

ISSN 2499-4553

IJCoL

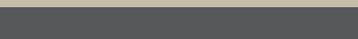
Italian Journal
of Computational Linguistics

Rivista Italiana
di Linguistica Computazionale

Volume 7, Number 1-2
june-december 2021
Special Issue

Computational Dialogue Modelling:
The Role of Pragmatics and Common Ground in Interaction

aA
ccademia
university
press



editors in chief

Roberto Basili

Università degli Studi di Roma Tor Vergata

Simonetta Montemagni

Istituto di Linguistica Computazionale “Antonio Zampolli” - CNR

advisory board

Giuseppe Attardi

Università degli Studi di Pisa (Italy)

Nicoletta Calzolari

Istituto di Linguistica Computazionale “Antonio Zampolli” - CNR (Italy)

Nick Campbell

Trinity College Dublin (Ireland)

Piero Cosi

Istituto di Scienze e Tecnologie della Cognizione - CNR (Italy)

Giacomo Ferrari

Università degli Studi del Piemonte Orientale (Italy)

Eduard Hovy

Carnegie Mellon University (USA)

Paola Merlo

Université de Genève (Switzerland)

John Nerbonne

University of Groningen (The Netherlands)

Joakim Nivre

Uppsala University (Sweden)

Maria Teresa Paziienza

Università degli Studi di Roma Tor Vergata (Italy)

Hinrich Schütze

University of Munich (Germany)

Marc Steedman

University of Edinburgh (United Kingdom)

Oliviero Stock

Fondazione Bruno Kessler, Trento (Italy)

Jun-ichi Tsujii

Artificial Intelligence Research Center, Tokyo (Japan)

editorial board

Cristina Bosco

Università degli Studi di Torino (Italy)

Franco Cutugno

Università degli Studi di Napoli (Italy)

Felice Dell'Orletta

Istituto di Linguistica Computazionale "Antonio Zampolli" - CNR (Italy)

Rodolfo Delmonte

Università degli Studi di Venezia (Italy)

Marcello Federico

Fondazione Bruno Kessler, Trento (Italy)

Alessandro Lenci

Università degli Studi di Pisa (Italy)

Bernardo Magnini

Fondazione Bruno Kessler, Trento (Italy)

Johanna Monti

Università degli Studi di Sassari (Italy)

Alessandro Moschitti

Università degli Studi di Trento (Italy)

Roberto Navigli

Università degli Studi di Roma "La Sapienza" (Italy)

Malvina Nissim

University of Groningen (The Netherlands)

Roberto Pieraccini

Jibo, Inc., Redwood City, CA, and Boston, MA (USA)

Vito Pirrelli

Istituto di Linguistica Computazionale "Antonio Zampolli" - CNR (Italy)

Giorgio Satta

Università degli Studi di Padova (Italy)

Gianni Semeraro

Università degli Studi di Bari (Italy)

Carlo Strapparava

Fondazione Bruno Kessler, Trento (Italy)

Fabio Tamburini

Università degli Studi di Bologna (Italy)

Paola Velardi

Università degli Studi di Roma "La Sapienza" (Italy)

Guido Vetere

Centro Studi Avanzati IBM Italia (Italy)

Fabio Massimo Zanzotto

Università degli Studi di Roma Tor Vergata (Italy)

editorial office

Danilo Croce

Università degli Studi di Roma Tor Vergata

Sara Goggi

Istituto di Linguistica Computazionale "Antonio Zampolli" - CNR

Manuela Speranza

Fondazione Bruno Kessler, Trento

Registrazione presso il Tribunale di Trento n. 14/16 del 6 luglio 2016

Rivista Semestrale dell'Associazione Italiana di Linguistica Computazionale (AILC)
© 2021 Associazione Italiana di Linguistica Computazionale (AILC)



Associazione Italiana di
Linguistica Computazionale



direttore responsabile
Michele Arnese

isbn 9791280136770

Accademia University Press
via Carlo Alberto 55
I-10123 Torino
info@aAccademia.it
www.aAccademia.it/IJCoL_7_1-2



Accademia University Press è un marchio registrato di proprietà
di LEXIS Compagnia Editoriale in Torino srl

Computational Dialogue Modelling: The Role of Pragmatics and Common Ground in Interaction

Invited editors: *Hendrik Buschmeier and Francesco Cutugno*
co-editors: *Maria Di Maro and Antonio Origlia*

CONTENTS

Editorial Note <i>Francesco Cutugno, Hendrik Buschmeier</i>	7
Knowledge Modelling for Establishment of Common Ground in Dialogue Systems <i>Lina Varonina, Stefan Kopp</i>	9
Pragmatic approach to construct a multimodal corpus: an Italian pilot corpus <i>Luca Lo Re</i>	33
How are gestures used by politicians? A multimodal co-gesture analysis <i>Daniela Trotta, Raffaele Guarasci</i>	45
Toward Data-Driven Collaborative Dialogue Systems: The JILDA Dataset <i>Irene Sucameli, Alessandro Lenci, Bernardo Magnini, Manuela Speranza e Maria Simi</i>	67
Analysis of Empathic Dialogue in Actual Doctor-Patient Calls and Implications for Design of Embodied Conversational Agents <i>Sana Salman, Deborah Richards</i>	91
The Role of Moral Values in the Twitter Debate: a Corpus of Conversations <i>Marco Stranisci, Michele De Leonardis, Cristina Bosco, Viviana Patti</i>	113
Computational Grounding: An Overview of Common Ground Applications in Conversational Agents <i>Maria Di Maro</i>	133
Cutting melted butter? Common Ground inconsistencies management in dialogue systems using graph databases <i>Maria Di Maro, Antonio Origlia, Francesco Cutugno</i>	157
Towards a linguistically grounded dialog model for chatbot design <i>Anna Dell'Acqua, Fabio Tamburini</i>	191
Improving transfer-learning for Data-to-Text Generation via Preserving High-Frequency Phrases and Fact-Checking <i>Ethan Joseph, Mei Si, Julian Liaonag</i>	223

Analysis of Empathic Dialogue in Actual Doctor-Patient Calls and Implications for Design of Embodied Conversational Agents

Sana Salman*

Faculty of Science and Engineering –
Macquarie University, Sydney

Deborah Richards**

Faculty of Science and Engineering –
Macquarie University, Sydney

Patrina Caldwell†

University of Sydney, The Children’s
Hospital at Westmead

Embodied Conversational Agents (ECAs) are virtual agents that exhibit humanlike verbal and non-verbal behaviours. When it comes to eHealth, ECAs can provide vital support to patients by being more reachable. In order to make ECAs more effective, humanlike empathy expressed during conversation through relational cues is essential. Empathy revolves around a wide range of verbal and non-verbal behaviours that include, for example, the choice of words in social dialogues. Owing to the COVID-19 situation, there was an opportunity to record online consultations in the Incontinence Clinic and Sleep Clinic at the Children’s Hospital at Westmead in Sydney, Australia. The present study analysed these human dialogues using qualitative research methods to understand the role of empathic dialogue used by the medical team. The qualitative analysis of the live calls used psychology-based relational cues derived from conversational characteristics of humans to build a coding framework around the most relevant themes. Statistical analysis was used to compare relational cue usage between healthcare roles. Finally, using the framework dialogues of the medical team and two ECAs known as Dr Evie (eVirtual Agent for Incontinence and Enuresis) and SAM (Sleep Adherence Mentor) were compared to provide recommendations for health practitioners and future ECA dialogue development.

1. Introduction

Embodied Conversational Agents (ECAs) are virtual agents that exhibit humanlike verbal and non-verbal behaviours. They are increasingly being applied in contexts where the main mode of interaction is a dialogue between two or more humans (Bickmore, Gruber, and Picard 2005). In eHealth, ECAs can provide vital support to patients by being more reachable and available in their time of need (Richards and Caldwell 2017). The agents can not only act as a source of information on health issues (Lisetti et al.

* Faculty of Science and Engineering – Macquarie University, Sydney
E-mail: sana.salman@students.mq.edu.au

** Faculty of Science and Engineering – Macquarie University, Sydney.
E-mail: deborah.richards@mq.edu.au

† University of Sydney, The Children’s Hospital at Westmead
E-mail: patrina.caldwell@health.nsw.gov.au

2012) and their prevention or cure (Yin, Ring, and Bickmore 2012), but also motivate the patients to adhere to treatments (Bickmore et al. 2010); (Richards and Caldwell 2017). For teenagers and children, a virtual agent can act as an educator, buddy and motivator (Looije, Neerinx, and Lange 2008), such as those designed for childhood obesity intervention (Kowatsch et al. 2017) or for promoting well-being and positive thoughts in young people undergoing cancer treatment (Greer et al. 2019).

In order to make ECAs more effective, humanlike empathy expressed during conversation is a vital component (Bickmore, Gruber, and Picard 2005). Empathy is an essential part of building rapport and creating a bond, or a working alliance, between the patient and therapist to improve patient adherence and satisfaction (Bennett et al. 2011). Empathy has been defined as ‘an observer reacting emotionally because he perceives that another is experiencing or about to experience an emotion’ (Paiva et al. 2005), (p.4), or ‘the process whereby one person feels her/himself into the consciousness of another person’ (Wispé 1987). Empathy is expressed through a wide range of non-verbal behaviours, such as mirroring head nods, and verbal behaviours, such as the choice of words in social dialogues, the use of greetings and farewell rituals and the art of bringing continuity in the conversation (Laver 2011). In this paper, we focus on empathy expressed through verbal dialogue.

Empathy has been researched as a vital component of emphasis devices in human computer interaction to emphasize particular qualities and context (Wright and McCarthy 2008) which leads to building dialogue sets that contain the empathic cues, such as confirmation and adherence encouragement, that act as emphasis monitors. Rather than being specialized in a single quality or domain, a good open-domain conversational agent should be able to seamlessly deliver the necessary features into one cohesive conversational flow. Facebook’s blender bot is an example of such a conversational agent that has been specially trained on empathic dialogue sets (Roller et al. 2020). As in the Blenderbot approach, there is a tendency in artificial intelligence approaches to use mimicry and to replicate the behaviour of humans based on corpora without analysis of whether that behaviour is appropriate. However, mimicking/replicating human responses to a tragic event, for example, will not always be the best response (Lundqvist 1995) and will depend on the context such as the relationship between the parties. The importance of empathy in human dialogue has led to interest in ECAs expressing empathy particularly to bring about behaviour change (McRorie et al. 2009), (Ochs, Pelachaud, and Mckeown 2017), (Ravichander and Black 2018).

In this paper we evaluate the use of relational cues in recorded patient-doctor dialogues and the ECA’s known as Dr Evie (eVirtual agent for Incontinence and Enuresis) and as SAM (Sleep Adherence Mentor) designed to increase adherent behaviour in the domains of paediatric incontinence and sleep disorders, respectively. Both conditions have long specialist waitlists. Paediatric Incontinence affects up to 20% of school aged children (Malhotra et al. 2020). It often leads to avoidance of social interactions, low self-esteem and poor quality of life. Sleep disorders also impacts quality of life for many children (Roth 2007).

Due to COVID-19, there was an opportunity to record 30 online consultations (15 minutes to 2 hours) in the Incontinence Clinic and Sleep Clinic. We used qualitative research methods that help in building structured linguistic frameworks by analysing conversations (Alam, Danieli, and Riccardi 2018) and further utilizing them in building dialogue sets for ECAs. The objective is to semantically and pragmatically analyse the actual human dialogues involving a coding framework comprised of 16 relational cues identified in the literature using NVIVO and Discursis. Relevant to the context, our analysis identifies differences between new and follow-up patients and different

clinician roles (paediatrician, psychologist, physiotherapist, nurse). The results from the coding framework are further analysed statistically to more deeply understand the relationships, context and use of relational cues between different health specialists and patients towards specification of an empathic dialogue framework that supports complex interactions. We also identify design features for future ECA dialogue development and improvement in their emphatic structures.

In the next section we describe the background domain relevant to the online consultations and evaluated ECAs (Section 2), followed by review of the literature in empathy and relational cues in dialogue (Section 3). The methodology is presented in Section 4 followed by results in Section 5. Finally, discussion, conclusions and future work are presented.

2. Background

According to the International Children Continence Society (ICCS), the medical condition of incontinence refers to intermittent or continuous bed wetting during the day or night or both (Maternik, Krzeminska, and Zurowska 2015). Paediatric incontinence is a common condition affecting up to 20 percent of school-aged children (Malhotra et al. 2020) in many of their social activities like sports and sleepovers, which often leads to avoidance of social interactions. The children feel embarrassed and anxious, which leads to frustration and low self-esteem (Theunis et al. 2002). Children report a negative self-image owing to the physical and psychological impact of having incontinence, which is often unrecognised and seen as just another milestone in their growth (Butler 1998);(Harter 1982).

Despite the fact that incontinence impacts the patient's quality of life and is a cause of stress for them and their families (Malhotra et al. 2020);(Thibodeau et al. 2013) long waiting times to receive treatment are common, up to two years. This is because of a shortage of specialists, as incontinence is categorised as non-life threatening. Hence, ECAs could provide more timely support (Richards and Caldwell 2017); (Laranjo et al. 2018).

To address the problem of long public hospital waitlists, an incontinence specialist for children aged 3-18 at the Children's Hospital at Westmead (CHW) in Sydney, Australia, created an interactive eHealth program known as eADVICE (electronic Advice and Diagnosis Via the Internet following Computerised Evaluation). eADVICE enables young patients accompanied by their parents to get an online consultation regarding incontinence treatment factoring in the patient's medical history and encoded algorithmic response scenarios that capture the domain knowledge of the health experts. Developed in 2016, the website was evaluated in several pilots that found adherence to the six possible recommended treatments was around 50 percent. To allow patients and families to 'discuss' their treatments, eADVICE was enhanced through the addition of an ECA known as Dr Evie (eVirtual agent for Incontinence and Enuresis), which gave a human embodiment to the online consultation experience. Possessing the actual voice of the incontinence specialist, this ECA significantly improved the adherence and health outcomes of patients on the hospital waiting list (Richards and Caldwell 2017). The success of Dr Evie can be attributed to its availability and its empathic and empowering dialogue (Bickmore 2004).

Owing to its success for incontinence patients, the eADVICE approach – involving a website to provide tailored recommended treatments and an ECA to discuss the treatments – has been deployed for sleep disorders (eADVICE-sleep), another condition that is not life-threatening but significantly reduces quality of life (Roth 2007). Roth

(2007) associated sleep disorders with “the presence of long sleep latency, frequent nocturnal awakenings or prolonged periods of wakefulness during sleep periods”. This condition is considered chronic if the sleep environment is comfortable but the daytime routine is full of distress, light headedness and anxiety due to lack of sleep (Kredlow et al. 2015).

As sleep disorder patients also suffer from long waiting periods to access specialists and poor treatment adherence, they can potentially benefit from an ECA (Horsch et al. 2012); (Yin, Ring, and Bickmore 2012). Known as SAM (Sleep Adherence Mentor), the ECA in eADVICE-sleep acts as a virtual sleep coach. SAM has eight dialogue sets to cover the range of treatments and to ensure the dialogues are appropriate for the child’s age.

3. Empathy, Relationship Building and Relational Cues

This review first defines empathy and then briefly reviews its role in human relationships and past ECA work involving empathy and human-ECA relationship building. To provide the basis for the coding themes in the methodology, we provide a brief review from linguistics to identify and define verbal relational cues and types of dialogue expressions that have been found to assist relationship building.

Empathy is a complex human behavioural phenomenon defined by Hoffman (2001) as ‘the cognitive awareness of another person’s internal states that is, his thoughts, feelings, perceptions and intentions’ (p.29). Hoffman refers to empathy as any emotional reaction compatible with (but not necessarily similar to) the other’s situation. (Rogers and others 1959) theory of positive psychology and his client-centred framework emphasise that ‘for a person to ‘grow’, they need an environment that provides them with genuineness (openness and self-disclosure), acceptance (being seen with unconditional positive regard) and empathy (being listened to and understood)’(Mamarimbing 2021), (p.8).

3.1 Empathy and Relationship Building with ECAs

There has been more than a decade of research on the importance of empathy in human dialogues, which has led to interest in how ECAs can express empathy to bring about behaviour change (McRorie et al. 2009);(Ochs, Pelachaud, and Mckeown 2017);(Ravichander and Black 2018). ECAs typically have a particular purpose referred to as task-oriented empathy (Bickmore, Caruso, and Clough-Gorr 2005); (Bickmore et al. 2010), which is more easily detectable in focused dialogues. However, social empathy, which is not task based, is also important, as it offers comfort and encourages long-term relationships (Bickmore 2004). According to Halpern (2007), task-oriented empathy comes more naturally in doctor-patient dialogues while a doctor gathers the patient’s background information or recommends a certain treatment. Non-task-based, or social, empathy is more generic to the conversational themes in daily life. In designing ECAs, the component of social empathy is more complex and has been less commonly analysed (Halpern 2007).

Owing to the importance of empathy in human relationships, many researchers have created and evaluated empathic ECAs such as GRETA (Hartmann, Mancini, and Pelachaud 2005) and REA, the Real Estate Agent (Cassell et al. 1999). Building and maintaining human-ECA relationships, however, is broader than congruent expression of empathy and includes other behaviours. Long-term relationships are highly influenced by the use of the right relationship-building dialogues (Stafford and Canary 1991).

According to the psychology or medical literature, a working alliance is important for successful therapy (Halpern 2007). Many scales have been developed that emphasise the use of empathic and social dialogues during health-related consultations (Looije, Neerinx, and Lange 2008); (Yin, Ring, and Bickmore 2012).

ECAs with empathic dialogues have been studied across a diverse range of health programs such as relational agents for anti-psychotic medication adherence (Bickmore et al. 2010), avatar-based health intervention to modify unhealthy lifestyles (Lisetti et al. 2012), exercise advisors that interact with older adults (Bickmore, Caruso, and Clough-Gorr 2005) and ECAs that can help cancer patients to adopt a positive lifestyle after chemotherapy (Greer et al. 2019). Research has suggested frameworks that determine the useful verbal and non-verbal behaviours for virtual agents, such as 10 cues including empathy, social dialogues and continuity (Bickmore, Gruber, and Picard 2005), the Big Five model of personality traits (Neff et al. 2010), five dialogue characteristics that exhibit relationship building (Richards and Caldwell 2017) and annotation schemes for negative emotions' handling in customer care bots (van Velsen et al. 2019).

3.2 Relational cues in Linguistics and Psychology

Bardovi-Harlig (2010) defines pragmatics as 'the study of language from the point of view of users, especially of the choices they make, the constraints they encounter in using language in social interaction and the effects their use of language has on other participants in the act of communication' (p. 221). Research on social cognition through pragmatics (Bosco et al. 2015) has resulted in empathic pragmatic models (Zhanghong and Qian 2018) that refer to empathy in the context of verbal utterances that help build strong relationships.

This paper focuses on identifying a set of verbal relational cues that have been reported in the literature, particularly in psychology, to build a working alliance or strong rapport between the patient and the health specialist. The literature has identified a number of relational cues as in the following examples.

Empathy as a cue: Empathic phrases can be divided into three categories – Queries, Clarifications and Responses – which will be the foundation of empathic cue detection in the dialogue set (Coulehan et al. 2001).

Social dialogue: Conversation can be broken into three phases: the opening, middle and closing phases with social dialogues playing their role in opening and closing phase (Laver 2011).

Reciprocal self-disclosure: Conversations use a key social strategy defined as self-disclosure to build relations and increase conversational depth and as a process of disclosing details about yourself to the listener (Ravichander and Black 2018).

Meta-relational Communication: The specific talk that results in enjoyable relationships, cooperation, building up self-esteem, giving compliments, being courteous and polite, mitigating criticism, fostering patience and forgiveness in the participants, encouraging openness and talking about the relationship's quality and needs, and helping to acknowledge the relationship is categorized as the meta-relational communication (Stafford and Canary 1991).

Continuity: During a conversation, the behavioral units that exhibit continuity are of three types: prospective, introspective and retrospective. For a doctor-patient interaction, prospective and retrospective are covered under verbal interaction (Gilbertson, Dindia, and Allen 1998).

Reference to mutual knowledge: Some examples of mutual knowledge include knowledge of participant's/patient's biography, present life or habits, recent or future events (Planalp and Benson 1992).

Affirmation: Something emotionally challenging or complex to comprehend is related to an affirmative response from the listener. The complexity can be due to multiple reasons ranging from whether or not take up a medicine or follow a recommended treatment that is not fully understood or experienced. The main purpose of affirmation is to enable to patient to express their frustration openly. Letting the patient voice out their concerns and responding in understandable utterances is the essence of affirmation. (Cameron et al. 2015).

Confirmation: Means to reiterate the facts or validating the correctness of something previously believed or suspected to be the case is defined as confirmation. In eHealth, through confirmation, the physician recognises the situation conveyed by the patient and in return receives confirmation from the patient who sees the specialist as someone to confide in and understand their issues (Abramovitch and Schwartz 1996).

4. Methodology

4.1 Data Collection

CHW recorded its live clinical calls for research purposes during the COVID-19 period from June to November 2020. In these calls, the patients are children aged 3–18 and they have specific issues such as incontinence of urine. The objective of this study is to analyse these actual human dialogues using qualitative research methods and draw findings focused on the element of empathy in doctor-patient conversations. This live-call dataset provides a unique view of how actual dialogues, recorded in a real-life environment, can suggest changes to the existing ECA dialogues and help to validate the existing component of empathy in these ECAs. The health specialists consists of two paediatricians (senior paediatrician is later referred to as senior doctor), one nurse, two physiotherapists and one registrar. All of them are experienced health specialists with experience ranging from 10-40 years. The data is composed of 23 unique patients and consultation sessions, with a total of 50,000 utterances, collected over six months from the incontinence and sleep clinics.

4.2 Data Pre-processing

The process of analysing live calls begins with transcribing the recorded calls. Transcription involves generating text files from the audio recordings followed by character identification (e.g. physio, nurse, doctor, patient or relative of patient) and validating the dialogue assignment to the respective character. To ensure privacy, elimination of the patient's personal information (e.g. name, contact number and email) is the next step. The process involved listening to the recording and removing any personal detail manually. We replaced it with random names and emails. The tools used for transcribing the data include Temi¹ and Transcribe Wreally².

1 www.temi.com

2 <https://transcribe.wreally.com/>

4.3 Qualitative Analysis Overview

Qualitative data analysis is an approach to finding patterns in conversations and interviews. The data itself can be unstructured, which means there are no predefined questions or predefined answers to choose from. The qualitative analysis method chosen depends upon not only the level of structure in the data but also how acquainted a researcher is with the subject being analysed. Figuring out the context of the data in the form of variables is known as coding. The description follows Bengtsson (2016), who used content analysis in the domain of nursing, which is close to the domain of interest. Coding has two main approaches, deductive and inductive. In the deductive approach, the researcher is familiar with the content of the discussions and has developed an understanding of the context. Hence, the codes or themes are known beforehand. In the inductive approach, the researcher reads through the conversation and figures out the common words, semantics and context before grouping them into themes to define the coding framework. The next step is to decide whether to code the exact words or phrases from the conversation as codes or themes or to go deeper and understand the underlying meaning of the dialogue content and define that as a code. The former is known as manifest analysis, where the codes are the exact content; the latter is latent analysis, where the researcher goes under the surface and defines the codes according to the research aims (Bengtsson 2016).

4.4 Coding and Annotation Approach

The present research approach is deductive latent analysis, which means that codes relevant to verbal behaviours leading to empathy will be extracted from a literature review and will have associated dialogues for further analysis. Each dialogue can be placed in multiple codes as well, depending upon the hidden context of utterances. Ensuring coding's credibility increases the measure of trustworthiness of the coding (Graneheim and Lundman 2004). Credibility mechanism is how similar or dissimilar the coding results are when another person tries to reproduce the results. While qualitative methods do not claim to be reproducible, one way to improve credibility is to seek agreement between different researchers who do the coding independently and then establish a consensus. The approach is not to validate the coding quantitatively but to open the forum for discussion in which all experts come to consensus about each other's way of coding (Woods and Catanzaro 1988).

4.4.1 Coding Process

The coding phase in qualitative analysis begins by analysing the dialogues one by one to find themes according to the literature review on verbal and non-verbal behaviours found during conversations. In this study, the focus is on verbal behaviours because audio recordings cover only the verbal aspects of conversation. While finding themes, the focus is on the element of empathy; hence, all behaviours that exhibit empathy will be considered (e.g. politeness, inclusive pronouns). It is important to consider both task-oriented and social empathy. After defining and distinguishing themes from the literature to avoid overlaps, 16 themes were identified in relation to expressing empathy through dialogue: ten themes from Bickmore et al. (2005) and six themes from Richards and Caldwell (2017) as shown in Figure 1.

For each relational cue, we have definitions, common key words and examples of sentences from the literature that can act as a set of guidelines for the application of the

annotation scheme to other data sets. These details can be made available by contacting the first author.

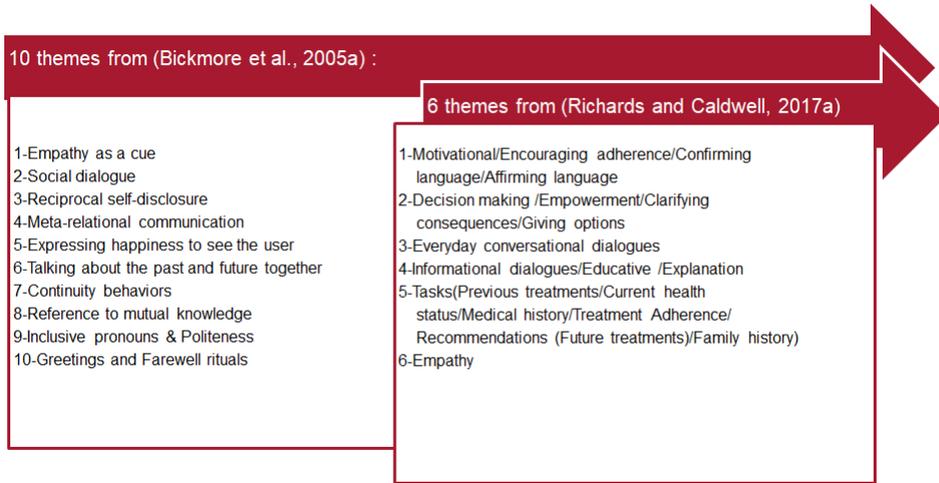


Figure 1
A breakdown of parent and sub-themes derived from the literature

4.4.2 Coding in NVIVO

Coding was done in NVIVO which began with building a framework of nodes, cases and roles based on the research questions. The detailed review of sixteen themes from literature resulted in 13 parent codes, out of which 4 parent codes were further categorized into 14 child codes as shown in Figure 2. For simplicity and for further relative comparison of codes within roles each dialogue was coded in one category only. Also the child codes were given preference over the parent codes. For example, in case of motivational dialogue, the child codes were encouraging adherence, giving option and clarifying consequences. Only in rare cases where the dialogue was motivational but did not fit any child code was it assigned to the parent code. In the end, child code dialogues were aggregated to create the total count of the parent code.

4.4.3 Independent Coders' Agreement through Cohen's Kappa

In order to seek agreement between how similar or dissimilar coding is, two annotators (SS and DR) took part in the validation process. SS was the main coder who had the context of approximately 50,000 dialogues. DR is an expert in qualitative analysis approaches but had not coded the whole dataset. After 200 dialogues were randomly selected from the first coder's assessment, the second coder was briefed on the codes' description, but the context remained missing until the first kappa was calculated so that consensus could be established in later discussions. The idea is to see how much the dialogues resonate with the theme even if the context is not given so that a more generalised coding can be obtained in the first iteration. The aim was not to validate the coding quantitatively but to open the forum for discussion in which all experts agree to each other's way of coding (Woods and Catanzaro 1988). Various coder's agreement techniques have been used in research including Krippendorff's alpha as well as Scott's pi and Cohen's kappa (Artstein and Poesio 2008). Cohen's kappa was selected as a

Parent Codes	Child Codes
1 : Continuity	
2 : Decision making	2.1 : Clarifying consequences
	2.2 : Empowerment
	2.3 : Giving options
3 : Empathy	
4 : Expressing happiness to see the Patient	
5 : Greetings and Farewells	
6 : Informational dialogues	6.1 : Educative
	6.2 : Explanation
7 : Meta-relational Communication	
8 : Motivation	8.1 : Affirmation
	8.2 : Confirmation
	8.3 : Encourage adherence
9 : Reciprocal self-disclosure	
10 : Reference to mutual knowledge	
11 : Social dialogues	
12 : Talking about the past and future together	
13 : Tasks	13.1 : Current health status
	13.2 : Family history
	13.3 : Medical history
	13.4 : Previous treatments
	13.5 : Recommendations (Future Treatments)
	13.6 : Treatment adherence

Figure 2
Relational Cues used in the Coding Process

measure of the agreement between the coder’s independent coding. An unweighted kappa score is calculated, which calculates the percentage agreement and the measure of coding by chance in each theme (Warrens 2015).

Initially, a kappa score of 0.61 was calculated on 200 dialogues with 24 themes. The coders examined the percentage agreement within each theme and agreed that three themes – everyday conversational dialogues, inclusive pronouns and politeness strategies – should be re-coded since the dialogues that were coded in them were multi-thematic and were more appropriately placed in other themes like empowerment and motivation. Both coders re-coded the original 200 dialogues in these themes. Some confusion remained concerning dialogues that were specifically questions like ‘Do you have any more questions?’ or ‘Are you with me so far?’, as they were part of multiple themes. Based on a literature review, it was agreed that the appropriate classification was confirmation. These changes and clarifications resulted in a revised kappa score of 0.75. The remaining dialogues were reviewed one by one. After discussion about the categories, the second coder agreed to change allocations from affirmation to empathy, encouraging adherence to recommendation and from child theme to parent theme, if a dialogue had more than one child theme representation. Final coding resulted in agreement on 171 dialogues and a kappa value of 0.84.

The remaining cases were resolved through discussion. In most cases, the context was mandatory for the assignment because the dialogues had more than one coding

Relational Cue	Dialogue Snippet
Continuity	Oh, good. Okay. Well, um, any questions from Sam about it. He's doing very well. So let's talk again in about three months and hopefully will discharge you from our care.
Decision making	You could be just having the wet night occasionally. I don't know. So I'm going to send you your own chart.
Clarifying consequences	I want you to check what sorts of foods you're doing.
Empowerment	So, um, how what your side effects of the medicine is it can make you thirsty. And so I think that's a Good, because you need to learn to be the boss of your own body. So questions around toilet training.
Giving options	What age was Sam toilet trained as a toddler?
Empathy	And we'll just leave it at that. You don't have to keep going, unless you absolutely want to and it's easy. need to work there.
Everyday conversational dialogues	Okay, look, it's its 5050. I find everyone neither loves it nor hates it.
Expressing happiness to see the Patient	Okay, so kind of spot on. Glad you're there with mom!
Greetings and Farewells	That's a pleasure. Talk to you soon. Take care. Have a good day. Bye bye.
Inclusive pronouns	Then we might be making some progress by doing bedwetting alarm training if bladder is not normal. So there's a whole lot of stuff we can do together.
Informational dialogues	It's harder to do it when it's wet and cold and dark. It's true, you just want to roll out and go back to sleep. But I think what's happening is your bladder is telling your brain that it's cool.
Educative	And teenagers behaviors and habits and bodies are different. So what I know about teenagers that they all need a good nine hours sleep in order to function.
Explanation	You've got the gene. And if you haven't got the pieces in place, that's when things are easier to disrupt if you think, and it's more likely to have an accident.
Meta-relational Communication	Okay, and that's what we want. We want a bladder that is well and is stable and empties well when you go to the toilet and okay we're getting close to the end of it.
Motivation	Well, so we want her to be prompted every two hours to go to the toilet and she clearly needs a little bit more. We need to try and work out, but that's what we're aiming for a bladder big enough to hold well
Affirmation	You are doing absolutely wonderfully because wearing the nappies helps you get better sleep at night.
Confirmation	Okay. So, normally you sleep at nine o'clock. Is that right, okay.
Encourage adherence	It gets better, you find that initially you would have a large wet and then the wet gets smaller and smaller, which means you're responding better
Politeness strategies	Sorry, what was that!
Reciprocal self-disclosure	But three to four hours. Probably he would have to pee 20 times a day. I would, which is what he's doing.
Reference to mutual knowledge	But as mum rightly noticed you are not quiet but you are very close to turning into a teenager.
Social dialogues	I'm just going to briefly tell her about you and you correct me if you think I've got it wrong. Okay, I just missed the fun side that she can hear you.
Talking about the past and future together	You visited back in October last year and it may be worthwhile to come back in. So we can physically examine you and scan you to make sure your bladder is normal.
Recommendations	You can just put money in your savings, slow down, just to make sure you can do it, you can practice on food first. And then if you're good, then you can take the tablet.

Figure 3
Examples of dialogues' annotation after coders' agreement

category in them such as affirmation or explanation and the context placed it correctly in empathy.

Following consensus on the 200 dialogues, the main coder reclassified any dialogues in the 50,000 dialogues that had been reassigned to three themes – everyday conversational dialogues, inclusive pronouns and politeness strategies. Finally, the main coder confirmed that all of the dialogues used everyday conversational language and that use of personal pronouns would be automatically calculated by searching for the terms 'us' and 'we' as a rough but quick method of assessment that would enable comparison between roles and individuals. The coding process resulted in parent and child codes listed in Figure 2. A few annotated dialogues after consensus are shown in Figure 3 and Figure 4 to give a glimpse of dialogues and their chosen relational cues.

4.5 Empathic Cues: Analysing the Dr Evie and SAM Dialogue Sets

Dr Evie and SAM both have structured dialogues with empathic cues already being a part of their semantics. The hypothesis of this study is that live unstructured sessions can bring out different sentence structures that are more beneficial in creating a level ground of adherence for the patients. This includes a validation process in which Dr

Relational Cue	Main Annotator	Second Annotator	Final Annotation
that's okay. That's okay. So currently, I'm thinking you don't wear any pull ups or protection to bed. Correct.	Empathy	confirmation	Empathy
So if we can break her toileting into daytime going to wee, nighttime wetting and Poo's, the plan will be more managable.	Inclusive pronouns	Educative	Educative
Make sure you're getting enough sleep, and I know that's a big ask for teenagers. But honestly, if you want to get Dry. That's really important.	Reference to mutual knowledge	Every day conversational dialogues	Empathy

Figure 4
Inter coder disagreement example dialogues

Evie and SAM would be thoroughly screened for empathic cues finalised in the prior coding process. A similarity scale would determine the threshold, and based on the similarity scores further empathic cues could be recommended. The same coding process used in Sections 4.4 was followed to evaluate these. After normalising the results, a comparison has been done to show how empathic both dialogue sets are.

5. Results

This section covers the analysis of 23 patient’s consultations. The age of the patients ranges from 9-12 years with 10 patients being females while the rest are male. The consultations consist of 12 follow-ups and 11 new patient sessions. In 22 session, patients are accompanied by one or both parents and the average session length is 40 minutes. Out of 23, 16 sessions have one health specialist while 7 sessions have 2 health specialists with one always being the senior doctor. The role of the medical specialist and whether the consultation was a new or follow-up meeting are used for the comparison of results. The findings from the recorded consultations are discussed in Sections 5.1 to 5.3. Section 5.4 analyses results from the recorded consultations with the Dr Evie and SAM dialogues. In Section 5.5 the dialogue structure and topic sequencing is formulated into the coding framework.

5.1 Percentage Distribution of Relational Dialogues

The percentage of dialogue usage was compared for the following six categories:

1. Senior doctor’s consultations with first-time patients
2. Senior doctor’s consultations with follow-up patients
3. Average usage of relational cues by the senior doctor with follow-up versus new patients (Figure 5)
4. Physiotherapist’s consultations with patients
5. Nurse’s consultations with patients
6. Usage of relational cues in sessions where both the senior doctor and physiotherapist are present (Figure 6).

It is evident from the coding percentages shown in Figure 5 that patients who visit the clinic for the first time need more information about their health issue and its remedies. The senior doctor uses more informational and motivational dialogues and decision-making is also encouraged. The health specialist also needs to ask about a new patient’s medical history and current health status more than for the follow-up patients.

For follow-up patients, the percentage usage of cues is more variable as it depends on how many sessions the patient has had before and their progress as to date. For some patients, it is more about encouraging adherence to a treatment discussed in previous sessions. For others, the effects of new recommendations need clarification and dialogues confirm their understanding. Figure 5 also shows that the senior doctor’s empathic cues’ usage is the same for both new and follow-up patients but the difference lies in encouraging adherence and informational and motivational dialogues. Informational dialogues are more for new patients, whereas motivational and decision-making cues are more for follow-up patients. The reciprocal self-disclosure and social dialogues feature less in percentage usage but they are relatively more used for first-time patients.

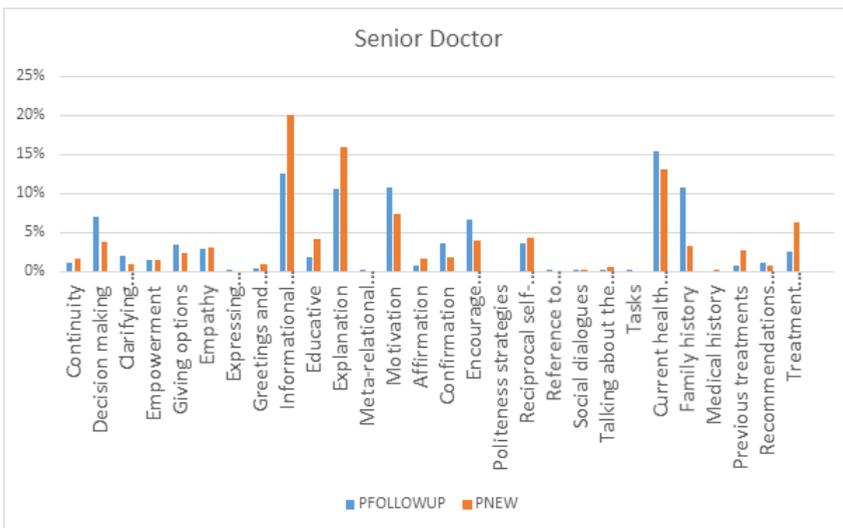


Figure 5 Senior doctor’s average usage of relational cues with first-time (PNEW) and follow-up (PFOLLOWUP) patients

The physiotherapist and nurse are the other most frequent roles found in the consultations. Most of the consultations contain one therapist role but in consultations where multiple specialists are present, as shown in Figure 6, where both senior doctor and physiotherapist appear, the physiotherapist uses more task-based dialogues and the senior doctor uses more adherence and decision-making dialogues. Social dialogues are uttered more by the physiotherapist, but empathic cues are uttered more by the senior doctor.

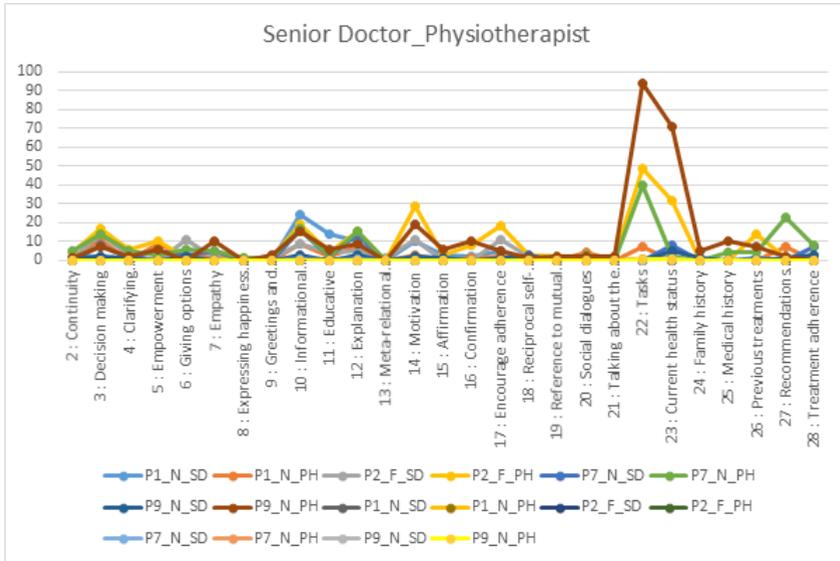


Figure 6 Senior doctor’s and Physiotherapist’s average usage of relational cues with first-time (PNEW) and follow-up (PFOLLOWUP) patients

5.2 Role Differences and Similarities in the Use of Relational Dialogues

The data helps in analysing the use of relational cues in multiple roles. In total, there were six roles to analyse, with more consultation sessions for the senior doctor, physicians and nurse. The remainder of the medical team – a paediatrician, registrar and male physiotherapist – had only one consultation each. More data is needed to analyse their use of relational cues, but for the purpose of this research they are grouped into one category under ‘others’. Owing to few data samples and lack of confirmation of normal distribution in the usage of a particular theme within a category under observation, non-parametric (i.e. Mann Whitney U) tests were chosen to understand the differences and similarities in the use of relational dialogues. The Mann Whitney U test compares outcomes between two independent groups based on the median of two distributions. The test was performed on all relational cues. The results that are significant at 90 percent confidence level or have high U-values include empowerment, explanation, social dialogues and reciprocal self-disclosure when comparing senior doctor’s usage of relational cues with junior doctors. These categories were further analysed to determine whether the senior or junior doctors used more of these cues.

Dominance is defined as the higher usage percentage of the relational cue in most of the sample points for that role. It is evident that Senior doctor is dominant in using relational cues like explanation, reciprocal self-disclosure and clarifying consequences while junior doctors are dominant in using relational cues like empowerment, social dialogues and task-based dialogues. Examples of visualisations that support the statistics are presented in Figure 7.

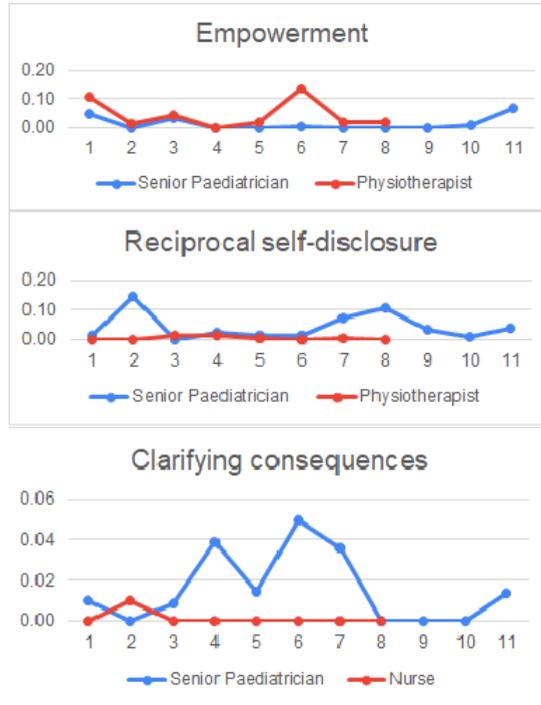


Figure 7
Comparison of Senior Paediatrician and Physiotherapist. X-axis represents specific patients, Y-axis is percentage of relational cue usage

5.3 Demographic Influence on Use of Relational Dialogues

The influence of patient demographics including gender, age and first time visitation on the use of relational dialogues has been analysed using Mann-Whitney U scores in Figure 8. The higher the U-value, more significant is the difference of relational cues usage in the respective groups. The p-value and z-score further determine the significance of the differences in the given empathic cue’s group. Figure 8 also shows U-values of new and follow-up consultations. Although first time visitation is not the demographics of the patient but it has a major contextual influence on the usage of relational cues. The new patients needed to be walked through the whole treatment details and were encouraged to speak up about their current health issues. The follow-up patients were either asked for time and volume charts or about their adherence to a certain treatment that had been discussed in previous consultations.

5.4 Comparison of Dr Evie’s and SAM’s Dialogues with Live-call Dialogues

Dr Evie’s dialogue set consists of multiple treatment-based dialogue streams, including alarm training, bowel program, caffeine intake, fluid increase, medication and time voiding. The same coding framework used for the recorded consultations was applied to the dialogues used in Dr Evie to facilitate a comparison between the usage of rela-

Group	U-value	p-value	z-score
Males versus Females			
Motivational dialogues	49	0.044	-2.014
Explanation	54	0.077	1.771
Age <10 versus Age >=10			
Previous treatments	39	0.014	-2.464
Clarifying consequences	56	0.095	-1.674
New versus follow-up			
Motivational dialogues	35	0.007	-2.693
Encourage adherence	50	0.049	-1.965
Medical history	51	0.055	1.917

Figure 8
Mann Whitney (U-scores) for significant relational cues in patient’s six demographically variant groups.

tional cues on live clinical calls and Dr Evie. Since the dialogue sets are those used for creating the Dr Evie application, they cannot be compared in terms of actual sessions that patients have with Dr Evie. The whole dataset provides a statistical presence of relational cues, which are shown in Figure 9.

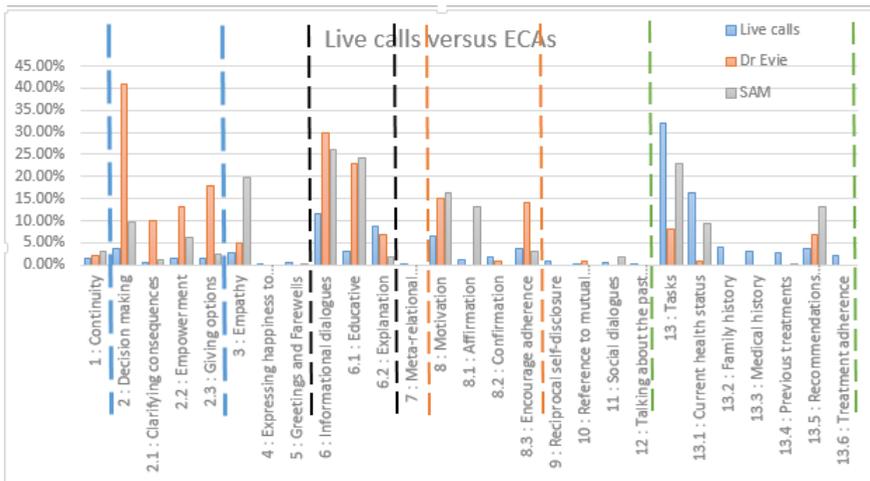


Figure 9
Coding comparison of live calls, Dr Evie and SAM. Dotted segregation is for each parent theme and corresponding sub themes. Sub themes counts add up to make parent theme counts.

The dialogue set of Dr Evie was designed by experienced health professionals who are experts in incontinence and other related disorders; hence, percentage usage of dialogue cues in live clinical calls is similar to that of relational cues in Dr Evie, especially for education, empowerment, encouraging adherence and giving options. The live clinical calls had a lower percentage usage of cues such as reference to mutual knowledge, greetings and farewells and the same can be seen with Dr Evie. The other cues such as empathy, explanation, recommendation, confirmation and clarifying consequences are also present in Dr Evie in good proportion and comparable with their usage in live clinical calls. Relational cues such as talking about the past and future together and social dialogues are missing in Dr Evie's context, Dr Evie's use of relational cues could be enriched by recommendations from live calls.

SAM's dialogue sets are mostly around sleep routine management and diet options: caffeine intake, regular sleep, night terrors, sleep hygiene and snoring issues. The recommendations are also about the sleep routine and diet habits. The dialogue flows have informational dialogues, current health assessment and recommendation dialogues incorporated in empathic language (e.g. 'I understand' and 'I know this is hard but'). SAM's dialogues are rich in empathic cues and affirmation dialogues, but lack cues for encouraging adherence, clarifying consequences and explanation. Live calls dialogue cues for these themes can be used to enrich SAM's dialogue set.

5.5 Structure, Topic and Inter-speaker Relationship Analysis

Health consultations have a defined and specific structure that exhibits the ontology of the subject area (Bickmore, Gruber, and Picard 2005). In order to embed relational cues in a logical manner, it is important to understand the conversational structure, topic variance and inter-speaker contribution in the overall consultation. The recurrence of topics and the time taken by speakers determine the level of engagement and understanding among the participants.

The dialogue structure in health consultations presented in Baker, Richards, and Caldwell (2014) places the dialogue cues from Bickmore et al. (2010) into a structure found in real consultations. Starting with greetings and farewells, social dialogues and previous treatment-related dialogues, it continues into more empathic dialogues and reciprocal self-disclosure cues. The last part of the conversation is more about future recommendations, adherence and continuity of the consultations. Our dataset is also mapped onto this logical structure to validate its existence in live calls, which suggests that even if the ECAs are built on natural language instead of structured questions, they would follow a similar structure as shown in Figure 10.

6. Discussion

The role of relational cues was analysed by the percentage of their occurrence in conversations. The analysis showed that two relational cues were used in all conversations but their percentage of usage in terms of utterance count by each role is very low – greetings, farewells and continuity. This confirms the (Laver 2011) finding that social dialogues are uttered mainly during the greeting and farewell phases of a conversation. Four relational cues were used very rarely in all conversations – expressing happiness to see the users, reference to mutual knowledge, talking about the past and future together and reciprocal self-disclosure. These cues are highly connected to how long term the relationship is with the patient and the level of comfort. The highest usage of relational cues is for empathy and motivational, informational and decision-making

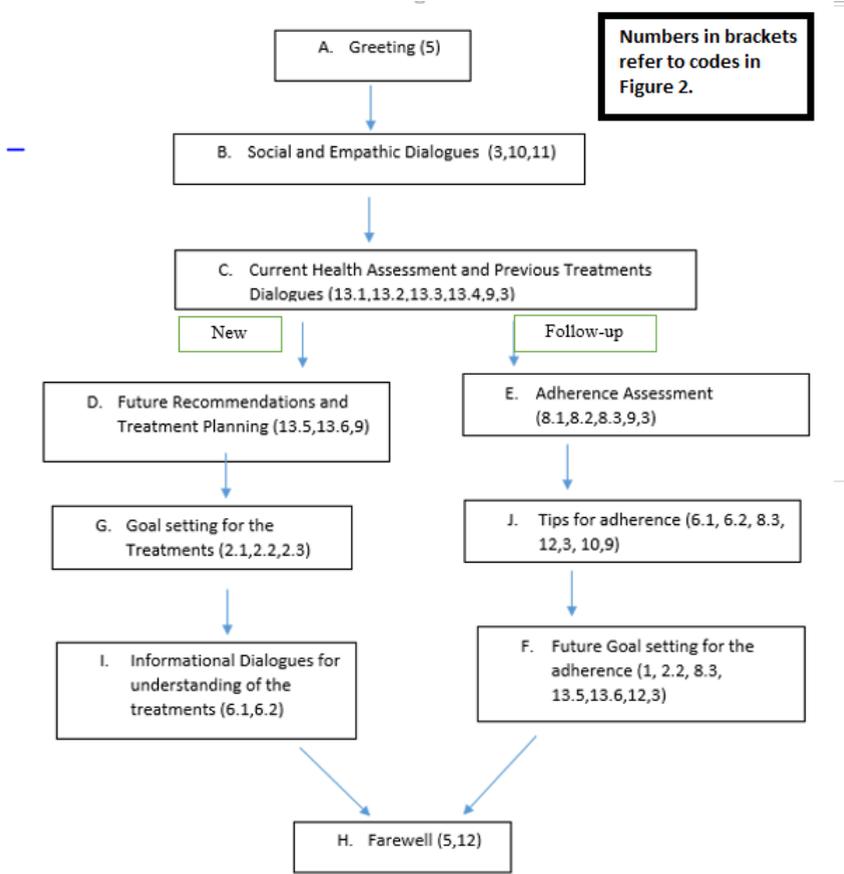


Figure 10
The Dialog Structure of the Live Calls

codes. Relational cues differ according to the session type and whether it is with a new or follow-up patient (Norfolk, Birdi, and Walsh 2007). A new patient needs to be onboarded with more communication related to treatment mechanics and education of the health issues. A follow-up patient needs more encouragement and dialogues related to clarifying consequences.

This analysis provided detailed findings on roles and demographics in the live calls. These insights can be used to create customised responses for different gender and age groups. Two factors that make the ECAs humanlike are their appearance and intelligence (Raval 2020), which depend highly on dialogue management producing dialogues similar to actual human dialogues. The live calls were more contextualised and customised. They included cues such as self-disclosure, confirmation and explanation. If the ECA dialogues were modified to capture more of a patient’s context, these cues could be used to enrich Dr Evie and SAM dialogues.

Multiple roles can be introduced in Dr Evie and SAM, as the senior doctor uses more explanation and empowering cues and the junior doctor takes care of certain routine tasks. This can make dialogue sets more empathic. To implement an approach where multiple ECAs can support a patient in different ways to provide holistic care, the

Council of Coaches platform could be used (op den Akker et al. 2018) where multiple ECAs review the patient's situation together and have separate conversations with the user according to their specific area of expertise or role (e.g. dietician, physiotherapist, friend).

The above suggested extensions to Dr Evie and SAM, such as more data gathering leading to more personalised use of relational cues, can be applied more widely to the design of other ECAs. This would involve the inclusion of user/patient models that persist between consultations, which allow the knowledge of the ECA to grow and be updated each time it meets the patient. This would also allow the ECA to tailor its social dialogue to the interests of each human. The relational cues and approach used in this article can be used to evaluate other ECAs and more relational cues could be incorporated to improve the relationship built with the user, with the aim of improving health outcomes. Another contribution of this paper is the capture of valuable datasets. The dataset and findings can be used by others in different ways. It is evident that live calls contain relational cues (e.g. disclosure and social dialogue) and dialogue patterns that could be used to design relational dialogue for ECAs. This dataset can also be used in machine learning and AI-based agents (Van Welbergen, Yaghoubzadeh, and Kopp 2014), which can learn to respond and formulate conversation using natural language generation. The live-call transcripts could also be used for the medical training of patient-doctor communication. This would be an alternative or supplement to approaches that offer guidance for health practitioners such as those provided by Rogers' (1959) client-centred therapy, which includes empathy, genuineness and unconditional acceptance. There is clear overlap in some of the relational cues and in the approach suggested by Rogers.

Finally, live clinical calls provide utterances from various roles. These roles include not only health specialists but also patients. Hence, the data can be used to build multiple ECAs that have different roles to facilitate the training of practitioners. While the quality of the recordings of patients and family members was too poor to allow transcription and qualitative analysis and outside the scope of this study due to its focus on practitioner use of relational cues, some specific calls in the live-call dataset can be used to build virtual patient ECAs, so that doctors can practise and refine their patient-doctor conversational expertise.

7. Limitations and Future Work

The dataset of live clinical calls was collected during the COVID-19 timeframe when normal clinical practice was disrupted. Live online clinical consultations became the new norm, which made their recording possible. Nevertheless, not all patients or practitioners were comfortable with this form of consultation, and it is possible that the dialogues were different to what might have been recorded in live face-to-face sessions in consultation rooms. The delay at the start to obtain consent prior to recording may have also inhibited the naturalness of the conversation and relationship. The ECAs for incontinence and sleep have been studied in conjunction with live calls, but more ECAs in domains specific to children should be explored to establish the use of relational cues especially in terms of the health specialist. As recommended by an anonymous reviewer, in the future it would be interesting to analyse empathy in a diachronic perspective, analysing the evolution of elements of empathy in doctor-patient conversations over time.

Conversational unit interfaces (CUIs) in health care are able to analyse natural languages (Laranjo et al. 2018) and to build responses according to the patient's situation

and history. As we move into the digital era, reliance on virtual agents that talk and understand like humans is a big research area (Sas, Whittaker, and Zimmerman 2016). Dr Evie uses scripted dialogues, whereas SAM uses more sophisticated technology that takes into account the user's goals and beliefs. The architecture SAM uses also allows preferences, medical history and other contextual features to be included in the ECA's reasoning and to provide explanations (Abdulrahman and Richards 2019). As Sam was developed later, we tried to incorporate more empathy and affirmation cues into its dialogues. Now that we have the data from live calls, Sam can further be enriched with encouraging adherence and clarification cues.

SAM and Dr Evie use fixed choice responses, primarily to ensure patient safety and accuracy, which is a current risk in health domains due to limitations in natural language processing (Xu et al. 2020). In the future when these limitations are addressed, safe and reliable solutions that use natural language input can be evaluated with a mix of controlled responses.

In the future, the recommendations for ECA dialogue design can be utilised to produce more dialogue sets that can be generalised over certain situations and cultures. Since negative thoughts can aggravate health issues, an empathic ECA that is personalised to the individual could help both mental and physical well-being. Hence, future agents for all health issues can potentially benefit from the relational cues and their usage presented in this paper.

8. Acknowledgments

Thanks to the health professionals, patients and families who agreed to have their consultations recorded

References

- Abdulrahman, Amal and Deborah Richards. 2019. Modelling working alliance using user-aware explainable embodied conversational agent for behaviour change: Framework and empirical evaluation. In *40th International Conference on Information Systems, ICIS 2019*, pages 1–17, Munich, Germany, December. Publisher Association for Information Systems.
- Abramovitch, Henry and Eliezer Schwartz. 1996. Three stages of medical dialogue. *Theoretical Medicine*, 17(2):175–187.
- Alam, Firoj, Morena Danieli, and Giuseppe Riccardi. 2018. Annotating and modeling empathy in spoken conversations. *Computer Speech & Language*, 50:40–61.
- Artstein, Ron and Massimo Poesio. 2008. Inter-Coder Agreement for Computational Linguistics. *Computational Linguistics*, 34(4):555–596, 12.
- Baker, Scott, Deborah Richards, and Patrina Caldwell. 2014. Putting a new intelligent virtual face on a medical treatment advice system to improve adherence. In *Proceedings of the 2014 Conference on Interactive Entertainment*, pages 1–9, Newcastle NSW Australia, December.
- Bardovi-Harlig, Kathleen. 2010. Exploring the pragmatics of interlanguage pragmatics: Definition by design. *Pragmatics across languages and cultures*, 7:219–259.
- Bengtsson, Mariette. 2016. How to plan and perform a qualitative study using content analysis. *NursingPlus Open*, 2:8–14.
- Bennett, Jennifer K., Jairo N. Fuertes, Merle Keitel, and Robert Phillips. 2011. The role of patient attachment and working alliance on patient adherence, satisfaction, and health-related quality of life in lupus treatment. *Patient education and counseling*, 85(1):53–59.
- Bickmore, Timothy, Amanda Gruber, and Rosalind Picard. 2005. Establishing the computer–patient working alliance in automated health behavior change interventions. *Patient education and counseling*, 59(1):21–30.
- Bickmore, Timothy W. 2004. Unspoken rules of spoken interaction. *Communications of the ACM*, 47(4):38–44.
- Bickmore, Timothy W., Lisa Caruso, and Kerri Clough-Gorr. 2005. Acceptance and usability of a relational agent interface by urban older adults. In *CHI'05 extended abstracts on Human factors in*

- computing systems, pages 1212–1215, Portland OR USA, April.
- Bickmore, Timothy W., Kathryn Puskar, Elizabeth A. Schlenk, Laura M. Pfeifer, and Susan M. Sereika. 2010. Maintaining reality: Relational agents for antipsychotic medication adherence. *Interacting with Computers*, 22(4):276–288.
- Bosco, Francesca M., Ilaria Gabbatore, Claus Lamm, Rosalba Morese, Giorgia Silani, and Soile Loukusa. 2015. Social cognition: from empathy to pragmatic ability. In *11th International Conference on Cognitive Science*, Torino, Italy, September.
- Butler, Richard J. 1998. Night wetting in children: Psychological aspects. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 39(4):453–463.
- Cameron, Rachel A., Benjamin L. Mazer, Jane M. DeLuca, Supriya G. Mohile, and Ronald M. Epstein. 2015. In search of compassion: a new taxonomy of compassionate physician behaviours. *Health Expectations*, 18(5):1672–1685.
- Cassell, Justine, Timothy Bickmore, Mark Billingham, Lee Campbell, Kenny Chang, Hannes Vilhjálmsson, and Hao Yan. 1999. Embodiment in conversational interfaces: Rea. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 520–527, Pittsburgh Pennsylvania USA, May.
- Coulehan, John L., Frederic W. Platt, Barry Egner, Richard Frankel, Chen-Tan Lin, Beth Lown, and William H. Salazar. 2001. “Let me see if I have this right...”: words that help build empathy. *Annals of Internal Medicine*, 135(3):221–227.
- Gilbertson, Jill, Kathryn Dindia, and Mike Allen. 1998. Relational continuity constructional units and the maintenance of relationships. *Journal of Social and Personal Relationships*, 15(6):774–790.
- Graneheim, Ulla H. and Berit Lundman. 2004. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse education today*, 24(2):105–112.
- Greer, Stephanie, Danielle Ramo, Yin-Juei Chang, Michael Fu, Judith Moskowitz, and Jana Haritatos. 2019. Use of the chatbot “Vivibot” to deliver positive psychology skills and promote well-being among young people after cancer treatment: randomized controlled feasibility trial. *JMIR mHealth and uHealth*, 7(10):e15018.
- Halpern, Jodi. 2007. Empathy and patient–physician conflicts. *Journal of general internal medicine*, 22(5):696–700.
- Harter, S. 1982. A developmental perspective on some parameters of self-regulation in children. *Self-management and behavior change: From theory to practice*, pages 165–204.
- Hartmann, Björn, Maurizio Mancini, and Catherine Pelachaud. 2005. Implementing expressive gesture synthesis for embodied conversational agents. In *International Gesture Workshop*, pages 188–199, Berder Island, France, May. Springer.
- Hoffman, Martin L. 2001. *Empathy and moral development: Implications for caring and justice*. Cambridge University Press.
- Horsch, Corine, Willem-Paul Brinkman, Rogier van Eijk, and Mark Neerinx. 2012. Towards the usage of persuasive strategies in a virtual sleep coach. In *The 26th BCS Conference on Human Computer Interaction 26*, pages 1–4, Birmingham United Kingdom, September.
- Kowatsch, Tobias, Marcia Nißen, Chen-Hsuan Iris Shih, Dominik Rügger, Dirk Volland, Andreas Filler, Florian Künzler, Filipe Barata, Dirk Büchter, Björn Brogle, et al. 2017. Text-based healthcare chatbots supporting patient and health professional teams: Preliminary results of a randomized controlled trial on childhood obesity. In *Conference: Persuasive Embodied Agents for Behavior Change (PEACH 2017) Workshop, co-located with the 17th International Conference on Intelligent Virtual Agents (IVA 2017)*, Stockholm, Sweden, August.
- Kredlow, M. Alexandra, Michelle C. Capozzoli, Bridget A. Hearon, Amanda W. Calkins, and Michael W. Otto. 2015. The effects of physical activity on sleep: a meta-analytic review. *Journal of behavioral medicine*, 38(3):427–449.
- Laranjo, Liliana, Adam G. Dunn, Huong Ly Tong, Ahmet Baki Kocaballi, Jessica Chen, Rabia Bashir, Didi Surian, Blanca Gallego, Farah Magrabi, Annie YS Lau, et al. 2018. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association*, 25(9):1248–1258.
- Laver, John. 2011. Communicative functions of phatic communion. In *Organization of behavior in face-to-face interaction*. De Gruyter Mouton, pages 215–238.
- Lisetti, Christine, Ugan Yasavur, Claudia De Leon, Reza Amini, Ubbo Visser, and Naphtali Rishe. 2012. Building an on-demand avatar-based health intervention for behavior change. In *Twenty-Fifth International FLAIRS Conference*, Marco Island, Florida, May.

- Looije, Rosemarijn, Mark A. Neerincx, and Vincent de Lange. 2008. Children's responses and opinion on three bots that motivate, educate and play. *Journal of Physical Agents*, 2(2):8.
- Lundqvist, Lars-Olov. 1995. Facial emg reactions to facial expressions: A case of facial emotional contagion? *Scandinavian journal of psychology*, 36(2):130–141.
- Malhotra, Neha R., Karen A. Kuhlthau, Iliana Rosoklija, Matthew Migliozi, Caleb P. Nelson, and Anthony J. Schaeffer. 2020. Children's experience with daytime and nighttime urinary incontinence—a qualitative exploration. *Journal of Pediatric Urology*, 16(5):535–e1.
- Mamarimbing, Stevanus Natanael. 2021. *The humanistic approach by Erin Gruwell in the freedom writers movie*. Ph.D. thesis, Widya Mandala Surabaya Catholic University.
- Maternik, Michal, Katarzyna Krzeminska, and Aleksandra Zurowska. 2015. The management of childhood urinary incontinence. *Pediatric Nephrology*, 30(1):41–50.
- McRorie, Margaret, Ian Sneddon, Etienne de Sevin, Elisabetta Bevacqua, and Catherine Pelachaud. 2009. A model of personality and emotional traits. In *International Workshop on Intelligent Virtual Agents*, pages 27–33, Amsterdam, The Netherlands, September. Springer.
- Neff, Michael, Yingying Wang, Rob Abbott, and Marilyn Walker. 2010. Evaluating the effect of gesture and language on personality perception in conversational agents. In *International Conference on Intelligent Virtual Agents*, pages 222–235, Philadelphia, PA, USA, September. Springer.
- Norfolk, Tim, Kamal Birdi, and Deirdre Walsh. 2007. The role of empathy in establishing rapport in the consultation: a new model. *Medical education*, 41(7):690–697.
- Ochs, Magalie, Catherine Pelachaud, and Gary Mckeown. 2017. A user perception-based approach to create smiling embodied conversational agents. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 7(1):1–33.
- op den Akker, Harm, Riëks op den Akker, Tessa Beinema, Oresti Banos, Dirk Heylen, Björn Bedsted, Alison Pease, Catherine Pelachaud, Vicente Traver Salcedo, Sofoklis Kyriazakos, et al. 2018. Council of coaches a novel holistic behavior change coaching approach. In *4th International Conference on Information and Communication Technologies for Ageing Well and e-Health, ICT4AWE 2018*, pages 219–226, Porto, Portugal, March. SciTePress.
- Paiva, Ana, Joao Dias, Daniel Sobral, Ruth Aylett, Sarah Woods, Lynne Hall, and Carsten Zoll. 2005. Learning by feeling: Evoking empathy with synthetic characters. *Applied Artificial Intelligence*, 19(3-4):235–266.
- Planalp, Sally and Anne Benson. 1992. Friends' and acquaintances' conversations i: Perceived differences. *Journal of Social and Personal Relationships*, 9(4):483–506.
- Raval, Raturaj. 2020. An improved approach of intention discovery with machine learning for pomdp-based dialogue management. *arXiv preprint arXiv:2009.09354*.
- Ravichander, Abhilasha and Alan W. Black. 2018. An empirical study of self-disclosure in spoken dialogue systems. In *Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue*, pages 253–263, Melbourne, Australia, July.
- Richards, Deborah and Patricia Caldwell. 2017. Improving health outcomes sooner rather than later via an interactive website and virtual specialist. *IEEE journal of biomedical and health informatics*, 22(5):1699–1706.
- Rogers, Carl Ransom et al. 1959. *A theory of therapy, personality, and interpersonal relationships: As developed in the client-centered framework*, volume 3. McGraw-Hill New York.
- Roller, Stephen, Emily Dinan, Naman Goyal, Da Ju, Mary Williamson, Yinhan Liu, Jing Xu, Myle Ott, Kurt Shuster, Éric M Smith, et al. 2020. Recipes for building an open-domain chatbot. *arXiv preprint arXiv:2004.13637*.
- Roth, Thomas. 2007. Insomnia: definition, prevalence, etiology, and consequences. *Journal of clinical sleep medicine*, 3(5 suppl):S7–S10.
- Sas, Corina, Steve Whittaker, and John Zimmerman. 2016. Design for rituals of letting go: An embodiment perspective on dispositional practices informed by grief therapy. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 23(4):1–37.
- Stafford, Laura and Daniel J. Canary. 1991. Maintenance strategies and romantic relationship type, gender and relational characteristics. *Journal of Social and Personal relationships*, 8(2):217–242.
- Theunis, Marleen, Eline Van Hoecke, S Paesbrugge, Piet Hoebeke, and J Vande Walle. 2002. Self-image and performance in children with nocturnal enuresis. *European urology*, 41(6):660–667.
- Thibodeau, Betty Ann, Peter Metcalfe, Priscilla Koop, and Katherine Moore. 2013. Urinary incontinence and quality of life in children. *Journal of pediatric urology*, 9(1):78–83.

- van Velsen, Lex, Marijke Broekhuis, Stephanie Jansen-Kosterink, and Harm op den Akker. 2019. Tailoring persuasive electronic health strategies for older adults on the basis of personal motivation: Web-based survey study. *Journal of medical Internet research*, 21(9):e11759.
- Van Welbergen, Herwin, Ramin Yaghoubzadeh, and Stefan Kopp. 2014. Asaprealizer 2.0: The next steps in fluent behavior realization for ecas. In *International Conference on Intelligent Virtual Agents*, pages 449–462, Boston, MA, USA, August. Springer.
- Warrens, Matthijs J. 2015. Five ways to look at cohen’s kappa. *Journal of Psychology & Psychotherapy*, 5(4):1.
- Wispé, Lauren. 1987. History of the concept of empathy. *Empathy and its development*, 2:17–37.
- Woods, Nancy Fugate and Marci Catanzaro. 1988. *Nursing research: Theory and practice*. Mosby Incorporated.
- Wright, Peter and John McCarthy. 2008. Empathy and experience in hci. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 637–646, Florence, Italy, April.
- Xu, Jing, Da Ju, Margaret Li, Y-Lan Boureau, Jason Weston, and Emily Dinan. 2020. Recipes for safety in open-domain chatbots. *arXiv preprint arXiv:2010.07079*.
- Yin, Langxuan, Lazlo Ring, and Timothy Bickmore. 2012. Using an interactive visual novel to promote patient empowerment through engagement. In *Proceedings of the International Conference on the foundations of digital Games*, pages 41–48, Raleigh North Carolina, May.
- Zhanghong, Xu and Wang Qian. 2018. Pragmatic empathy as a grand strategy in business letter writing. *English Language Teaching*, 11(8):14–27.