, M. Kölch, and J. Fegert, *Assessing change in the behavior of children and adolescents in youth welfare institutions using goal attainment scaling,* Child and Adolescent Psychiatry and Mental Health **7** (2013).

[25] E. A. Locke and G. P. Latham, *New directions in goal-setting theory,* Current directions in psychological science **15**, 265 (2006).

[26] L. S. Vygotsky, *Mind in society: The development of higher psychological processes* (Harvard university press, 1980).

[27] H.-u. Krieger, R. Peters, B. Kiefer, M. A. V. Bekkum, F. Kaptein, and M. A. Neerincx, *The Federated Ontology of the PAL Project Interfacing Ontologies and Integrating Time-Dependent Data,* in *8th Int. Jt. Conf. Knowl. Eng. Ontol. Dev.* (SciTePress, 2016).

[28] R. Peters, J. Broekens, and M. A. Neerincx, *Guidelines for tree-based learning goal structuring,* in *Proceedings of the 22Nd International Conference on Intelligent User Interfaces, IUI*, Vol. 17 (ACM, Limassol, Cyprus, 2017) pp. 401–405.

[29] M. Van Welie, G. C. Van der Veer, and A. Eliëns, *An ontology for task world models,* in *Design, Specification and Verification of Interactive Systems' 98* (Springer, 1998) pp. 57–70.

[30] 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 child's diabetes self-management,* (2016).

[31] J. Preece, Y. Rogers, and H. Sharp, *Interaction design: beyond human-computer interaction* (John Wiley & Sons, 2015).

[32] W. Lidwell, K. Holden, and J. Butler, *Universal principles of design, revised and updated: 125 ways to enhance usability, influence perception, increase appeal, make better design decisions, and teach through design* (Rockport, 2010).

[33] J. Nielsen, *Nielsen norman group - children's websites: Usability issues in designing for young people,* (2010).

[34] J. Nielsen, *Usability Engineering* (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993).

[35] J. Nielsen, *Nielsen norman group - introduction to usability,* (2012).

[36] M. Soegaard, *Usability: A part of the user experience,* (2018).

[37] S. Gilutz and J. Nielsen, *Usability of websites for children: 70 design guidelines,* (2002).

[38] M. A. Khanum and M. C. Trivedi, *Take care: A study on usability evaluation methods for children,* CoRR **abs/1212.0647** (2012), arXiv:1212.0647 .

[39] R. Tahir and F. Arif, *Framework for evaluation the usability of mobile educational applications for chidlren,* Society of Digital Information and Wireless Communications , 156 (2014).

[40] A. Hussain and E. Ferneley, *Usability metric for mobile application: A goal question metric (gqm) approach,* in *Proceedings of the 10th International Conference on Information Integration and Web-based Applications & Services*, iiWAS '08 (ACM, New York, NY, USA, 2008) pp. 567–570.

[41] van Solingen Rini, B. Vic, C. Gianluigi, and R. H. Dieter, *Goal question metric (gqm) approach,* in *Encyclopedia of Software Engineering* (American Cancer Society, 2002) https://onlinelibrary.wiley.com/doi/pdf/10.1002/0471028959.sof142 .

[42] N. Bevan, *Practical issues in usability measurement,* interactions **13**, 42 (2006).

[43] J. Nielsen, *Nielsen norman group - 10 usability heuristics for user interface design,* (1995).

8

9

Conclusion

This thesis examines stylised behaviour of socially interactive agents and personalised content in intelligent tutoring systems in respectively part I and II. Regarding the expression of pedagogical interaction style by a robot, as discussed in Part I, we found the following: First, previously validated perception with considerable support is limited to personality trait extraversion, and positive or negative emotion. Second, in perception studies we showed that expression of dominance is possible by manipulation of gesture and posture expansiveness, and that warmth and competence modelled in parallel in gesture, posture and prosody can be perceived even by young children based upon short observation or interaction, but that perception is influenced by context (the robot was perceived more competent at school than at a camp). Finally, we did not find any effect of different pedagogical interaction styles of a robot on learning outcomes, while observing children not looking at the robot. With respect to the personalisation of learning content as discussed in Part II we found the following. First, learning objectives are a valuable means to present the child's developmental levels and interest to both human users and the system. Second, objectives can be structured domain independently and this structure can be filled with domain specific knowledge generated in a multidisciplinary team. Third, specific user-interfaces design for either the learner (child), supervisor (parent) and educator (health care professional) can communicate the objectives and/or tasks. Finally, we found that clear understanding of the objective (goal) displays and an abundant amount of learning content are crucial to the success of long-term engagement with the system.

9.1. Findings

9.1.1. Part I

The main research question for part I is: *How to design SIA behaviours that express different pedagogical styles and what is the effect on learning outcomes?*

Research question 1; Chapter 2
What are the possible parameter-based manipulation of bodily shape and motion for social-affective factors and how are these perceived by humans?

We found that research in this field is multidisciplinary and shows a large variety in concept definitions, behavioural manipulations and evaluation methodologies. Our TAX-MOD taxonomy is a starting point to develop a shared understanding and interpretation of research objectives and outcomes, to position and compare research, to explicate progress in this area, and to formulate a road-map. We report the following key findings as guidelines for future behaviour design: 1) the expression of personality traits using virtual or robot bodies seems limited to the trait extraversion; 2) the expression of social dimensions seems possible, but only when using the whole body, and more research is needed to disentangle individual effects on friendliness, competence and dominance; 3) the expression of emotion seems possible but restricted to generic positive versus negative signals; and 4) context seems important for users for the correct interpretation of the expressive behaviour.

Research question 2; Chapter 3
To what extent do humans perceive warmth and competence expressions of a NAO robot?

We showed that style expression by a humanoid robot can be perceived by (young) children. Bodily posture, hand gestures and paralinguistic cues can be manipulated to evoke a desired expression of competence. Warmth manipulations had the intended effect only for high-competence robots. Moreover, context influenced children's perceptions: at school the robot was perceived warmer and more competent than at camp.

Research question 3; Chapter 4
To what extent do humans perceive dominance expressions of a NAO robot?

We have shown the validity of body expansiveness modulation for dominance expression in both postures and gestures and proven that with a limited set of parameters we can moderate dominance expression. We found that specific postures and gestures have a natural tendency towards being perceived as more or less dominant. Moreover, the manipulation effect was consistent across various behaviours, except for a sitting pose. This study provides evidence that body expansiveness is an important factor for dominance expression and that this effect is behaviour independent and independent of view angle.

Research question 4; Chapter 5
What is the effect of pedagogical interaction style (Expert and Facilitator) shown by a NAO robot on children's a) learning approach, and b) learning gain?

We did not find any effect of robot interaction style on children's learning approach or gain. Based solely on verbal behaviour variations, children perceived the explaining robot as more competent than the robot providing evidence descriptions; however, this perception difference did not impact learning approach or gain in the present study. No evidence for differences in perception based on non-verbal behaviour only was found. We did find that the presence of a robot giving feedback on children's rolling trials did cause children to play longer and do more informative experiments compared to absence of robot feedback during play, however, this difference in learning approach did not impact learning gain from the rolling trials. We conclude that the task selection and attention towards the robot is critical: the task shouldn't compete for the robot's attention and it must be challenging enough to show a learning effect (avoid ceiling effect).

To answer the main research question (*How to design SIA behaviours that express different pedagogical styles and what is the effect on learning outcomes?*) we conclude that expression of pedagogical interaction style can be achieved by varying the expression of friendliness, competence and dominance. Each dimension is associated with certain non-verbal behavioural modulations. We present evidence that dominance expression is possible by modulation of bodily expansiveness. We did not find evidence that the non-verbal behavioural style influences learning gain or outcomes. We argue that attention towards the robot was limited. To develop effective pedagogical interaction style study of long-term engagement is needed. Detection of attention and attention loss, and strategies to cope with this loss are needed. Further research is needed to answer this second part of the main research question.

9.1.2. Part II

The main research question for part II was: *How to personalise learning content based on personal learning objectives?*

Research question 1; Chapter 6
How to model learning objectives and implement these in a intelligent tutoring system?

Effective learning goals must attune to the appropriate level, a way to structure this is Bloom's taxonomy [1, 2]. Further, a learning goal must have attributes presenting relations and descriptions (e.g., [3–6]). Formalisation in an ontology supports shared understanding and automated reasoning [7]. In our upper ontology we structured the objective classes (i.e., achievements, learning goals and accompanying tasks), relations between them and defined domain independent parameters (i.e., level and topic). Our upper ontology adequately supported the formalisation of implicit knowledge of health care professionals on diabetes self-management training. For the formalisation of these domain specific learning goals, achievements, tasks and materials the following design guidelines have been recommended: work in a multidisciplinary team (include domain- and pedagogic experts next to knowledge engineers to make an inventory of important learning goals and define learning activities); formulate achievements from logical learning units (e.g., daily challenges) that require a subset of the knowledge and skills encapsulated in the goals to improve relevance; formulate achievements and goals from the perspective of the child to facilitate ownership and increase experienced relevance; and define user characteristics relevant to goals and/or tasks selection. We implemented the ontology successfully in an multi-modal (robot, tablet, and web applications) system providing personalised learning content. The system was developed iteratively, and both the model and interfaces were refined based on user evaluations (see Chapter 7 and 8). We showed that our ontology contributed to the support of children's basic needs (autonomy, competence, and relatedness) [8]. The ontology and its implementation supported all children in goal attainment, however, not all children were supported equally. This difference is likely related to the degree in which the children consciously used the system and understood the system. For the integration of the ontology in a multi-modal intelligent tutoring system the following design guidelines have been recommended: provide instruction and explanation to the child on how achievements, goals and tasks are selected and can be attained (i.e., that progress on a goal is gained by task completion and benefits earned by this); embed the objectives in the ITS application by making them easily accessible to the (child) user and integrate them in other system functionality such as feedback provided by an pedagogical agent; and offer enough learning content (games, quizzes) to maintain interest and engagement.

Research question 2; Chapter 7
How to design a user interface for health care providers representing a domain specific knowledge base providing a structure for learning objectives to facilitate personal goal setting and progress monitoring?

We found that tree-based presentation of learning objectives is effective for health care professionals to set personal goals and monitor progress. The following design guidelines for an authoring tool have been proposed: provide clear, visual feedback on goal structure, and active state and progress; consistently use a different representation (e.g., shape) for different concepts of the model (e.g., goal and achievement); cover the full domain and different skill levels in the goal content; and support assessment of current abilities next to goal setting, progress monitoring and attainment registration.

Research question 3; Chapter 8
How to design a user interface for children that integrates educational objectives with gamification features to advance the interests and engagements of children?

We found that the following design elements were well understood: colouring indicating status; and navigation between layers of information. Children experienced difficulties in: interpreting the meaning conveyed in iconic presentations; understanding layered information; and navigating the interface. Based on reported usability issues, we presented guidelines for the design of a dashboard for children: provide descriptive labels next to visual elements because children lack experience using apps and thus understanding of icons and such; connect elements accordingly by placing them in close proximity and in boxes with appropriate labels; and ease navigation between layers when hiding detailed information to avoid cluttering and cognitive overload and avoid cluttering elements such as navigation bars.

To answer the main research question (*How to personalise learning content based on personal learning objectives?*) we conclude that content can be personalised by personalisation of learning objectives that are related to certain tasks and content. For formalisation of objectives we looked into theories from the education field. For development of the domain specific knowledge base of learning objectives and content both existing checklist and implicit expert knowledge have be accessed. To be effective, learning objectives should carry personal relevance and progress must be visible. Personal relevance can be achieved by both pre-selection based on user characteristics (such as device usage in the case of diabetes self-management education) and personal goal setting in collaboration with a care provider. One of the main challenges to foster long-term engagement is the inclusion of sufficient and challenging content.

9.2. Overall Contribution

In this thesis, we have made a first step towards the design of pedagogical interaction styles for socially interactive agents (SIAs) and towards the formalisation of learning objectives for personalisation of learning content, all in intelligent tutoring systems (ITSs). The designs were implemented and evaluated with the intended user groups in smaller studies and 'in the wild' in a three month long study evaluating the PAL system with children with T1DM and their caretakers. Our research shows that it is possible to design stylised behaviours perceived as intended by children aged from 8 to 13, implement these SIA behaviours into an ITS, and select behaviours based on knowledge of the learning activity. Further, we show that a domain independent ontology of learning objectives and content can be filled with domain specific knowledge and can be implemented in an ITS to select personalised learning content based on specific user knowledge (i.e., the user's developmental levels, as reflected by the current state of objectives and disease management preferences). We conclude that stylised behaviours can express the desired level of competence, friendliness and dominance, thereby making it possible to express different teaching styles (see Chapters 3 and 4). However, to have an effect on learning outcome these behaviours must be noticed by the child, meaning that the SIA must be the centre of attention (see Chapter 5). For a SIA-ITS to be motivating and engaging in the long-term, full coverage of the domain and an abundance of learning content in various learning activities is necessary (see Chapter 6).

9.3. Limitations

Existing work on behaviour manipulations is multidisciplinary and therefore lacks a standardised vocabulary or methodology. Further, conflicting results arise between individual studies. Another major limitation in existing work is that different parameters are often manipulated in parallel, making it impossible to disentangle the effect of individual parameters. We have attempted to overcome these challenges by relating existing studies to our TAXMOD, however, this required us to abstract away from many details in existing studies on both the bodily manipulation and outcome measures.

The NAO robot and it's avatar have limited non-verbal expression capabilities. The NAO robot does not have facial expressions, therefore we could only use bodily cues.

The NAO robot does have colour leds, but existing research on the expressive capabilities of these leds is very limited. Also the NAO robot is restricted in its spacial movement, meaning that it walks very slow and unstable and larger movements may be unstable in general. This results in a limited repertoire for behaviour manipulation.

Evaluations were limited in participant number and in interaction. In the perception studies not many children/participants participated. In the effect study we had a total of 73 children participating, but due to the many variables a maximum of 20 children participated in the unique conditions. The perception studies involved observation rather than interaction scenarios. Also, in the effect study the robot was giving (passive) guidance in a WOz setup rather than having a real conversation with the child. The behaviours were implemented in the PAL system and thereby part of the long-term study. However due to the many variables in this system and limited usage of the system we could not derive enough data to study the effect of robot behaviours on learning outcome in this study. (This is why we set up an effect study in a different, more controlled context.) In the PAL study we did evaluate the learning objectives ontology and its user interfaces. Here again the number of participants was low and overall usage of the system was limited.

Questionnaires for children are limited. There are existing toolkits[1], but these focus on concepts such as likeability rather than perception of social factors. Further, even more so then adults children experience difficulties filling in questionnaires [9]. Therefore we had to develop our own questionnaire. Even though we did this thoroughly and carefully it is not a standardised instrument.

Inclusion of minorities and cultural differences were not considered in the present studies. All studies were conducted mostly with western participants. Chinese participants were included in the study in chapter four, however, these were all participants with a high economic status studying in the Netherlands. Therefore, we cannot draw any conclusions on the applicability of results in non-western cultures and developing countries. Further, it is well known that neurodivergent children perceive and learn differently. Even though these children have been included as participants we have not balanced or controlled for this factor.

9.4. Future Work

The TAXMOD taxonomy of behaviour modulations needs to be discussed with experts in related fields in order to create a shared understanding and vocabulary between disciplines. We proposed a first taxonomy based on the vocabulary presented in studies included in our literature review. The current taxonomy is simplistic. Creating an advanced, generally accepted taxonomy would help the field to provide comparable results, and further and faster advance knowledge about the effect of behaviour modulations on how people perceive a robot or virtual avatar.

[1] https://www.puppyir.eu/results/user-evaluation-toolkit

The effect study was limited in the interaction scenario and thereby attention towards the robot. We would like to see follow-up studies that include richer and possibly prolonged interactions so that the child pays more attention to the robot or is exposed to it more often. This way the effects of different interaction styles become more visible. This would require both engaging activities and activities where there is visual exposure towards the SIA. In the effect study presented in Chapter 5 we have chosen a between-subject set-up and one single activity. A study including multiple activities and within-subject set-up such as the PAL study [10] will allow us to study the suitability of various styles for certain types of activities. However, care should be taken not to include too many other variables as was the case in the PAL study. To get reliable answers focus of the study must be the style manipulations. One possibility for the future is to include this type of studies at a larger scale which is possible with the growing usage of ITSs and robots in classrooms and hospitals. Further, the effect study in Chapter 5 was done in a classroom or after school care setting, and not in the diabetes domain. General principles can be evaluated with children in general but to find conclusive results for implementation in a system for a specific user group (i.e., children with T1DM) the effect of stylised behaviours should be tested with the intended user group in the intended ITS.

Available learning content was limited in the present studies. In the current implementation of the PAL system often the number of possible tasks for selected objectives and the amount of content per task were minimal. This declines prolonged interaction and motivation to use the system. An increased number of learning activities (i.e., tasks and accompanying materials such as questions and carddecks) is necessary to maintain interest and engagement. Likewise, the acquisition and formalisation of domain specific knowledge elements is a time consuming human effort which might be supported by better tooling. The currently used WiSCE Confluence page appeared still rather 'technical' for experts to directly create the domain model content. User friendly tooling for 'filling' the domain independent learning objectives model is needed.

The learning objectives ontology was used to select learning tasks, determine SIA behaviour style, and provide progress feedback to the child. The user model, specifically current state of knowledge, was entered manually. Having user assessment of current knowledge embedded in the system both provides possibilities to keep learning activities relevant for the child and lowers the burden on human assessors. Of course human monitoring remains necessary but the system can support this task. For example, a regular scheduled, repeated quiz can be used for assessment of cognitive knowledge on the level of remembering and understanding, while conversations or physical tasks can be used for more complex skills. The challenge is to design activities that can be supervised automatically —instead of the trust-based manual check implemented now for real-world tasks.

9.5. A Final Note

This thesis presents the research done on two topics as part of the larger PAL project. Besides these directed studies participation in various studies —amongst presence in multiple week-long diabetes camps —within the PAL project offered a lot of insight into how children interact with a social robot and mHealth app. These insights did not contribute to the research questions in this thesis but are nonetheless valuable. One of the (perhaps a bit frustrating) learnings is that children in general are more enthusiastic about the robot performing a dance than the carefully crafted learning experiences. (Note should be taken that this interest wears of quickly, the 'novelty effect'.) Further, younger children's imagination makes them create a whole personality and range of behaviours for the robot. These intrinsically motivating experiences can and should be exploited when designing child-robot experiences. (Care should be taken never to show them a decapitated robot.) Children are willing to participate in all our studies and are even so willingly to talk to us researchers outside these formal moments. Within these moments valuable opinions and insights into their minds are shared. For example, while a child answered in the post-experiment questionnaire to consider one robot a father rather than a friend, in an informal talk he mentioned his father to be his best friend. This, of course, should change how we interpret results of the questionnaire. Doing research with children is challenging and needs a lot of consideration and perhaps more trial and error than was done in this thesis, but above all it is fun and full of possibilities to advance the field of SIA-ITSs.

References

[1] B. Bloom, *Taxonomy of educational objectives. Vol. 1: Cognitive domain,* (1956).

[2] D. R. Krathwohl, *A revision of bloom's taxonomy: An overview,* Theory into practice **41**, 212 (2002).

[3] S. Ebbens and S. Ettekoven, *Effectief leren* (Noordhof uitgevers BV, 2005).

[4] R. F. Mager, *Preparing instructional objectives. 2d. rev. ed,* (1984).

[5] H. Chung and J. Kim, *An ontological approach for semantic modeling of curriculum and syllabus in higher education,* International Journal of Information and Education Technology **6**, 365 (2016).

[6] L. Ng, *Ontological model for representation of learning objectives,* Ph.D. thesis, Interactive Arts and Technology-Simon Fraser University (2005).

[7] S. Staab and R. Studer, *Handbook on ontologies* (Springer Science & Business Media, 2010).

[8] M. Neerincx, W. Vught, O. Blanson Henkemans, E. Oleari, J. Broekens, R. Peters, F. Kaptein, Y. Demiris, B. Kiefer, D. Fumagalli, *et al.*, *Socio-cognitive engineering of a robotic partner for child's diabetes self-management,* Frontiers in Robotics and AI **6**, 118 (2019).

[9] M. Ligthart and R. Peters, *The challenges of evaluating child-robot interaction with questionnaires,* What Could Go Wrong (2018).

[10] F. Kaptein, B. Kiefer, A. Cully, O. Celiktutan, B. Bierman, R. Rijgersberg-Peters, J. Broekens, W. Van Vught, M. Van Bekkum, Y. Demiris, *et al.*, *A cloud-based robot system for long-term interaction: Principles, implementation, lessons learned,* ACM Transactions on Human-Robot Interaction (THRI) **11**, 1 (2021).

ACKNOWLEDGEMENTS

All good things must come to an end.

Finally, here it is: my dissertation. Enthusiastic, eager —and possibly a bit naive— I started this challenge in 2015. The plan became clearer, but the road long. For four years I enjoyed working at TU Delft and TNO and with the PAL team. I loved the collaborations, experiments, and hackathons we did within the PAL project. In hindsight I would have planned it differently because following my passions resulted in a lot of work done that did not end up in this thesis and a lot of work for this thesis done after the project and contract ended. Nonetheless, for a long time I regarded this as the best time of my life. I worked long days and seemingly never got tired. Until I did: this PhD, finishing my teaching degree, caring for my grandfather, and other family circumstances eventually got to me, and I could no longer hold up the fight against depression and burnout... Somewhere in between there and then the best time of my life changed into one of the worst. For a while I doubted if I would ever make it, but I was too stubborn to quit. Who would have guessed that getting a baby and a new job would be the solution to this burnout? Though not the best ingredients to finish a PhD thesis, I did regain my energy and with many small steps I did progress and finally finished this dissertation.

Of course I didn't do all this alone. Firstly, I want to thank my promotors Mark Neerincx and Joost Broekens who despite everything always kept believing we could bring this to a successful end. Sometimes the trust you put in me was hard for me because I felt insecure and wished for more guidance. But, you were right in the end. Independent committee members, thank you for taking the time to review this dissertation.

This work heavily relies on the PAL project, I would like to take this space to thank all the PAL project members for the fun and inspiring meetings, collaborations, and experiments. With the risk of forgetting someone I would specifically like to thank Bert for his help in programming the NAO robot and Bernd and Uli (who is sadly no longer with us) for introducing me into the world of ontologies. Thank you Willeke and Michael for your contributions to the final POM ontology, Elettra and Diego for the myPAL design and study and the Italian recipes, and Frank, Sylvia, and Olivier for the pleasant time during experiments in the hospitals and camps.

The experiments would never have been possible without the participation of the schools, daycare, hospitals and diabetes association. Many thanks for your dedication and time. Miriam, you were truly important for the success of the learning objectives and tasks. Your dedication to the project is admirable and inspirational.

For five years I did all this work at TU Delft and one day a week at TNO. I would like to thank all my colleagues —who I won't be naming in fear of missing out someone— there for making this a pleasant time. Thank you for the chats, coffee breaks, and lunch walks! There are two persons special to me I do wish to mention: Ruud who was always present to help with the NAO, server or any technical as well as personal issues. It was truly a

pleasure to meet and work with you! Rosemarijn, my Thursday colleague and friend. The many hours of chatting about the PAL project, promotion, and life in general may have been key to keeping my sanity through it all :)

Voor mijn familie is dit hele proces lastig te volgen geweest. Opa, je steunde me waar je kon ook al had je soms geen idee waar ik mee bezig was. Helaas maak je dit moment niet meer mee, ik weet zeker dat je de trotste ooit geweest zou zijn en zou opscheppen tegen iedereen in het dorp. Mams, bedankt voor het vele oppassen! Mimi, Puckie, Kayla en Loeki, bedankt voor jullie gezelschap en warmte.

Mijn kinderen, Ravi en Aurora. Jullie kwamen er pas later bij, maar zijn de sleutel geweest om weer op te krabbelen en mijn proefschrift af te kunnen schrijven. Ravi, sorry dat mama soms moest werken in het weekend en de vakantie en niet met jou kon spelen, dat is nu voorbij. Aurora, jij hebt van dit alles weinig meegekregen maar bent wel de kracht achter het laatste loodje! Jullie zijn mijn alles <3

List of Publications

Published work from the PAL-project

15. **R. Rijgersberg-Peters, T.J.P. van Schijndel, J. Broekens, and M.A. Neerincx**, *Effect of Robot Interaction Style on Children's Learning Approach and Cognitive Outcomes,* Submitted to International Journal of Social Robotics.

14. **R. Rijgersberg-Peters, J. Broekens, M.A. Neerincx**, *Parameter-based Modulation of Bodily Shape and Motion in Socially Interactive Agents: A Taxonomy and Review*, Submitted to ACM Transactions on Human-Robot Interaction.

13. **R. Rijgersberg-Peters, W. van Vught, J. Broekens, M.A. Neerincx**, *Goal Ontology for Personalized Learning and its Implementation in Child's Health Self-Management Support,* IEEE Transactions on Learning Technologies (2023).

12. **F. Kaptein, B. Kiefer, A. Cully, O. Celiktutan, B. Bierman, R. Rijgersberg-Peters, J. Broekens, W. Van Vught, M.A. Van Bekkum, Y. Demiris, and M.A. Neerincx**, *A cloud-based robot system for long-term interaction: Principles, implementation, lessons learned*, ACM Transactions on Human-Robot Interaction (2021).

11. **R. Peters, J. Broekens, K. Li, M.A. Neerincx**, *Robots expressing dominance: Effects of behaviours and modulation*, 8th International conference on affective computing and intelligent interaction (2019).

10. **R. Peters, J. Broekens, K. Li, M.A. Neerincx**, *Robot dominance expression through parameter-based behaviour modulation*, Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents (2019).

9. **M.A. Neerincx, W. Van Vught, O. Blanson Henkemans, E. Oleari, J. Broekens, R. Peters, F. Kaptein, Y. Demiris, B. Kiefer, D. Fumagalli, and B. Bierman**, *Socio-cognitive engineering of a robotic partner for child's diabetes self-management*, Frontiers in Robotics and AI (2019).

8. **R. Peters, E. Oleari, F. Sardu, and M.A. Neerincx**, *Usability of the PAL Objectives Dashboard for Children's Diabetes Self-Management Education*, Proceedings of the 5th International Conference on e-Society, e-Learning and e-Technologies, (2019).

7. **M. Ligthart, R. Peters**, *The Challenges of Evaluating Child-Robot Interaction with Questionnaires*, What Could Go Wrong workshop, (2018).

6. **R. Peters, J. Broekens, and M.A. Neerincx**, *Robots Educate in Style: The Effect of Context and Non-verbal Behaviour on Children's Perceptions of Warmth and Competence*, 26th IEEE international symposium on robot and human interactive communication, (2017).

5. **R. Peters, J. Broekens, and M.A. Neerincx**, *Guidelines for Tree-based Collaborative Goal Setting*, Proceedings of the 22nd International Conference on Intelligent User Interfaces, (2017).

4. **H. Krieger, R. Peters, B. Kiefer, M.A. Van Bekkum, F. Kaptein, and M.A, Neerincx**, *The Federated Ontology of the PAL Project. Interfacing Ontologies and Integrating Time-Dependent Data*, IC3K 2016: 8th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, (2016).

3. **M.A. van Bekkum, H. Krieger, M.A. Neerincx, F. Kaptein, B. Kiefer, R. Peters, and S. Racioppa**, *Ontology Engineering for the Design and Implementation of Personal Pervasive Lifestyle Support*, SEMPDS-2016 poster&demo at SEMANTiCS2016, (2016).

2. **M.A. Neerincx, F. Kaptein, M.A. Van Bekkum, H. Krieger, B. Kiefer, R. Peters, J. Broekens, Y. Demiris, and M. Sapelli**, *Ontologies for social, cognitive and affective agent-based support of child's diabetes self-management*, workshop on AI for Diabetes at ECAI2016, (2016).

1. **R. Peters, J. Broekens, and M.A. Neerincx**, *Let Me Guide You! Pedagogical Interaction Style for a Robot in Children's Education*, WONDER workshop at ICSR2015, (2015).

Other Publications

4. **R. Peters, D. Plass-Oudebos**, *Identifying and addressing common programming misconceptions with Variables-Part II*, Master Thesis (2018)

3. **R. Peters**, *How turn-taking influences the perception of a suspect in police interviews*, Master Thesis (2014)

2. **R. Peters, M. Bruijnes, R. op den Akker**, *How Turn-Taking Influences the Perception of a Suspect in Police Interviews*, Proceedings of Chi Sparks (2014).

1. **R. op den Akker, M. Bruijnes, R. Peters, T. Krikke**, *Interpersonal stance in police interviews: content analysis*, Computational linguistics in the Netherlands journal (2013)