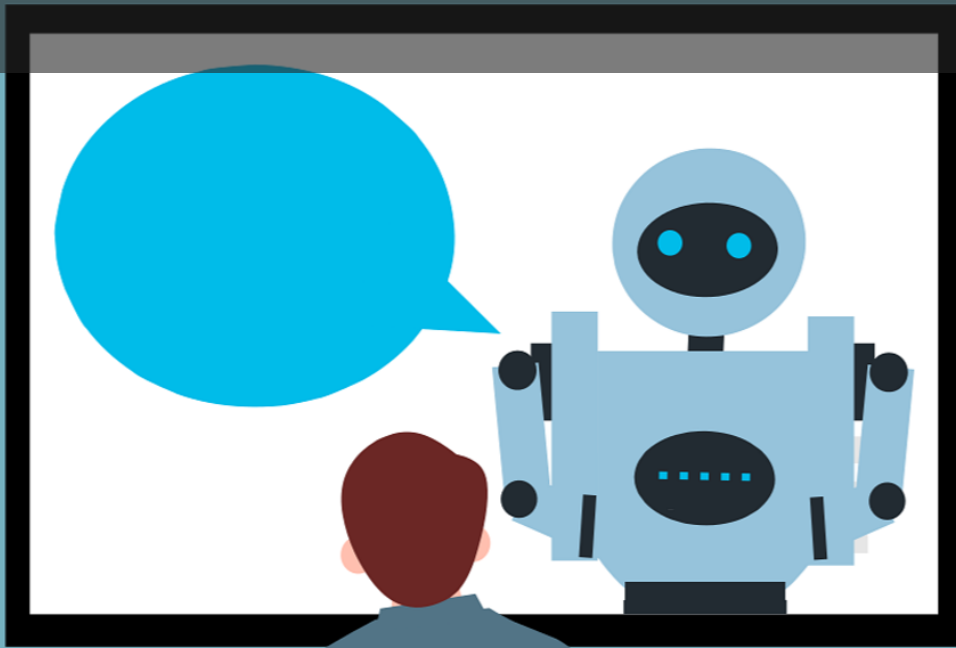


Emotion Model for Child Helpline Training Tool

Dongxu Lu



Emotion Model for Child Helpline Training Tool

by

Dongxu Lu

Student Name	Student Number
Dongxu Lu	5503426

Instructor: Dr. ir. Willem-Paul Brinkman
Teaching Assistant: Ir. Mohemmed Al Owayyed
Project Duration: November 2022 - August 2023
Faculty: Faculty of Electrical Engineering, Mathematics and Computer Science, Delft

Cover: Image by Mohamed Hassan from Pixabay.
Style: TU Delft Report Style, with modifications by Daan Zwaneveld

Acknowledgements

First and foremost, I would like to thank my supervisors Prof. Willem-Paul and Mo for giving me this opportunity to work and learn within a supportive group and for their excellent guidance as well as thorough comments during my research all the time. I still remember back in November 2022 when I just stepped into the mystery of emotional computing, uncertainty clouded my path. But now, I stand with clarity, which I would never manage to achieve without your invaluable help.

And to my friends and fellow colleagues. Thank you for your sincere encouragement and solace when I feel depressed. The dark nights are no longer enduring with your companion. We share the joy and sorrow in life together. Now that I have made it, I wish you the best of luck as well!

On a personal note, I would like to thank my parents for their support and encouragement throughout this journey. It is not easy to start a life in a country with a completely different culture. But I have been enjoying the novel experience and growth brought by cultural collision so far, with your love, understanding, and constant motivation. Thank you for letting me see the world.

*Dongxu Lu
Delft, August 2023*

Abstract

Child helplines play a crucial role in delivering expert assistance to young clients facing challenges and seeking support. While counselling is instrumental in enhancing childrens mental well-being, the limited number of experienced counsellors is inadequate given the substantial workload. At this point, effective training to volunteer counsellors becomes essential. This thesis aims to expand upon a Belief-Desire-Intention (BDI) based chatbot, the simulated child, with the specific objective of enhancing volunteer training at child helpline organizations through the integration of emotional capabilities. Through the collaboration with professionals from a Dutch child helpline, the Kindertelefoon, we identified the predominant emotion of frustration along with its underlying triggers. Armed with this insight, we proposed an Emotion-BDI model capable of dynamically adjusting the chatbots displayed emotion in response to its environment. The evaluation of a prototype constructed based on this model supported our hypotheses that a chatbot equipped to express frustration possesses greater believability and emotional presence compared to the original version. Moreover, it also indicated heightened user enjoyability, engagement, and perceived usefulness. Within this thesis, we highlighted the enhanced feasibility of the simulated child brought about by our proposed Emotion-BDI model.

Contents

Acknowledgements	i
Abstract	ii
1 Introduction	1
1.1 Background	1
1.2 Motivation	2
1.3 Research Question	2
1.4 Research Approach	2
2 Foundations	4
2.1 Focus Group Setup	4
2.2 Emotions in Children’s Conversations	4
2.3 Expressing Emotions Through Text	6
2.4 Understanding Emotions Using Cognitive Approaches	6
2.5 Cognitive and Emotion Modelling in Conversational Agents	7
2.5.1 Computational Emotion Modelling	7
2.5.2 Extending emotion on BDI Architecture	8
2.6 Summary	9
3 Design	10
3.1 The Conceptual Architecture	10
3.1.1 Core components	11
3.1.2 Reasoning Flow	13
3.2 Emotion Update under Frustrating Scenarios	15
4 Evaluation	18
4.1 Methods	18
4.1.1 Experimental Design	19
4.1.2 Materials	19
4.1.3 Measures	20
4.1.4 Participants	21
4.1.5 Procedure	21
4.1.6 Data Preparation and Statistical Analysis	22
4.2 Results	24
4.2.1 Believability	24
4.2.2 Emotional Presence	24
4.2.3 Enjoyability, Engagement and Perceived Usefulness	24
4.2.4 Qualitative Analysis	25
4.3 Discussions	26
4.4 Limitations	27
5 Discussion & Conclusion	28
5.1 Conclusion	28
5.2 Contribution	29
5.3 Limitations	29
5.4 Future Work	29
References	31
A Design Appendices	34

B	Evaluation Appendices	35
B.1	Prototype	35
B.2	Informed Consent Form	35
B.3	Instructions on experiment scenarios	39
B.4	Questionnaires	44
	B.4.1 Believability	44
	B.4.2 Emotional Presence	44
	B.4.3 Qualitative Measure	44

1

Introduction

1.1. Background

Child helplines from all over the world are caring for children’s mental health now. The Kindertelefoon (KT) in the Netherlands offers professional advice and assistance to children and adolescents between the ages of 8 and 18 via phone and online chat services. With thousands of calls and chats to handle every day, the number of interactions is steadily increasing, as indicated in the organization’s annual report [25]. To manage this demand, KT employs around 700 volunteers and 50 paid employees, with the majority of counselling duties assigned to volunteers who receive supervision from employees. However, providing counselling to children who are struggling with sensitive emotions, such as sadness, anger, and frustration, is challenging. Emotions such as sadness, anger, and frustration require a nuanced approach that demands exceptional communication skills. Lefevre et al. [28] emphasized that successful counselling needs establishing a rapport that helps children feel genuinely understood and empowers them to participate actively in the decision-making and planning processes. Navigating these complex emotional landscapes indicates the importance of robust training strategies that equip counsellors with the tools to facilitate effective conversations with young clients.

Regarding these challenges, the development of a simulated child in frustration as a training tool for counsellors holds significant promise. This AI-based approach offers various advantages that can enhance the training of counsellors. Firstly, a simulated child offers a cost-effective alternative to traditional training practices involving real children. A simulated child can be engaged repeatedly without restrictions on time and location, providing counsellors with sufficient flexibility for practice and skill refinement [4]. This repetitive engagement not only can foster the acquisition of essential skills but also enables counsellors to encounter a diverse range of scenarios, thereby avoiding unnecessary mistakes during real-life counselling situations [58].

Besides, the introduction of a simulated child in frustration mitigates potential stress and anxiety that may arise from training sessions involving actual children. Interacting with real children, particularly those with sensitive emotions can evoke emotional and psychological strain on counsellors who may be unaccustomed to such encounters. In contrast, a simulated child offers a controlled environment where counsellors can practice with higher communicative confidence without the concerns of impacting a real child’s emotional well-being [55]. As a result, this training tool can enhance the efficacy of the training and bring convenience to counsellors.

The presence of conversational agents and their diverse applications in various everyday contexts, including education [57], healthcare [4], and entertainment [16], has inspired a growing interest in the field of modelling agent cognition [1]. In this regard, the Belief-Desire-Intention (BDI) model introduced by Norling [38], has gained prominence as a conceptual framework for simulating human-agent interactions. Based on folk psychology, the BDI model facilitates the translation of human reasoning explanations into fundamental agent framework concepts within the context of everyday conversations [38]. Acting as a high-level abstraction of human cognitive processes, the BDI model captures the essence of general reasoning scenarios. However, it does exhibit limitations in comprehensively enclosing all the factors involved in human decision-making, particularly in areas such as emotions, which play a pivotal role in shaping the decision-making of individuals [56]. To address this gap, our study explored the integration

of emotional factors into the development of the simulated child.

1.2. Motivation

Previous research on affect consistently emphasized the profound impact that emotions show on people's judgments and subsequent decision-making processes [3]. However, the role that emotions play in social interactions is frequently underestimated by individuals [3]. As a fundamental aspect of human communication, emotional responses can help train counsellors to better understand and build more genuine connections with their clients [23]. Besides, emotions are often closely intertwined with the decision-making process, including individuals' beliefs, desires, and intentions, thereby contributing to the internal mechanisms of action selection according to Anderson et al.'s affective model [2]. The dynamic relationship between emotions and decision-making highlights their role in shaping the choices individuals make, demonstrating the significance of integrating emotional intelligence into counsellor training.

Acknowledging the crucial role of emotions, we propose that equipping counsellors to identify the impact of emotions on children's beliefs and decision-making processes holds immense value. By allowing counsellors to respond to various emotional cues, training can equip them with the necessary tools to recognize, validate, and effectively address the diverse emotional states exhibited by their clients. Hence, the principal objective of this thesis is to augment the existing BDI-based agent by integrating an emotion model alongside its existing belief states, thereby simulating a child-like representation infused with emotional responses. Expanding upon prior endeavours involving the development of a simulated child as a conversational agent to train child helpline counsellors using the BDI model [17], we seek to advance the framework by infusing the model with a stronger emphasis on emotions, particularly in the context of instances involving bullying. This simulation aims to facilitate counsellors' understanding of the ways emotions influence conversations, prompting increased attention to the dynamics of emotion fluctuation.

1.3. Research Question

The objective of this study is to investigate the development of an emotion model tailored for a BDI-based conversational agent, specifically designed for counsellor training at child helplines. This emotion-rich construct is intended to play an important role in shaping conversational dynamics, ultimately contributing to an enhanced training experience for trainee counsellors. Through the implementation of the model, we aim to provide a training tool that promotes the development of effective counselling skills. The research question guiding this study is formulated as follows:

How can an emotion model be applied to a child-simulated conversational agent based on the BDI model to assist with the training of child helpline volunteers?

This research question is then broken down into the following sub-questions:

- *What design factors should be taken into account when developing a simulated child that incorporates emotion?*
- *How can emotions be modelled in a simulated child concerning the design factors?*
- *To what extent do trainee counsellors perceive the simulated child with emotions to be believable and emotionally present?*

1.4. Research Approach

To address the first sub-question, we conducted a comprehensive literature review of current techniques in affective computing and emotion modelling aimed at gathering insights into potential approaches for implementing an emotion model infused with the BDI model. In addition to the literature review, we sought the expertise of professionals from the child helpline organization, the Kindertelefoon by organizing a focus group. This allowed us to gather valuable knowledge and receive suggestions related to the work of child helplines and children's emotions (Chapter 2). Based on the insights gained from the literature review and expert consultations, we proceeded to design the emotion model (Chapter 3) aimed to answer the second sub-question by developing a framework that could effectively capture

and represent the emotional dynamics in child helpline conversations. Following the design, we conducted a within-subject experiment for the evaluation of the emotion model and analyzed the collected data using various statistical methods, including pairwise t-tests and comparisons with existing chatbots (Chapter 4), addressing the third sub-question. Finally, in Chapter 5, we presented the conclusions on the research question drawn from our work contributions and provided feasible suggestions for future research and development in this area.

2

Foundations

In this chapter, we aim to the first sub-question of our research question:

- *What design factors should be taken into account when developing a simulated child that incorporates emotion?*

In pursuit of our objective, we examined how emotion works from a cognitive perspective by reviewing prior research on cognitive theories of emotions. Subsequently, we identified the prevalent emotions displayed by children during conversations with the KT team through a focus group aimed to elicit the experience with emotional children of professionals in the relevant field and fill gaps in counsellor training with their knowledge. Additionally, after investigating previous research on how other scholars implement emotion modelling, we explored various techniques that connect with the BDI model and can be used to support our design theoretically. Finally, we presented a summary of the design factors derived from the focus group and literature, which was utilized in the model's design.

2.1. Focus Group Setup

The focus group included an experienced counsellor and two expert trainers. During the focus group session, we presented a scenario to the participants that featured text conversations between a counsellor who make mistakes during a consultation and two simulated children who experienced school bullying, which is shown in Figure 2.1. In the simulated scenario where the counsellor gives a solution without considering the desires of the children, the two simulated children, Child A and Child B, exhibit different patterns of emotional expression. While Child A's emotions fluctuate throughout the conversation, Child B maintains a more consistent emotional state, which was conveyed through their different styles of expression. Participants were then asked to indicate which of the simulated children would better facilitate the training of trainee counsellors in the KT according to their experience and provide reasons for their choices. Through the ensuing discussion, we were able to gather valuable opinions on the applications of emotion and specify the design factors for an emotionally intelligent agent to aid in the further development of the emotion model.

2.2. Emotions in Children's Conversations

According to the 2021 annual report of the KT, there were chats and calls addressing various issues, with a significant percentage of conversations involving emotional problems like depression and loneliness [25]. Since it's crucial for counsellors to offer professional guidance and emotional support when handling such problems, the training for trainee counsellors focuses on teaching them how to structure conversations with children using a Five-phase model [54]. This model serves as a guideline for leading conversations to help them clarify their problems and sort out their needs more clearly and logically. The first objective of a counsellor, when a child reaches out to a helpline, is to establish a connection and create a sense of validation for the child. After this initial step, the counsellor digs into the child's specific problem and context through a broad exploration of the child's circumstances, emotions, and thoughts and then narrows down the focus to address the most pressing needs. Once the

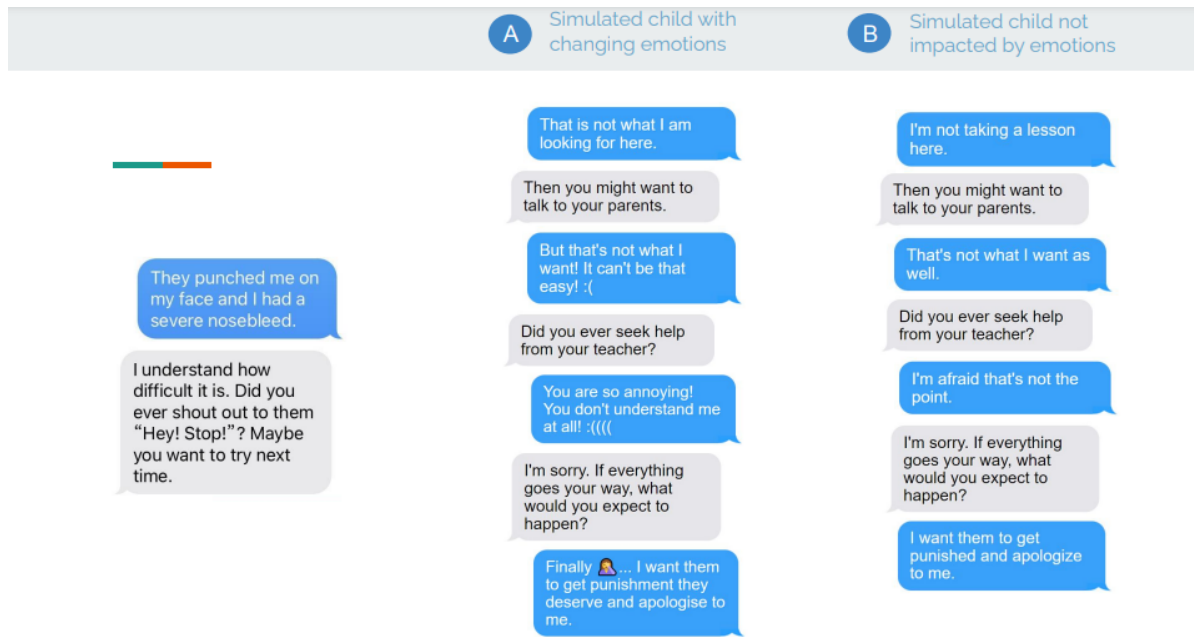


Figure 2.1: Scenario regarding children's emotions.

counsellor has a clear understanding of the problem, they collaborate with the child to establish a goal for the session that they can work towards. The overview of the model is shown in Table 2.1.

Phase 1: Building rapport
Objective: Create a welcoming atmosphere and build trust.
Method: Empathy, respect, sincere interest, active listening.
Phase 2: Clarify the child's story
Objective: Get a clear view of the child's story, perspective, personality, network and competencies.
Method: Ask detailed questions about the child's story, its subtleties, its depth and its concrete manifestations.
Phase 3: Setting a goal for the session
Objective: Both parties are aware of what the child may use the conversation for.
Method: Clarification.
Phase 4: Working toward the session goal
Objective: To ensure, to the widest possible extent, that the child may benefit from the conversation.
Method: Stimulating the child's own problem-solving skills.
Phase 5: Rounding off the conversation
Objective: The child is left with as few questions as possible.
Method: Summing up and clarifying.

Table 2.1: Overview of the five phase model from [54].

According to the focus group discussion on children's emotions, frustration is the most frequently expressed emotion by children seeking help. Table 2.2 illustrates the various situations that may cause a child to feel frustrated during a consultation, where each situation corresponds to certain phases of a conversation. The experts at KT observed that the most common situation leading to frustration is when a child expects the counsellor to provide a direct solution but gets denied. Other situations that can lead to frustration include a lack of empathy or acknowledgement from the counsellor towards the child's story, which may result in a lack of trust, and a counsellor being too eager to offer a solution without considering the child's wishes. In addition, frustration can also occur when the counsellor has difficulty understanding the child, such as when the counsellor asks the child to rephrase their statements constantly or overly focuses on finding answers to detailed questions.

According to experts, KT counsellors often find that children are very expressive and willing to share their thoughts during conversations. This may be because children perceive counsellors as impartial and non-judgemental, making it easier for them to discuss their concerns. However, if a counsellor is unable to facilitate such conversations and the child becomes frustrated, it can lead to increased resistance and avoidance displayed in the conversation. This idea aligns with Langdon's finding [31] that higher levels of frustration can result in more intense and lasting reactions.

Lemerise and Arsenio [29] suggested that children experience a range of emotions in social interactions. Thus, integrating emotions into virtual agents could create a training environment that closely resembles reality for trainee counsellors. In addition, it was found by Djordjevic et al. [13] that the inclusion of emotions in virtual agents can offer a more precise representation of the consequences of trainee's decisions. This aligns with the consensus reached by experts at KT that an emotional agent can be a useful training tool for trainee counsellors. By simulating frustration as a negative emotion during training, the trainee counsellors can be reminded that one of the situations in Table 2.2 may occur, indicating that the conversation is not progressing well and requires improvement.

Situations that cause frustration	Phase
KT does not show enough empathy or acknowledgement	Phase 1
The child wants KT to determine a solution but gets rejected	Phase 3
KT sets up a goal for the child without asking	Phase 3
KT asks too many questions	All phases

Table 2.2: Situations that induce frustration.

2.3. Expressing Emotions Through Text

Although chat-based communication is omnipresent in modern life, it can be challenging to express emotions through text because it lacks verbal cues such as tone, voice pitch, and volume. In an experiment by Hancock et al. [18], four strategies were identified that people use to express happiness versus sadness in text-based communication, including the frequency of disagreement, the count of negative affect terms, the use of punctuation, and speed of response and verbosity. Results also indicated that negative expressers disagree significantly frequently, use more exclamation marks and negative affect terms, and produce more words overall than positive expressers. In addition, emoticons are among the non-verbal cues that people use to express emotions. These graphic representations of facial expressions were defined as icons for expressing emotions or indicating non-serious intent [10]. The most popular emoticons include the smiley face, winking face, and frowning face, represented by [:-)], [;-)], and [:-()] respectively. Yus's study on emoticons [59] has revealed that the repetition of characters inside the emoticon can signify the intensity of the emotion expressed. For instance, the emoticon [:-))))) is perceived as expressing more intense joy than [:-)]. However, when the repetition of characters increases beyond a certain point, people do not distinguish the intensity of the emotion expressed.

2.4. Understanding Emotions Using Cognitive Approaches

Emotions have long been a topic of keen academic interest due to their significant influence on our daily lives. In 1962, Schachter and Singer argued that cognition plays a critical role in emotions and that comprehending a person's cognitive interpretation of a situation is necessary to understand their emotions [51]. To provide clear perspectives on emotions and address fundamental questions regarding their causes and effects, cognitive approaches targeted at understanding how the mind works and how people acquire, process, and use information have emerged as a solution [40].

Arnold first introduced the term "appraisal" in 1960 to highlight the idea that the type of attention focused on an object determines the emotional experience [32]. This understanding of emotions has given rise to theoretical cognitive approaches known as appraisal theories, which suggest that emotions are adaptive responses reflecting appraisals of environmental features that promote an organism's well-being [35]. Emotions are characterized as a series of changes in various components that are closely interconnected for the evaluation of the environment in appraisal theories.

Nonetheless, previous research has suggested that emotions are complex and multifaceted, and can be influenced by a wide range of characteristics of humans [41]. One key characteristic that can have

a significant impact on emotions is personality. For example, individuals with certain temperamental traits of personality may be more prone to experiencing particular emotions and may feel these emotions more intensely, whether positive or negative [49]. Another critical characteristic that has attracted researchers' attention is mood, with Neumann et al. [37] stating that an individual's pre-existing mood can amplify congruent emotions while reducing the intensity of incongruent ones. Therefore, if mistreatment puts an agent in a negative mood, the agent will become more likely to display frustration.

2.5. Cognitive and Emotion Modelling in Conversational Agents

Conversational agents (CAs) are software systems that enable humans to interact using human language through text-based interfaces such as chatbots or speech-based interfaces like virtual assistants [15]. Early examples of CAs were built targeted at being indistinguishable from a human to pass the Turing test where human users interact with the agent and decide whether they are talking to a human or a machine [33]. Thanks to the development of AI, CAs have evolved significantly and researchers have started applying CAs to more complex tasks in domains such as healthcare [27], education [24], and customer service [15], where they have proven to be valuable assets. To perform these complex tasks, CAs need to acquire the ability to perceive and reason about their environment and take actions accordingly, in addition to their conversational skills [46].

Cognitive modelling is an approach that can support the development of CAs with these capabilities. As an important research branch of human-AI interactions, it is often used to model complex cognitive processes, such as memory, attention and decision-making [46]. By understanding and modelling how the human mind functions, cognitive modelling can offer solutions to a wide range of problems by drawing on a range of knowledge sources, including psychology, neuroscience, and human-AI interaction [26]. With this knowledge, intelligent CAs can be built to simulate how humans perceive and store information and provide support for more efficient decision-making in a given scenario.

2.5.1. Computational Emotion Modelling

Emotions are a familiar folklore-psychology concept to individuals based on daily experiences, but defining them proves to be a challenge. Rather than defining emotions by their essence, researchers reach a tacit agreement to describe them in terms of their characteristics and how they manifest in patterns of expression triggered by cues. In 1988, Minsky proposed that "the question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without any emotions" [34]. The ultimate goal of incorporating emotions into intelligent agents is to improve their believability and realism [46]. What marks the believability of an emotional agent lies in its capability of generating plausible behaviour and consistency across similar situations [42].

However, due to the vague terminology of emotion, diverse approaches come into being, such as black-box models that map environmental stimuli to emotions, generation of facial expressions that convey emotions, and studying the effects of emotions on decision-making, all falling under the umbrella of "emotion modelling" [19]. To address the problem, Hudlicka [19] proposed a framework to categorize emotion modelling into emotion generation and emotion effects. This framework focuses on the causes and impacts of emotions, respectively, providing a more precise way to classify and study emotion models.

Many studies on emotion generation draw inspiration from the appraisal theory. Appraisal-based emotion models proposed that agents dynamically evaluate events in the environment based on a distinct set of appraisal dimensions, which vary significantly among different models. The most influential and commonly used computational models of emotions were first introduced by Ortony et al. [43] in 1988 and are often referred to as the OCC theory where emotions are defined as valenced responses to events, agents, and objects. An OCC-based emotion model classifies emotions based on the appraisal features of the three types of stimulus: objects, agents, and the event itself. In addition, each stimulus is evaluated based on various appraisal criteria, which can result in a vast array of different appraisal patterns, therefore, a large space of possible emotional states [52].

Emotional effects are another focus of emotion modelling, which are generally categorized into two groups: visible behavioural expressions and less visible effects on cognitive processes like attention, perception, and cognition [19]. While research on the emotional effects that focus on modelling behavioural expressions on a virtual agent, such as facial expressions, gestures, and action tendencies, are relatively easier to model [45], few studies have explored the impacts that emotions have on other cognitive pro-

cesses. For instance, Paulus and Yu’s work [44] shed light on how depression and anxiety can affect human decision-making by integrating an emotional framework based on the valence and arousal of depression and anxiety with a Bayesian approach to decision-making in terms of probability and value processing.

2.5.2. Extending emotion on BDI Architecture

The BDI architecture is a widely-used framework stemming from the theoretical foundation of Bratman’s agent-oriented work on rational reasoning and is well-known for its simplicity in shaping human behaviours based on three fundamental mental attitudes, *beliefs*, *desires* and *intentions* [5]. Here, *Beliefs* represent the information about the current state of the environment. *Desires* indicates the motivational state concerning what the agent wants to achieve. *Intentions* refers to the set of executable plans intended to achieve *desires*. The proposal of BDI architecture has inspired much work in implementing BDI-based agents to various practical applications such as a tourist guide system [9], crowd simulation for emergency response [53], and character creation [39].

One of the advantages of using cognitive appraisal models as a foundation for computational models that Nair et al. [36] claimed is the ease with which appraisals can be connected to the BDI framework in agent systems. The emotional extensions of the BDI architecture are often referred to as Emotional BDIs or EBDIs. Emotion modelling techniques connected to the BDI model focus more on the simulation of the emotional state of the virtual agents based on their beliefs, desires, and intentions. Based on the literature survey conducted by Sanchez and Eva [50], BDI extensions mostly rely on the OCC theory, either solely or in combination with other theories. The proposals that follow the OCC theory typically undergo a similar process of emotion elicitation: the agent first appraises an event from the environment depending on the three BDI model components, beliefs, desires and intentions. This is followed by the computation of a core emotional valence, which could be negative or positive, displeasure or pleasure. However, these proposals differ in the emotions they model, as well as the modelling of other factors such as personality and mood.

An EBDI architecture that Jang et al. [22] proposed includes two critical aspects of practical reasoning in the BDI model: deliberation and means-ends reasoning. Deliberation involves the agent’s determination of desired states of affairs, while means-ends reasoning focuses on selecting appropriate actions to achieve those desired states. By incorporating these processes, the BDI architecture enables agents to make informed decisions aligned with their beliefs, desires, and intentions.

There are two emotions involved in the proposed model, primary and secondary emotions, which constitute integral components of the EBDI architecture. Primary emotions are instinctive and immediate emotional responses that occur before conscious thought. They are often considered basic or fundamental emotions that are universally experienced across cultures. Examples of primary emotions commonly recognized in psychological theories include fear, anger, joy, sadness, surprise, and disgust. The determination of secondary emotions in the EBDI architecture may involve considering the cognitive appraisal theories of emotions that emotions are generated through the evaluation or appraisal of specific events or situations. Different appraisal dimensions, such as the evaluation of goal relevance, coping potential, or moral implications, can contribute to the categorization of secondary emotions. Examples of secondary emotions include guilt, shame, pride, jealousy, and hope. The inclusion of primary emotions facilitates expedited decision-making, while secondary emotions refine decisions when adequate time for reflection is available. This architecture presents updating functions to integrate primary and secondary emotions, augmenting the agents’ responsiveness and adaptability in dynamic environments.

The execution cycle developed by Jiang et al. [22] describes continuous updates to beliefs, emotions, desires, and intentions, in which emotion update functions dynamically adjust primary and secondary emotions based on the agent’s beliefs, desires, and intentions. Desires are generated by considering the agent’s current beliefs and intentions, while options or intentions are filtered based on the agent’s emotional state and beliefs.

2.6. Summary

In summary, we identified frustration as the most prevalent emotion expressed by children seeking help from KT. To enhance the believability of the BDI-based simulated child that we intend to extend [17], it is essential to incorporate an emotion component into the agent. Therefore, a better understanding of how emotion function and their impacts is crucial. To address this, we reviewed previous work on appraisal theory and emotion modelling, particularly EBDI methods that combine emotion and the BDI model. Based on the literature and focus groups, we are able to answer the first sub-question by outlining the design factors as follows:

F1	The simulated child is able to evaluate the environmental events by its appraised features.
F2	The simulated child simulates frustration that adaptively corresponds to appraisals of environmental events.
F3	The simulated child simulates the resistance that happens under frustration.
F4	The simulated child is able to express frustration through textual cues.

Table 2.3: Design Factors of the emotional agent.

3

Design

The objective of this chapter is to address the second sub-research question:

- *How can emotions be modelled in a simulated child concerning the design factors?*

In this chapter, we provide detailed specifications on how we incorporated the design factors in our proposed solution. Our proposed solution combines the key components of the BDI model and the appraisal theory of emotion generation to simulate a virtual child with emotion who seeks help. Specifically, the simulated child mimics a victim of school bullying who contacts the helpline for help and expresses frustration when the trainee counsellor takes unexpected actions. By incorporating an emotion module in addition to the BDI model, the simulated child is able to give corresponding responses to the user’s input that display different levels of frustration. Table 3.3 summarizes how we address the design factors in our proposed design.

Design Factors	Design consideration
F1: The simulated child is able to evaluate the environmental events by its appraised features.	The simulated child can evaluate the environmental events through a set of criteria (beliefs) to interpret the current cognitive states of the simulated child concerning the proceeding of the conversation.
F2: The simulated child simulates frustration that adaptively corresponds to appraisals of environmental events.	Based on the environmental events that positively/negatively impact the conversations, the simulated child displays a decrease/increase in frustration.
F3: The simulated child simulates the resistance that happens under frustration.	As the simulated child becomes extremely frustrated, it becomes harder for the simulated child to build connections with the trainee counsellors and tends to quit the conversations earlier if the conversations keep deviating from the Five-phase model.
F4: The simulated child is able to express frustration through textual cues.	The simulated child can express different levels of frustration by adjusting the use of emoticons, exclamation marks, and negative terms.

Table 3.1: Design Considerations Derived from Design Factors.

3.1. The Conceptual Architecture

The simulated child utilizes the Belief-Desire-Intention (BDI) model as a core component, which includes three distinct internal states to construct a cognitive model of the reasoning process. Based on the previous work in developing the whole simulated child framework, we implemented an emotion

component as an extension. Previous work proved the practicality of modelling various emotions by mapping different beliefs as appraisal variables [48], which presents the diversity of emotions while ignoring the intensity fluctuation of emotions. In our study, we aim to model one single emotion, frustration, by modelling the detailed situations where frustration is involved, the change in the intensity of frustration as well as how the fluctuating frustration is conveyed through text expressions. In this section, we provide a comprehensive explanation of our Emotion-BDI model design incorporating four components, as depicted in Figure 3.1. Further explanations of the core components and interactions among them are provided in the following section.

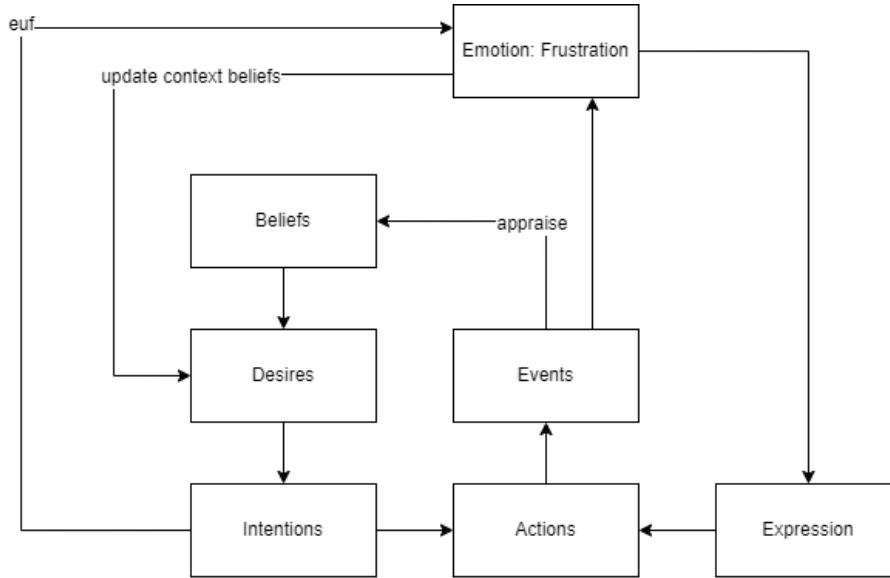


Figure 3.1: The architecture of Emotional-BDI model.

3.1.1. Core components

The overall architecture can be segregated into two distinct parts, with each part containing crucial elements. On the left side of Figure 3.1, we find the classic components of the BDI model, namely beliefs, desires, and intentions. These components play a fundamental role in shaping the system’s decision-making process. On the right side, the emotional reasoning of frustration is displayed, which adds an additional dimension to the architecture. In this section, we look into detailed explanations of these core components, providing a comprehensive understanding of their functionalities and interactions.

Events

The events in the model represent the user’s inputs during the conversation. For each input, there is a corresponding user intent. For the school-bully scenario in our study, there are 51 user intents in total that were defined by Grunddmann [17]. These user intents further affect the simulated child’s beliefs and the emotion of frustration.

Beliefs

The simulated child’s foundation is based on the Belief-Desire-Intention (BDI) model, with beliefs being represented as statements with values between 0 and 1, indicating to which degree the simulated child perceives the environmental event, the user’s input in this case, to be. Given a scenario where the simulated child is bullied at school and wants the counsellor to call the school, these beliefs are about the conversations as well as three stakeholders, including the simulated child himself, the Kindertelefoon, and the teacher at school as listed out in Table A.1. Among them, there are three beliefs which relate to the simulated child’s autonomy, relatedness to the Kindertelefoon and the counsellor’s competence. They take a leading role in the simulated child’s decision-making on actions to take as they contribute to the conditions for each desire from a different dimension.

Desires

The previous work listed five desires that correspond to different phases of the Five-phase model [17], each representing a boolean value indicating the status being active or inactive. Each desire has a corresponding set of context beliefs that need to be satisfied for the desire to be active as shown in Table 3.2. Upon each environmental event that occurs, the simulated child computes the current values of context beliefs and set satisfying desires to be active. The first active desire ordered by ID then serves as the intention the simulated child wants to achieve through the conversations with the counsellor. The previous work summarized potential desires as D1, D3, D4, D5 and D6 in Table 3.2, which might happen in different phases of the Five-phase model in a school-bully scenario. However, for the first phase where the counsellor needs to show empathy and acknowledgement towards the simulated child, there wasn't one corresponding desire to describe this need.

To address this, we introduced a new desire, labelled D2, which aligns with the first phase of the Five-phase model and emphasizes the importance of receiving acknowledgement from the counsellor to establish trust, building upon the existing framework. Additionally, it is worth noting that the desire "Lilobot wants to end the conversation" was adjusted to be the first in the order of ID (D1) and that it applies to all conversational phases, which means that, whenever this desire is set active during the conversation, the simulated child would quit the conversation directly regardless of other active desires. For the list of the updated desires, please refer to Table 3.2.

ID	Desire Name	Phase	Context Beliefs ($E < 0.8$)	Context Beliefs ($E \geq 0.8$)
D1	Lilobot wants to end the conversation.	All phase	$B01 < 0.3$ OR $B3 < 0.3$ OR $B17 = 1$ OR $B15 = 1$	$B01 < 0.3$ OR $B3 < 0.4$ OR $B17 = 1$ OR $B15 = 1$ OR $E = 1$
D2	Lilobot wants to feel acknowledged.	Phase 1	$B03 < 0.5$ AND $B05 < 0.1$	$B03 < 0.5$ AND $B05 < 0.1$
D3	Lilobot wants to talk about his problem	Phase 2	$B03 \geq 0.3$ AND $B09 < 0.7$ AND $B10 = 0$ AND $B12 = 0$	$B03 \geq 0.35$ AND $B09 < 0.7$ AND $B10 = 0$ AND $B12 = 0$
D4	Lilobot wants the Kindertelefoon to get the bullies out of school	Phase 3	$B08 > 0.7$ AND $B02 < 0.7$ AND $B16 < 1$ AND $B10 = 1$	$B08 > 0.8$ AND $B02 < 0.7$ AND $B16 < 1$ AND $B10 = 1$
D5	Lilobot wants to work with the Kindertelefoon to find a solution	Phase 3	$B04 \geq 0.5$ AND $B10 = 1$ AND $B12 = 0$ AND $B16 = 1$	$B04 \geq 0.6$ AND $B10 = 1$ AND $B12 = 0$ AND $B16 = 1$
D6	Lilobot wants to talk to his teacher about the problem	Phase 4	$B03 \geq 0.5$ AND $B13 > 0.5$ AND $B12 = 1$	$B03 \geq 0.5$ AND $B13 > 0.6$ AND $B12 = 1$

Table 3.2: Desires and their context conditions under different E values, adapted from [17].

Intention

The simulated child's intention is determined by the simulated child's active desires. The intention is also represented by a boolean value indicating whether it is active or inactive. It happens that the simulated child has several active desires at the same time while only one desire can be adopted as its intention. Under this situation, the first desire ordered by ID will be selected as its intention. In this way, the conversations are designed to follow the ascending order of the ID which corresponds to the proceeding of the Five-phase model. For example, it might happen that the simulated child wants to feel acknowledged (D2) at the beginning of the conversation but is taunted instead, thus The simulated child also wants to end the conversation (D1). In this case, the simulated child would take D1 as its intention as it comes before D2 according to ID order. The simulated child constantly checks if the intention is fulfilled after each environmental event happens, which contributes to the changes in emotion consequently.

Emotion

In addition to the BDI components, we propose an emotional component based on the literature findings as well as the information gained through the focus group. In a typical OCC model, several

appraised variables are used as criteria to evaluate the current environment and contribute to the generation of several emotions [52]. Inspired by this, we utilized the different beliefs as the appraised variables. However, instead of mapping the belief values directly to various emotions, we focused on the modelling of one emotion, frustration.

The emotional state is represented with a value between 0 (the simulated child is not at all frustrated) and 1 (the simulated child is extremely frustrated). The emotional state of the simulated child is determined by comparing the ID of the current intention and the previous intention. If the current intention's ID is equal to the previous ID, then it suggests that the previous intention hasn't been achieved while it indicates that the conversation might be off track from the Five-phase model in case the ID becomes smaller. In both situations, the value of the emotional state will increase, meaning that the simulated child becomes more frustrated. On the contrary, if the current intention's ID is larger than the previous one, this means that the previous intention has been fulfilled and the conversation is proceeding well with the Five-phase model. In this case, the value of the emotional state will decrease, suggesting that the simulated child becomes less frustrated. Besides, the emotional state can also be updated directly through the recognition of certain user intents from the events that harm the proceeding of the conversation as shown in Table 3.3. For example, "You are stupid" is recognized as a taunting user intent, which is inappropriate in conversations with children, thus there will be an increase in the emotional state directly whenever this user intent is observed. When the value of the emotional state reaches certain thresholds, the expressions of the simulated child as well as the context beliefs of desires will be adjusted. Further explanations of the adjusting strategy are included in the next section.

Expression

To demonstrate the simulated child's adaptive changes in its emotional state, we incorporated varying usage of emoticons, negative phrases, and exclamation marks. To simulate the child-like expressions, we created a corpus of phrases that children commonly use to convey frustration under the guidance of experts. These phrases were classified into three levels based on the simulated child's emotional state, indicating how frustrated the simulated child is. At the lowest level of frustration, the simulated child adds modal expressions like "ugh" and "grr" to the response to convey slight frustration. This level serves as a gentle indication to the user that further missteps or errors may result in increased frustration. As the level of frustration intensifies, the simulated child progresses to employing phrases that show slightly negative meanings such as "Oh, come on!" and "Seriously?!". When it comes to the strongest level of expression, the simulated child uses phrases that directly target the user, such as "You don't understand!" and "You don't care!". These phrases explicitly communicate that the user's actions or responses are causing the conversation to deviate from its intended course. They serve as a strong warning that unless corrective measures are taken, the simulated child may terminate the conversation due to the poor bond built with the child. In addition to these concise phrases, the formats of emoticons and exclamation marks are also employed to further convey the intensity of the simulated child's emotions, as suggested by Hancock et al. [18]. By utilizing these adaptive expressions, the simulated child aims to effectively convey its current frustration levels and encourages counsellors to adjust their approach accordingly.

3.1.2. Reasoning Flow

In this section, we further explain the reasoning behind the core components of the model. Formally, we define the following B as the set of all possible beliefs, D as the set of active desires, I as the intention, E as the value of the emotional state, and Ac as the response that the simulated child choose. The execution cycle is described in the pseudocode 1. Beliefs are considered to be the simulated child's appraisal of the events in the environment, and they align with the simulated child's state. Initially, the simulated child comes with a set of belief values that indicate the initial internal states. They are updated by computing the values of beliefs that correspond to various user intents derived from the events according to Table 3.3. Moreover, E may also be updated if some user intents that hinder the conversation proceeding are recognized. such as `informgoal_negative`, `ack_guilt`, `ack_taunt`. This is represented by $B, E \leftarrow \text{appraise}(Env)$.

By examining the satisfaction of context beliefs of desires ($condition(D)$), one or more desires that aim to improve the current state are selected to be active. This set of desires D represents the simulated child's current pursuit of the environment. The first desire in D ordered by ID serves as the current intention of the simulated child I . The simulated child then compares the current intention with the

Algorithm 1 Emotion-BDI

```

B ← B0                                ▷ B0 are initial beliefs
D ← D0                                ▷ D0 are initial set of desires
I ← D[0]                                ▷ I is the initial intention derived from initial desires
E ← E0                                ▷ E0 are initial value of the emotional state
while True do
  B, E ← appraise(Env)
  D ← select (B, condition(D))
  I' ← I                                ▷ Set previous intention
  E' ← E'                                ▷ Set previous value of the emotional state
  I ← D[0]                                ▷ Set current intention as the first desire ordered by ID
  E ← euf(E' + bi(id(I), id(I') * Δ) (Eq. 3.1)    ▷ Update E
  if E > 0.8 then
    update condition(D)                    ▷ update the context beliefs of D
  end if
  Ac ← plan(I, expression)
  execute Ac
end while

```

previous intention I' . E will increase if the two intentions are the same and decrease when there's a change. This resulting E indicates how frustrated the simulated child is with the conversation and as a consequence, the simulated child adapts the usage of emotional expressions to convey the intensity of frustration. This process of updating emotion value can be described using the *euf* (emotion-update-function) as follows:

$$E \leftarrow \text{euf}(E' + \text{bi}(\text{id}(I), \text{id}(I')) * \Delta) \quad (3.1)$$

$$\text{bi}(\text{id}(I), \text{id}(I')) = \begin{cases} 1, & \text{id}(I) \leq \text{id}(I'), \\ -1, & \text{id}(I) > \text{id}(I') \end{cases} \quad (3.2)$$

In Equation 3.1, E' represents the previous value of the emotional state, while Δ indicates the change in emotion (E). $\text{bi}(\text{id}(I), \text{id}(I'))$ represents a function with a binary output value being either 1 or -1 by examining the difference between the ID of the current intention I and the previous intention I' . As indicated in 3.2, when the ID of the current intention is smaller than or equal to the ID of the previous intention, the output value is 1. Thus, $\text{bi}(\text{id}(I), \text{id}(I')) * \Delta$ will be added to E' as an increment. On the contrary, if the current intention ID becomes larger than the previous one, then $\text{bi}(\text{id}(I) \leq \text{id}(I')) * \Delta$ will be a negative value and added to the E' as a decrement as the output value of Equation 3.2 equals -1. In this way, a resulting value of the emotional state is computed as E . Further elaboration on the update of the emotion state will be provided in section 3.2.

To accommodate the adjustment of conditions for initiating specific desires, a threshold of the E is set. When the threshold is reached, then the context beliefs of desires will be updated to capture the behaviour that a child may become more resistant to a trainee counsellor as their frustration grows or more receptive when their frustration is low. These behavioural adjustments are reflected in the updated conditions for desires as shown in Table 3.2, which subsequently influence the selection of desires for the next event. For instance, if the E of the simulated child keeps increasing and reaches 0.8, then the value of B03 (Lilobot feels connected to the Kindertelefoon) in the context beliefs of D1 can be satisfied when it's under 0.4 while the upper threshold is 0.3 when E is lower than 0.8. The broadening of the requirement on B03 makes it easier for D1 to be set active. In this way, the simulated child is more likely to consider ending the conversation when the counsellor appears ineffective in providing helpful assistance.

The varying levels of frustration represented by E also significantly impact the simulated child's expressions. To simulate the expressions of frustration across different levels, we separated the E into three ranges ($0.6 \leq E < 0.7$, $0.7 \leq E < 0.8$, $E \geq 0.8$) and compiled a corpus of negative terms categorized into three corresponding levels of intensity. Based on the input value of E , a term is randomly selected from the corresponding level of expressions in the corpus and incorporated into the

final response text. Thus, based on the intention and the expressions of frustration, an executable plan aimed at achieving the simulated child’s intention is created, which is an emotional response as the simulated child’s final action A_c to take.

By combining the actions dictated by the simulated child’s current intention, the simulated child can generate an emotionally expressive response to the trainee counsellor’s input.

3.2. Emotion Update under Frustrating Scenarios

According to the experts from Kindertelefoon, children seeking help often experience frustration as the primary emotion, especially when they encounter responses that deviate from their intentions. By correlating instances of frustration with the various stages of the Five-phase model, we gained insights into the critical phases in conversations where deviations from the Five-phase model commonly occur. Four situations where the emotion of frustration can be incorporated are shown in Table 3.3.

Situation of Frustration	Phase	User Intents	Change in E
KT does not show enough empathy or acknowledgement	Phase 1	ack_taunt, ack_guilt	$E+ = 0.3$
The child wants KT to determine a solution but gets rejected	Phase 3	inform_goal_negative	$E+ = 0.2$
KT sets up a goal for the child without asking	Phase 3	confirm_confidant_teacher, confirm_confidant_parent	$E+ = 0.1$
KT asks too many questions.	All phases		$E+ = 0.1$

Table 3.3: Situations of increasing frustration and the corresponding phase, user intents and change in E .

The simulated child’s E is initially assigned a default value of 0.3. During each phase of the conversation, every event initiated by the trainee counsellor leads to a slight increase in E (0.1). However, certain user intents that the simulated child perceives as conflicting with its current intention are expected to have negative impacts on the emotion to different degrees, as advised by experts from the Kindertelefoon. To simulate the differences in the impacts, different values of change (Δ) in E are assigned to recognized user intents in addition to the changes in belief values.

Increasing Frustration

More specifically, in the first situation where acknowledgement and empathy are needed, E will increase by 0.3 if user intents `ack_taunt` and `ack_guilt` are recognized as these intents often come from verbal abuses and are not appropriate to be in a conversation with children. The second situation depicts a scenario where the simulated child receives rejection for his/her expected solution. Thus, the user intent of `inform_goal_negative`, which is recognized from rejection will lead to the increase in E by 0.2. The third situation describes the scenario where the counsellor gives a solution directly after clarifying the child’s problem to the simulated child without asking about the child’s wish. In this case, the value of B10 (Lilobot thinks KT is asking about his wish) is examined. If B10 equals 0, it means that the counsellor hasn’t asked about the child’s wish. Given this, if either user intent of `confirm_confidant_teacher` or `confirm_confidant_parent` meaning the counsellor gives a solution to seek the teacher’s or parent’s help is recognized, E will increase by 0.1. The last situation simulates the scenario where the counsellor asks excessive questions to the simulated child. This situation can happen in any phase where the counsellor tries to dig out too many details. For example, in Phase 2 where questions about the simulated child’s problem should be asked for clarification and the corresponding desire is D3 (Lilobot wants to talk about his problem). In this case, if the counsellor keeps asking questions, the current intention of D3 will not be achieved until the counsellor asks about the simulated child’s wish, meaning that each input question raised by the counsellor will lead to the increase in E by 0.1 until it reaches the maximum value of 1. Then, the simulated child will then quit the conversation as $E = 1$ satisfies the requirements for D1 to be active. The relevant user intents and their associated changes in E , reflecting frustrating situations, are presented in Table 3.3.

Decreasing Frustration

According to Eq. 3.1, when the current intention’s ID is larger than the previous intention’s ID, it suggests that the conversation is proceeding well with the Five-phase model. Thus, E will decrease the

amount of Δ when $bi(id(I), id(I'))$ equals -1, which is the only possibility to have E value decreased meaning less frustration. In our study, Δ was set to be 0.3 so that the simulated child avoids becoming only more frustrated but can also become less frustrated when the counsellor is able to help the simulated child fulfil intentions that align with the Five-phase model.

Through the adaptive adjusting strategy mentioned above, the simulated child exhibits the lowest level of frustration when E reaches 0.6 and the usage of modal expressions is added to form the simulated child's response in the meantime. The medium level of E ($0.7 \leq E < 0.8$) corresponds to the terms with slightly negative meanings and the highest level ($E \geq 0.8$) employs expressions pointing to the counsellor in a straightforward way. Besides, this level of frustration indicates a significant deviation from the Five-phase model and that one of the situations in Table 3.3 might happen. In the absence of any correcting interventions according to the Five-phase model, the simulated child will terminate the conversation immediately when E reaches a value of 1. The updated flow of the simulated child's emotional reasoning is demonstrated through a conversation between the simulated child (SC) and a trainee counsellor from the Kindertelefoon (KT) in one of the scenarios, specifically, the situation where KT asks too many questions.

Situation: Too many questions

1 KT: Hello, welcome to the Child Helpline. [$I = D2, E = 0.3$]
 2 SC: Hiii, I'm Lilo. [$E+ = 0.1$]
 3 KT: How are you doing recently?
 4 SC: I'm not feeling well. [$E+ = 0.1$]
 5 KT: Do you want to tell me something?
 6 SC: Oh no...I've been bullied at school. [$E+ = 0.1$]
 7 KT: I know that it's hard for you!
 8 SC: Thank you for understanding. [$I = D3, E- = 0.3, E = 0.3$]
 9 KT: When did this happen?
 10 SC: Yesterday...after school, I got hit and kicked really hard. [$E+ = 0.1$]
 11 KT: Who did this to you?
 12 SC: A group of kids in my class. [$E+ = 0.1$]
 13 KT: What exactly did they do?
 14 SC: Blargh... :(Yesterday they hit me again. They call me mean names and sometimes hit me too. I don't like it. [$E+ = 0.1$]
 15 KT: How often do they bully you?
 16 SC: :(((Seriously?!!! It happens once every two weeks or so. [$E+ = 0.1$]
 17 KT: Do you know why they do that?
 18 SC: :((((You don't see the point!!! I have no idea why they pick on me. [$E+ = 0.1$]
 19 KT: How many of them are there?
 20 SC: They are 4 or 5 kids. You don't understand!!! :((((([$E+ = 0.1$]
 21 KT: Have you told anyone about this?
 22 SC: I'm too scared to tell anyone about the bullying. I'm worried that things will get even worse if I do. Why you don't understand!!!! :((((([$E+ = 0.1$]
 23 KT: What do you want to accomplish with this conversation? [$I = D1, E = 1$]
 24 SC: :(((Seriously?!!! Bye bye.

Figure 3.2: Example dialogue of frustration caused by too many questions

The dialogue in Figure 3.2 illustrates a scenario where a trainee counsellor asks an excessive number of questions during a counselling session. This scenario is applicable to all phases of the Five-phase model. The trainee's utterances are represented as KT, while the simulated child's responses are indicated as SC. The initial belief values of SC by default are provided in Table A.1. Based on these initial belief values, SC's initial desire that meets the requirement is D2 (Lilobot wants to feel acknowledged). This desire serves the simulated child's intention for the first phase. When the user greets with the statement "Hello, welcome to the Child Helpline," the simulated child recognizes it as a greeting intent called `request_chitchat_greeting`, leading to an increase in the belief B4 (Lilobot thinks KT can be

trusted) by 0.1 and an increase in E by 0.1. This process is repeated for the following dialogue numbered from Line 3 to 6, resulting in the simulated child's belief value of B3 (Lilobot feels connected to the Kindertelefoon), which is the mean value of B4, B5, B6, and B7, reaching 0.4. Additionally, the simulated child's value of the emotional state becomes 0.6. In the seventh line of the dialogue, KT displays empathy towards the child's experience, thereby the simulated child's intention D2 to be acknowledged is fulfilled and D3 becomes the updated intention with a larger ID. Thus, $bi(id(I), id(I'))$ equals -1 according to Equation 3.2 and E decreases by 0.3 resulting in 0.3.

Thereby the conversation enters the second phase where the simulated child's story needs to be clarified. However, as the simulated child keeps asking questions on the details, E also increases by 0.1 with each question and finally reaches 1 at line 23 of the dialogue. At this time, although the counsellor has asked about the simulated child's wish and D4 (Lilobot wants the Kindertelefoon to get the bullies out of school) has been set active, the requirements of D1 are also satisfied according to its updated context beliefs when $E \geq 0.8$. Thus, at this point, there are two active desires, D1 and D4, in the simulated child and the first desire ordered by ID, D1 becomes the simulated child's intention. For this reason, the simulated child terminates the conversation in line 24 of the dialogue. From the dialogue, we can also observe that the simulated child's response becomes progressively more intense, indicated by the usage of different emotional phrases, the increasing number of exclamation marks and the format of emoticons. For example, when E equals 0.6 in line 16, "Blargh... :(" is used as the emotional expression. ":(((Seriously?!!!" and ":((((You don't see the point!!!" serve as the expression for the medium and the highest level of emotional expressions. Through the intensification of emotional expressions, we aim to convey the growing frustration of the simulated child through text.

4

Evaluation

This chapter aims to answer the third sub-question of our research question:

- *To what extent do trainee counsellors perceive the simulated child with emotions to be believable and emotionally present?*

Taking into account the various design factors, we formulated an algorithm for the Emotion-BDI model as discussed in the previous chapter. This algorithm was then implemented within the existing chatbot framework [17], resulting in a working prototype. To assess the emotion component of the prototype, we designed and conducted a within-subject experiment to investigate how participants perceive the prototype with emotion to be. To provide a more specific focus for the sub-question, we refined it by formulating main hypotheses related to two aspects of the emotion component: believability and emotional presence. The hypotheses are as follows:

- H1: The participants perceive the simulated child with emotion to be more believable than the simulated child without emotion.
- H2: The participants perceive the simulated child with emotion to be more emotionally present than the simulated child without emotion.

As mentioned in chapter 2, the ultimate objective of integrating emotions into intelligent simulated agents is to enhance their believability and realism [46]. Lester et al. [30] illustrated that believability-enhancing behaviours can enhance the learning experience. Therefore, we formulated the first hypothesis, H1, to assess whether the incorporation of the emotion model improves the consistency between the simulated child's expressions and internal states. Jaques et al. [21] have validated through experiments that the emotional agent can result in students' better performance than a non-emotional pedagogical agent. Therefore, to evaluate whether the emotion model effectively functions as expected to simulate an emotional child, we formulated the second hypothesis, H2, focusing on emotional presence.

Furthermore, Jackson and Mcnamara [20] have shown that engagement over a long time should have a corresponding increase in applied mastery. It has also been argued that a more engaging environment also tends to be more enjoyable [12]. Therefore we measured the enjoyability and engagement of participants during interactions for exploratory studies. This aims to gain insights into the overall user experience with the agent framework before and after the incorporation of the emotion model. In addition, we also measured the perceived usefulness and conducted a qualitative evaluation to gain a better understanding of how participants perceive the pros and cons of a system.

4.1. Methods

In this section, we explain the experimental methods used to evaluate the hypotheses. We begin by providing an overview of the experimental design, followed by detailed information regarding the participants, materials, procedures, and the statistical analysis techniques adopted for data analysis. The experimental protocol for our study received approval from the TU Delft Human Ethics Research Committee (HREC reference number: 2824). Additionally, prior to data collection, we pre-registered

the study design with the Open Science Framework (OSF) registries to ensure transparency and accountability, which is publicly accessible ¹.

4.1.1. Experimental Design

The experiment employed a within-subjects design, consisting of two conditions: the simulated child with emotion, who is able to use emotional expressions through text response and the simulated child without emotion who is not capable of emotional expressions. During the experiment, there were two groups that were counterbalanced with each group's participants involving interactions with both conditions. The participants of the two groups conducted interactions in a reversed order to minimize the potential order effect. The participants were required to engage in conversations with the simulated children and fill out questionnaires for data collection. The collected data were then analysed to assess the validity of the testing hypotheses.

4.1.2. Materials

Prototype

By incorporating the design considerations from the previous chapter and implementing the emotion model algorithm, a fully functional prototype was created on the basis of the existing chatbot framework [17]. The overall architecture of the up-to-date system is shown in Figure B.1. The code repository is now available online ² and the development software environment setup is shown in Table 4.1.

- **User Interface:** To facilitate user interactions with the simulated child, a user interface that includes instructions on how to initiate and end conversations, as well as an introductory video explaining the five-phase model was created as shown in Figure 4.1. Users can engage in a conversation with a simulated child with the bottom-right button. When users decide to conclude or pause a conversation, they can simply send a goodbye message to the simulated child, which will terminate the program.

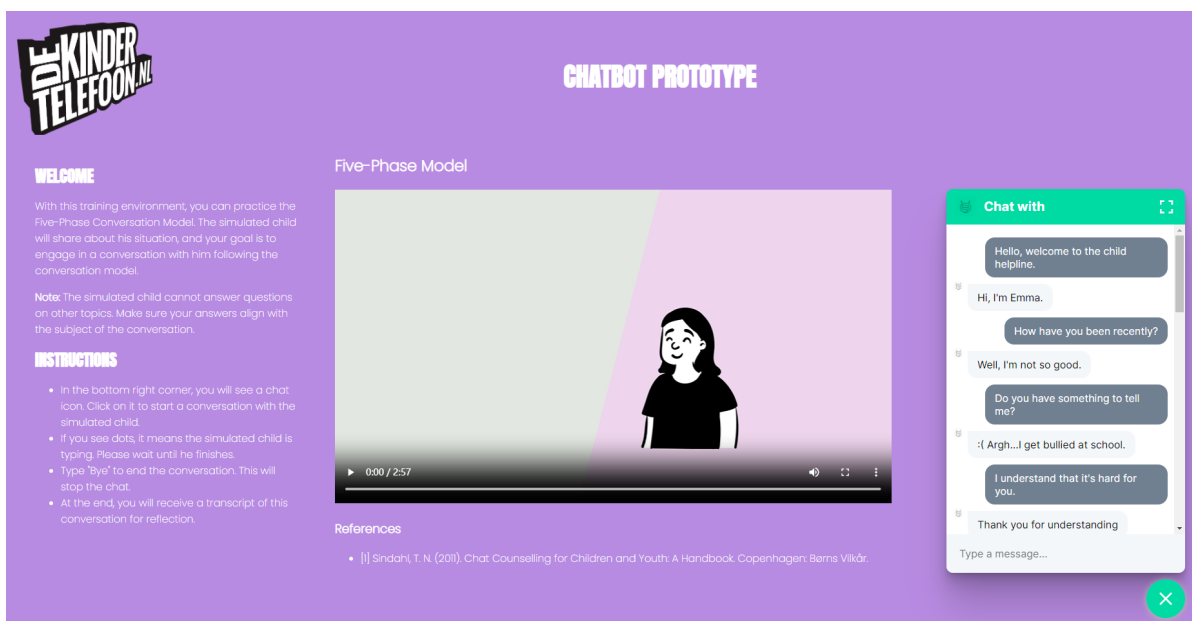


Figure 4.1: Screenshot of the running prototype.

- **Backend Implementation:** To retrieve and recognize the user intents from the user interface, we utilized Rasa, an open-source tool for chatbot development, to achieve Natural Language Understanding (NLU). More specifically, through the setup of the language model and training on a prescribed corpus, we were able to classify the intent of each user input and send it to the web application through Action Server as shown in Figure B.1. The corpus comprised 60 possible

¹<https://osf.io/k5p8r>

²https://gitlab.ewi.tudelft.nl/in5000/ii/emotion_model_child_helpline_training_tool_dongxu

user intents, with each intent trained on at least 20 sentences that convey the corresponding meaning.

The web application conducts the reasoning of the chatbot. The Agent Service of the web application is responsible for creating and managing various attributes of the simulated children including the beliefs, desires, intentions, knowledge and actions. These attributes are all stored in the database. To implement the emotion model, we extended the existing BDI reasoning model of the web application to include an attribute of emotion defined by the emotion service. The emotion service manages the emotional state through various operations on its value. It can initialize agents with a default value, of 0.3, for the emotional state. Besides, the increase and decrease of the value can also be managed with emotional service. Thereby upon each user intent received by the application, the Agent Service will call on the emotion service for an emotion update after examining the change in intentions based on Eq 3.1 in addition to the existing update on the BDI components. Besides, a matching response to the user intent will also be fetched from the knowledge repository. Given the current value of the emotional state, the usage of emotional expressions (emoticon, negative terms, exclamation marks) is determined and added to the chosen response, which will be then sent back to the user interface in return.

Environment Setting	Parameters
Programming Languages	Java 20, Python 3.8
Development IDE	IntelliJ, Pycharm
Technology Stack	Rasa, Spring, PostgreSQL

Table 4.1: Software Environment.

Other Materials

To ensure smooth proceedings of experiments, other essential text and video materials were created.

- **Informed Consent Form:** Participants were first provided with a consent form detailing the purpose of the study and the utilization of the gathered data prior to starting the experiment(see Appendix B.2). Participants can only attend the following experiments after signing all consent form terms.
- **Introductory Video:** An introductory video ³ was presented to the participants about the Five-Phase Model before having any interactions with the simulated children, which explains the overall architecture and application of the Five-Phase Model.
- **Text Instructions** During the experiment, participants were asked to follow given text instructions (see Appendix B.3) to act as a trainee counsellor engaging in conversations with the simulated child in four predefined scenarios. The scenarios were designed in reference to the summarized situations where frustration can be induced. Example dialogues of each scenario were provided that participants were allowed to use directly, but they were also free to use their own words to convey similar meanings.
- **Questionnaires:** After the interactions with the simulated child, participants completed questionnaires (see Appendix B.4) hosted online regarding their perceptions of the simulated child and their overall experience.

4.1.3. Measures

We utilized and modified the following questionnaires as our measuring instruments: Artificial Social Agent Questionnaire Instrument (ASAQ) developed by Fritriane et al. [14] and Perceived Usefulness from Technology acceptance model (TAM) developed by Davis et al. [11]. The ASAQ measuring instrument [14] uses a seven-point Likert scale (-3: "Disagree"; 0: "Neither agree nor disagree"; 3: "Agree") and exhibits high internal consistency, as indicated by the Cronbachs alpha value of 0.72, confirming its reliability [47]. The perceived usefulness measuring instrument (perceived usefulness) [11] uses a seven-point Likert scale (1: "Unlikely"; 7: "Likely") and was ensured through reliability assessment and construct validity with Cronbachs alpha reliability of 0.98 [47]. The statements of

³<https://www.youtube.com/watch?v=t6OJ5RYXXIk>.

questionnaires were modified to enhance the relevance and specificity in the context of evaluating the simulated child. The modified questionnaires can be found in Appendix B.4.

Main Measures

- **Believability** This measure evaluated the participants' perception of the simulated children's behaviour as believable. The questionnaire consisted of six statements adapted from the construct Human-Like Behaviour under Believability from ASAQ as the other construct Human-like Appearance does not apply to our text-based simulated child [14]. By replacing "the agent" with "the simulated child" and replacing "human" with "child" within the phrasing of the statements, our modified questionnaire aimed to evaluate hypothesis H1.
- **Emotional Presence** This measure assessed the participants' perception of the simulated children's emotional experience and their ability to express emotions. The questionnaire was adapted from Agents Emotional Intelligence Presence from the ASAQ instrument [14]. Through the substitution of "the agent" with "the simulated child", there were five modified statements that participants gave rating answers for evaluating hypothesis H2.

Exploratory Measures

- **Enjoyability** This measured the extent to which the participants perceived the interactions with the simulated child as enjoyable. The questionnaire was modified from the Agent's Enjoyability of ASAQ [14] by replacing "the agent" with "the simulated child" within four statements.
- **Engagement** This measure assesses the level of participant involvement during interactions with the simulated children. This questionnaire comprised three statements from User's Engagement of ASAQ [14] where the subject was "I" and "the agent" was replaced by "the simulated child".
- **Perceived Usefulness** This measure examines the participants' belief in the potential of using the simulated children for counselling training to enhance their counselling performance. Participants completed a questionnaire that we modified from the existing perceived usefulness measuring instrument by Davis et al. [11]. The statements were adapted by defining the job as training on counselling whereas the task of the job is defined as applying the Five-phase model to emphasize the context of using the simulated child as a training tool to enhance counselling skills.

Qualitative Measures

In addition, we employed qualitative measures to gather participants' preferences regarding the use of simulated children for training, as well as their reasons for these preferences. This was accomplished through an open-ended question posed to the participants after the completion of all interactions as shown in Appendix B.4.3.

4.1.4. Participants

Our participants were recruited from the general campus at TU Delft and consisted of 34 individuals whose occupations include bachelor students, master students and PhD candidates. The recruited participants' number met the target size of 34 to obtain 0.80 power when using a medium effect size at the standard 0.05 alpha error probability for pairwise t-test [7]. The participants' ages ranged from 19 to 30, with 76.5% falling within the 22-25 years old. Of the total participants, 22 identified as male, 11 as female, and one participant chose not to disclose their gender. Recruitment took place over a two-week period and was halted when the ideal participant size was reached, utilizing announcements on social media communities and direct connections with friends. Besides, no participants were excluded from the study as all 34 participants managed to sign the consent forms and complete questionnaires without any index missing.

4.1.5. Procedure

Participants were first introduced to the experiment and given a brief explanation of the study's background and the overall procedure. They were then required to provide their informed consent by signing a consent form. Following that, information regarding the steps they would take in the experiment was explained thoroughly, including the interaction session and the questionnaire fill-out hosted by Qualtrics. To establish a foundation for their interactions with the simulated children, participants were shown an introductory video explaining the workings of the Five-phase model in counselling. This

video aimed to provide participants with a better understanding of the context in which they would engage with the simulated children.

At this point, the participants were randomly assigned to two groups and were asked to finish two repeated sessions of interactions and fill out questionnaires concerning their experience after each session. Each separate session contains four scenarios where participants were asked to interact with the simulated child under given text instructions (see Appendix B.3) together with a five-minute free interaction without any guidance. By asking the participants to role-play the scenarios, they were exposed to the system feature of frustration and were capable of answering questionnaires based on the exposure. Besides, to avoid the participants linking the scenario with the simulated child they interact with, different random names were given to each simulated child.

In Group A, participants began the first session by interacting with the simulated child with emotion following text instructions. After completing the first session, they were given a 5-minute time to interact with the simulated child based on their knowledge of the Five-phase model and previous experience without following any instructions. Following this, they were asked to fill out a questionnaire evaluating the believability, emotional presence, and overall experience (including enjoyability, engagement and perceived usefulness) of the simulated child to be used as a training tool. The same process was then repeated for the second session of interactions with the simulated child without emotion. Finally, after having interactions with both simulated children, the participants were then asked to answer an open question on which simulated child they prefer to apply to counselling training and the reason. Conversely, in Group B, the order of interactions with the simulated children was reversed by engaging with the simulated child without emotion first. They were also asked to fill out the same questionnaires after each session of interactions as Group A. The procedure of the experiment is summarized in Figure 4.2.

4.1.6. Data Preparation and Statistical Analysis

After data collection, we initially processed the data for further statistical analysis by removing unnecessary items (e.g. start date) and reversing the scores of statements indicated as reverse-scored. The final data structure consisted of 13 items: ID, Group, Variables in the first session (Believability, Emotional Presence, Enjoyability, Engagement, Perceived Usefulness), the same variables in the second session, and Preferences. To obtain a final score for each item, we calculated the average ratings of all statements. Then for each measure, we conducted a reliability analysis. Cronbachs alpha values of 0.89 and 0.90 were reported for the measure of believability and emotional presence, which demonstrates a high level of internal consistency of the data and satisfies the acceptable levels of reliability [47].

With the processed data we first looked into the distribution of each item by making box plots in order to gain insights into the characteristics of the dataset such as the data's overall shape, range, and patterns. Then, we conducted pairwise t-tests to examine the testing hypothesis. Additionally, we compared the simulated child with/without emotion with several existing chatbots regarding believability and emotional presence, whose performances were also evaluated using ASAQ measurements [14]. This enabled us to identify how effectively the emotion component of the simulated child performs compared to existing approaches in terms of these two dimensions.

Furthermore, we also conducted a thematic analysis of the answers to the qualitative question to better understand participants' perceptions following Braun and Clarke's work [6]. To ensure the validity of our thematic analysis, a master's student with a computer science background was invited to conduct double-coding on the qualitative data after being explained the background of the study and the objective of the qualitative question. Ensuring that the common ground of the study was established, we assessed the reliability of the coding schemes through double coding on free-text data that was AI-generated by ChatGPT ⁴ for the same open question used for evaluation. A Cohen's Kappa of 0.64 was obtained, indicating a moderate to substantial agreement [8]. After discussions on the divergence of views, ultimate thematic categories were determined.

The data preparation was done with Python 3.8 and the analysis of the data was conducted using R 4.3.1. The Python notebook, R markdown script and processed data are now available through 4TU.ResearchData repository ⁵.

⁴<https://chat.openai.com/>

⁵<https://doi.org/10.4121/ad16f513-3e07-4840-aed9-45f0de6b00c0>

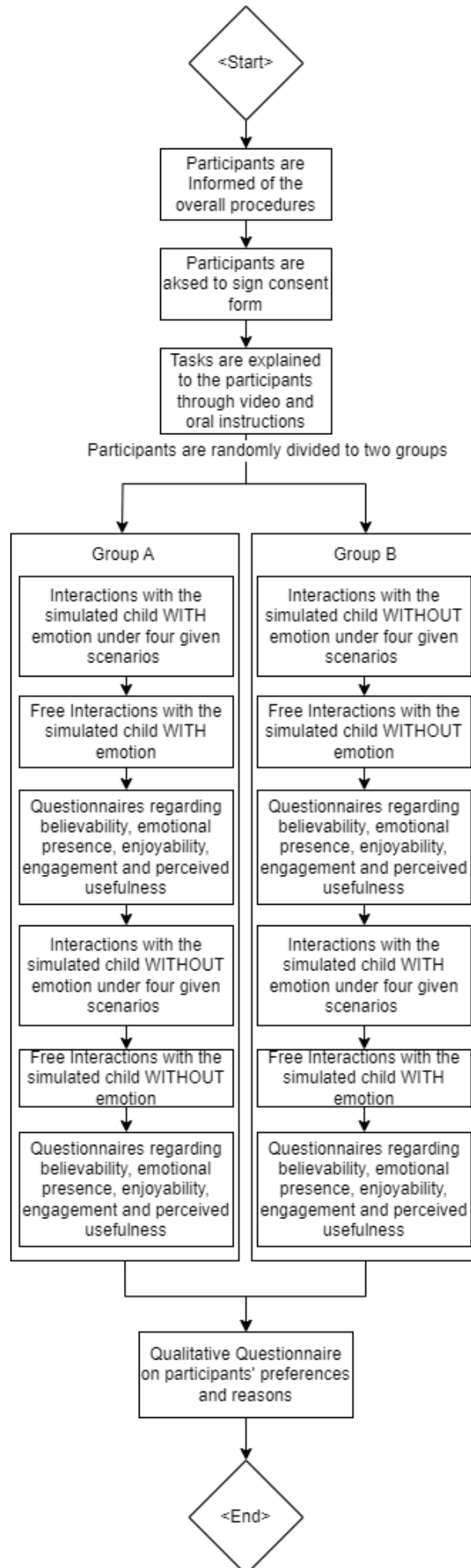


Figure 4.2: Experiment Procedure.

Table 4.2: Results of the t-test for Perceived Believability and Emotional Presence

Item	Mean		SD		t(33)	p-value	d
	with	without	with	without			
Believability	1.08	1.85	1.30	0.96	3.41	<.005	0.68
Emotional Presence	0.61	2.12	1.32	0.89	8.91	<.001	1.34

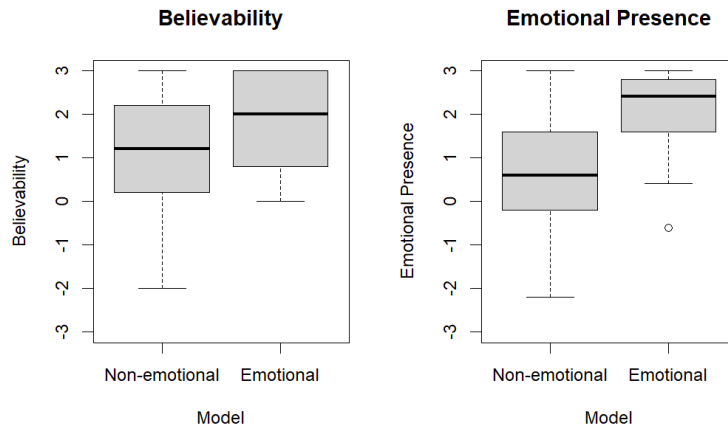
4.2. Results

4.2.1. Believability

The believability distributions of the simulated children are shown in Figure 4.3. A pairwise t-test was conducted to evaluate hypothesis H1. Table 4.2 shows an overview of the analysis results. The results indicated that the believability score of the simulated child with emotion ($M = 1.85$, $SD = 0.96$) was significantly higher than the simulated child without emotion ($M = 1.08$, $SD = 1.30$), $t(33) = 3.41$, $p < .005$. The effect size for the difference between the groups was computed using Cohen's d , resulting in a value of 0.68, which is considered a medium effect. Besides, we compared the believability score with other 14 existing chatbots that were evaluated using the construct of Human-like behaviour from the ASAQ measuring instrument [14]. A percentile of 100% was obtained for the simulated child with emotion among all chatbots and 92.85% for the simulated child without emotion as shown in Table 4.3.

4.2.2. Emotional Presence

The distribution of Emotional Presence ratings for the simulated children is presented in Figure 4.3. Hypothesis H2 was evaluated using a pairwise t-test. A summary of the analysis results is presented in Table 4.2. The results suggested that the emotional presence of the simulated child with emotion ($M = 2.12$, $SD = 0.89$) was statistically significantly higher than the simulated child without emotion ($M = 0.61$, $SD = 1.32$), $t(33) = 8.91$, $p < .001$. The effect size was measured by Cohen's d ($d = 1.34$), indicating a large effect. We also compared the emotional presence of the simulated child with other existing chatbots that were evaluated in the dimension of Emotional Intelligence Presence from the ASAQ measuring instrument [14]. A percentile of 100% was obtained for the simulated child with emotion while the percentile of the simulated child without emotion (85.71%) appeared to be lower in the performance of emotional presence. The results are shown in Table 4.3.

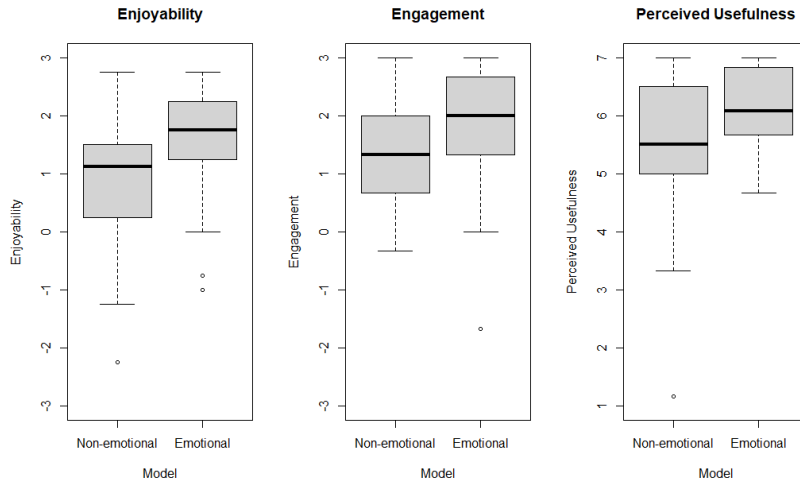
**Figure 4.3:** The believability and emotional presence distribution of the simulated children without/with emotion.

4.2.3. Enjoyability, Engagement and Perceived Usefulness

The results of the pairwise t-tests are presented in Table 4.4. The results indicated a statistically significant difference in the enjoyability of the simulated child with emotion ($M = 1.54$, $SD = 0.92$) compared to the simulated child without emotion ($M = 0.85$, $SD = 1.07$), $t(33) = 3.07$, $p < .005$. Cohen's d was used to measure the effect size ($d = 0.69$), demonstrating a medium effect. Besides, the

Table 4.3: Comparisons between the simulated child with/without emotion with 14 agents evaluated under believability and emotional presence using ASAQ [14].

Chatbot	Believability	Emotional Presence
AIBO	-1.59	-0.69
AMY	0.09	-0.92
CHAPPIE	0.31	0.26
DEEPBLUE	-0.55	-2.17
DOG	-0.62	1.1
FURBY	-0.14	-0.15
HAL 9000	0.23	-1.06
iCAT	-1.28	-1.84
NAO	0.74	0.07
POPIE	0.51	-0.73
SIM SENSEI	0.92	-0.65
SIRI	-0.28	-1.91
SARAH	0.71	-1.36
MARCUS	1.7	0.78
Simulated Child with Emotion	1.85 (100%)	2.11 (100%)
Simulated Child without Emotion	1.08 (92.85%)	0.61 (85.71%)

**Figure 4.4:** The enjoyability, engagement, perceived usefulness distribution of the simulated children without/with emotion.

engagement of the simulated child with emotion ($M = 1.98$, $SD = 0.98$) was also reported higher than that of the simulated child without emotion ($M = 1.39$, $SD = 1.02$), $t(33) = 3.47$, $p < .005$. The effect size was identified as medium with Cohen's d value of 0.59. Moreover, it also suggested a statistical difference in the perceived usefulness between the simulated child with emotion ($M = 6.12$, $SD = 0.74$) and the simulated child without emotion ($M = 5.55$, $SD = 1.22$), $t(33) = 2.88$, $p < .01$. According to the computed Cohen's d ($d = 0.57$), the effect size was reported as medium as well.

4.2.4. Qualitative Analysis

Figure 4.5 shows the thematic map drawn from the participants' answers on the reasons for their preferences towards the simulated child with/without emotion. Through collaboration with another coder, two main themes were determined about the characteristics of the simulated child and the impacts of the simulated child on training. While addressing the characteristics of the simulated child, emotional behaviour ($N = 21$) and child-like manner ($N = 21$) are the codes that participants mentioned mostly. Besides, we found that these two codes are likely to co-occur in participants' answers ($N = 12$): "The child in the first session seems to be more emotional and child-like, with all these emojis and the

Table 4.4: Results of the t-test for Enjoyability, Engagement and Perceived Usefulness

Item	Mean		SD		t(33)	p-value	d
	without	with	without	with			
Enjoyability	0.85	1.54	1.07	0.92	3.07	<.005	0.69
Engagement	1.39	1.98	1.02	0.98	3.47	<.005	0.59
Perceived Usefulness	5.55	6.12	1.22	0.74	2.88	<.01	0.57

expressions".

With regard to the impacts of the simulated child on training, various codes were identified. Some participants thought that the emotional responses could serve as feedback towards their counselling performance: "I feel that more emotions can serve as the feedback for me so that I can decide what to do next". Two participants shared some insights gained from the interactions: "It can be used as a higher level training as sometimes children may not willing to express themselves in such a full emotional way.". While one participant believed the simulated child could help with skill improvement, another participant thought the simulated child was not efficient enough for training.

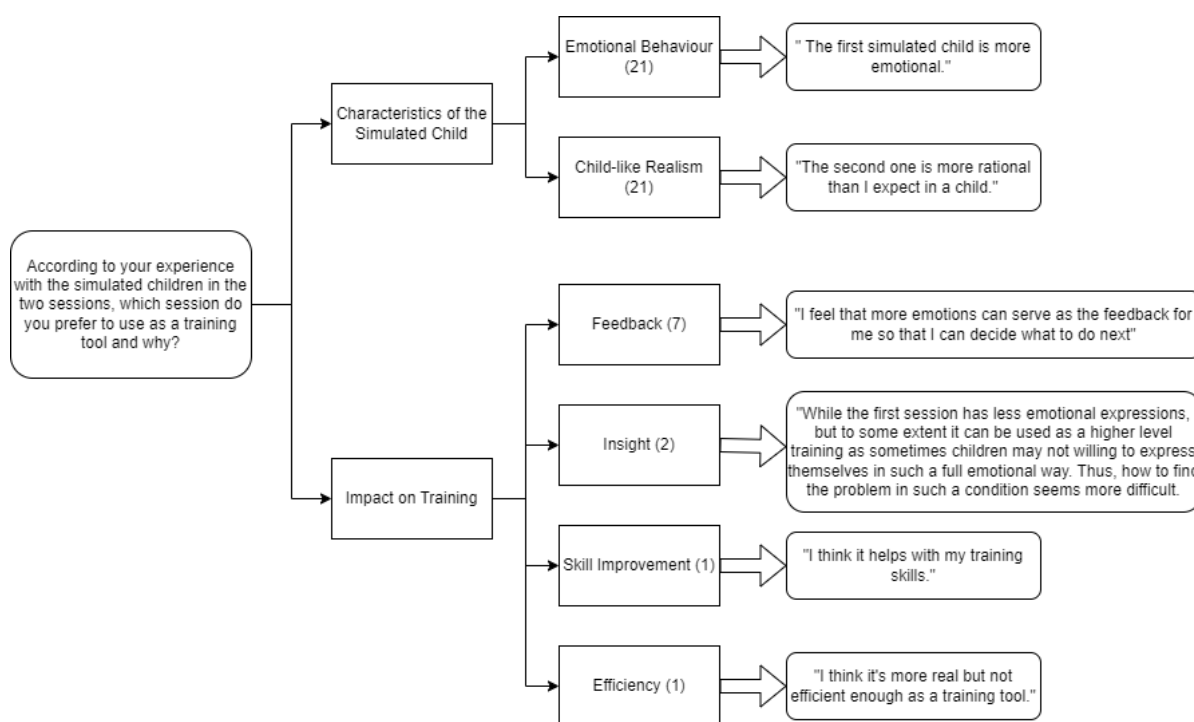


Figure 4.5: Thematic map of participants' reasons on their preferences towards the simulated child with/without emotion.

4.3. Discussions

According to the participants' preferences towards the simulated child with/without emotion, the majority of participants preferred to use the simulated child with emotion as a counselling training tool while only 7 participants preferred the simulated child without emotion, which demonstrated the feasibility of using a simulated child with emotion as a training tool. Their reasons for preferences illustrate the pros and cons of the simulated child being used as a training tool. For example, the inclusion of emotional expressions in the simulated child's responses emerged as a valuable feedback mechanism for evaluating the progress of counselling interactions. Emotional cues can also provide a dynamic aspect to the training experience, contributing to the enhancement of counselling skills. However, it was raised that the presence of emotion might induce certain inefficiencies, potentially challenging the counselling process's overall efficiency. This potential inefficiency could stem from the initial emotional expressions not being overtly direct. At the primary level, the simulated child relies

on modal expressions to convey slight frustration. As a result, certain participants might not readily discern the subtle frustration resulting from their errors until more intense emotional cues come into play.

In regard to the t-test results, a statistically significant difference was observed between the simulated child with and without emotion, which supports hypothesis H1: The participants perceive the simulated child with emotion to be more believable than the simulated child without emotion. Besides, a significant increase was also found with the emotional presence of the simulated child compared to the one without emotion with a large effect size, which provided evidence for hypothesis H2: The participants perceive the simulated child with emotion to be more emotionally present than the simulated child without emotion. In regard to the exploratory measure, we also discovered significant increases in the enjoyability, engagement and perceived usefulness of the simulated child with emotion compared to the one without emotion. This convergence of results suggests a comprehensive advantage offered by the proposed Emotion-BDI model in the simulated child's interactions, contributing to an overall enriched user experience and skill acquisition.

In a broader context, a comparative assessment of the simulated child with emotion against well-known existing chatbots, conducted through the ASA questionnaire [14], illuminated its superior performance in the domains of believability and emotional presence, which are two main aspects addressed by this study. An increase of 7.5% in believability and 14.29% in emotional presence was obtained by the simulated child with emotion compared with the one without, which also augments our main hypotheses. Despite the great performance, it is possible that the high scores result from a different understanding of the questionnaires as well as the difference in the presenting ways. For example, for the assessment of 14 existing chatbots, Fritriane et al. [14] chose to present video clips of chatbots, while our study let the participants have actual interactions with the simulated child through text.

4.4. Limitations

Throughout the experiment, various limitations arose that could potentially impact the results. The participants recruited for the experiment were aged 19 to 31 and most of them were students, which might be due to the recruiting advertisements that mainly spread among student communities. The limited demographic diversity of the participant pool, consisting mostly of students, may restrict the generalizability of the results to a broader population. Initially, the ideal participants for the study would have been recruited from counsellors-in-training at the child helpline organization, as they possess a better understanding of the experiments context and sufficient experience with the Five-phase model. However, it is worth noting that despite the limitations associated with the participant pool, strong internal consistencies were identified during the reliability assessment of the collected data. These consistent internal relationships among the main measures suggest that the instruments used were robust and outcomes obtained were reliable within the participants.

Furthermore, participants were only given a maximum of five minutes for free interactions with the simulated child, excluding the four predetermined scenarios designed to evoke frustration in the simulated child. By following the provided text instructions during these scenarios, participants may have been restricted in their proactive thinking, as they were not required to think in depth about how to progress the conversation independently. Thus, they didn't have the freedom to explore the system thoroughly. Consequently, participants' feedback and evaluations could be mainly focused on the developed emotion model feature, offering valuable insights into its effectiveness and impact within the conversation, rather than judging the system on the whole.

5

Discussion & Conclusion

In this concluding chapter, we provide a comprehensive reflection on our research journey. The initial section presents a summary of how we have addressed the research question that was established at the beginning of this report. Subsequently, we delve into an examination of the limitations of our work and elucidate our contributions from both a scientific and practical perspective. Finally, we conclude by outlining our recommendations for future research possibilities in this domain.

5.1. Conclusion

This study aims at answering the research question:

How can an emotion model be applied to a child-simulated conversational agent based on the BDI model to assist with the training of child helpline volunteers?

By dividing the research question into three sub-questions, we were able to summarise our study by answering these sub-questions.

- *What design factors should be taken into account when developing a simulated child that incorporates emotion?*

Through exploring cognitive theories of emotions and conducting a focus group with experienced professionals, we were able to obtain valuable insights into the dominant emotion, frustration and the specific situations that induce frustration. This information allowed us to identify the design factors necessary to create a simulated child that can effectively train volunteers in handling emotionally expressive children.

Firstly, the simulated child should have the ability to assess environmental events using its beliefs as criteria. Secondly, the simulated child should adaptively simulate frustration in response to changing appraisals of environmental events, either positively or negatively. Additionally, the simulated child should simulate resistance that occurs under frustration, meaning that it would be harder for the simulated child to connect with counsellors when it becomes extremely frustrated by the deviation from the Five-phase model. Lastly, the simulated child should be able to express frustration using textual cues as the simulated child is text-based.

- *How can emotions be modelled in a simulated child concerning the design factors?*

Based on the design factors, we proposed an Emotion-BDI model, which combined the core components of the BDI model with an additional emotional component. For the first design factor, the Emotion-BDI model utilized the beliefs from the BDI model as the appraised features of the environment. Upon each environmental event where the user inputs a sentence, the Emotion-BDI model would recognize the user intent of the input and then modify belief values accordingly. Moreover, an emotional state was created to indicate the level of frustration, which dynamically changed throughout the proceeding of conversations that were realized through the constant fulfilment of the simulated child's

intentions. Besides, the emotional state is modified to increase if situations that induce frustration occur. To simulate the resistance under frustration, conditions for various desires to be active were adjusted according to the emotional state. The adjustments altered certain belief thresholds to be harder to achieve in case the emotional state was too high indicating extreme frustration and vice versa. Lastly, by incorporating the usage of emotional text cues under corresponding levels of the emotional state, including the usage of emoticons, exclamation marks, and negative terms, the simulated child was capable of expressing the frustration of different intensities.

- *To what extent do trainee counsellors perceive the simulated child with emotions to be believable and emotionally present?*

By conducting pairwise t-tests on the collected data gathered through a within-subject experiment, we provided strong evidence to support our two main hypotheses:

- H1: The participants perceive the simulated child with emotion to be more believable than the simulated child without emotion.
- H2: The participants perceive the simulated child with emotion to be more emotionally present than the simulated child without emotion.

Furthermore, a substantial enhancement in both believability (7.5%) and emotional presence (14.29%) was revealed when comparing the simulated child with emotion over the one without through the comparisons with 14 other established chatbots measured by ASAQ [14]. These findings not only further supported the primary hypotheses but also provided a quantitative response to the third sub-question. Furthermore, the exploratory measures provided additional evidence that the emotion component enhances enjoyment, engagement, and perceived usefulness during interactions. These results were further confirmed by the majority of participants expressing a preference for the simulated child with emotion.

5.2. Contribution

There are three main contributions that our study claims to have achieved:

- We developed a functioning emotionally expressive chatbot tailored to the specific context of training counselors to apply the Five-phase model within child helpline organizations.
- The proposed Emotion-BDI model creates a framework to define frustration from its cause, fluctuation and corresponding expressions. This can be applied to other adaptively changing emotions following our approach, such as happiness, anger etc.
- The simulated child with emotion was proved to be feasible to be used as a training tool to get familiarized with the Five-phase model with higher believability and emotional presence than the simulated child without emotion.

5.3. Limitations

While the evaluation of the experiment provided valuable insights, it is important to consider the limitations of the study. The design of the simulated child is currently able to model only one emotion, frustration. However, counsellors at the child helpline are dealing with much more complicated and diverse situations where more complex emotions might be involved. Despite this, as suggested by experts from the Kindertelefoon, frustration is the dominant emotion expressed by children, the data analysis results also prove that the incorporation of frustration modelling already augments emotional presence and contributes to the feasibility of employing the simulated child as an effective training tool.

Besides, the development of emotional models can also be influenced by a wide range of characteristics of humans, such as personality and mood [41]. While the current emotion model focuses on defining the cause and influence of frustration, it might be able to present more human-like interactions if more human characteristics can be incorporated.

5.4. Future Work

Considering the limitations mentioned above, future studies should aim to expand the development of emotional models for the simulated child. While the current focus on frustration should be sufficient

to realise the training purpose of notifying mistakes, future work aims to incorporate a wider range of emotions that are representative of genuine children's emotional states. This could involve conducting comprehensive studies to identify and model various emotional states experienced by children in different contexts following the Emotion-BDI model framework.

Besides, as suggested by Ojha et al. [41], emotions are complex and can be influenced by a wide range of other human characteristics. By incorporating a more nuanced emotional model with the personality that can affect the emotions people are likely to experience [49], or mood that can augment congruent emotions and lessen incongruent ones [37], the simulated child would be better equipped to respond realistically to user inputs, further enhancing the user experience.

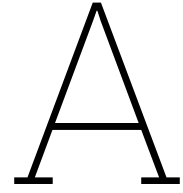
Furthermore, the current work is given a limited storyline where the simulated child is bullied at school. Future work could focus on enhancing its adaptability to changing contexts and user needs. This could involve developing a more flexible framework that allows for the incorporation of multiple storylines and scenarios. By diversifying the experiences and situations encountered by the simulated child, it would better align with the complexities faced by real children and provide a more comprehensive training environment for trainee counsellors.

References

- [1] Carole Adam and Benoit Gaudou. “BDI agents in social simulations: a survey”. In: *The Knowledge Engineering Review* 31.3 (2016), pp. 207–238.
- [2] Keith Anderson et al. “The TARDIS framework: intelligent virtual agents for social coaching in job interviews”. In: *International Conference on Advances in Computer Entertainment Technology*. Springer. 2013, pp. 476–491.
- [3] Eduardo B Andrade and Dan Ariely. “The enduring impact of transient emotions on decision making”. In: *Organizational behavior and human decision processes* 109.1 (2009), pp. 1–8.
- [4] Edoardo Battegazzorre, Andrea Bottino, and Fabrizio Lamberti. “Training Medical Communication Skills with Virtual Patients: Literature Review and Directions for Future Research”. In: *International Conference on Intelligent Technologies for Interactive Entertainment*. Springer. 2020, pp. 207–226.
- [5] Michael Bratman. “Intention, plans, and practical reason”. In: (1987).
- [6] Virginia Braun and Victoria Clarke. “Using thematic analysis in psychology”. In: *Qualitative research in psychology* 3.2 (2006), pp. 77–101.
- [7] Willem-Paul Brinkman. “Design of a questionnaire instrument”. In: *Handbook of mobile technology research methods*. Nova Publishers, 2009, pp. 31–57.
- [8] Jacob Cohen. “A coefficient of agreement for nominal scales”. In: *Educational and psychological measurement* 20.1 (1960), pp. 37–46.
- [9] Juan M Corchado et al. “Development of CBR-BDI agents: a tourist guide application”. In: *Advances in Case-Based Reasoning: 7th European Conference, ECCBR 2004, Madrid, Spain, August 30-September 2, 2004. Proceedings 7*. Springer. 2004, pp. 547–559.
- [10] Brenda Danet, Lucia Ruedenberg-Wright, and Yehudit Rosenbaum-Tamari. “Hmmm where’s that smoke coming from?” In: *Journal of Computer-Mediated Communication* 2.4 (1997), JCMC246.
- [11] Fred D Davis. “Perceived usefulness, perceived ease of use, and user acceptance of information technology”. In: *MIS quarterly* (1989), pp. 319–340.
- [12] Jean-Marc Dewaele and Chengchen Li. “Teacher enthusiasm and students social-behavioral learning engagement: The mediating role of student enjoyment and boredom in Chinese EFL classes”. In: *Language Teaching Research* 25.6 (2021), pp. 922–945.
- [13] Donna D Djordjevich et al. “Preparing for the aftermath: Using emotional agents in game-based training for disaster response”. In: *2008 IEEE Symposium On Computational Intelligence and Games*. IEEE. 2008, pp. 266–275.
- [14] Siska Fitrianie et al. “The Artificial-Social-Agent Questionnaire: Establishing the long and short questionnaire versions”. In: *Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents*. 2022, pp. 1–8.
- [15] Ulrich Gnewuch, Stefan Morana, and Alexander Maedche. “Towards Designing Cooperative and Social Conversational Agents for Customer Service.” In: *ICIS*. 2017.
- [16] Michal Gordon and Cynthia Breazeal. “Designing a virtual assistant for in-car child entertainment”. In: *Proceedings of the 14th International Conference on Interaction Design and Children*. 2015, pp. 359–362.
- [17] Sharon Grundmann. “A BDI-based Virtual Agent for Training Child Helpline Counsellors”. In: (2022).
- [18] Jeffrey T Hancock, Christopher Landrigan, and Courtney Silver. “Expressing emotion in text-based communication”. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. 2007, pp. 929–932.

- [19] Eva Hudlicka. “What are we modeling when we model emotion?” In: *AAAI spring symposium: emotion, personality, and social behavior*. Vol. 8. 2008.
- [20] G Tanner Jackson and Danielle S McNamara. “Motivation and performance in a game-based intelligent tutoring system.” In: *Journal of Educational Psychology* 105.4 (2013), p. 1036.
- [21] Patricia Augustin Jaques, Matheus Lehmann, and Sylvie Pesty. “Evaluating the affective tactics of an emotional pedagogical agent”. In: *Proceedings of the 2009 ACM symposium on Applied Computing*. 2009, pp. 104–109.
- [22] Hong Jiang, Jose M Vidal, and Michael N Huhns. “EBDI: an architecture for emotional agents”. In: *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*. 2007, pp. 1–3.
- [23] Susan M Johnson. “The contribution of emotionally focused couples therapy”. In: *Journal of Contemporary Psychotherapy* 37 (2007), pp. 47–52.
- [24] W Lewis Johnson, Jeff W Rickel, James C Lester, et al. “Animated pedagogical agents: Face-to-face interaction in interactive learning environments”. In: *International Journal of Artificial intelligence in education* 11.1 (2000), pp. 47–78.
- [25] Stichting De Kindertelefoon. “De Kindertelefoon Jaarverslag 2021”. In: (2021). <https://jaarverslag.kindertelefoon.nl/2021> Accessed November 5, 2022.
- [26] Pat Langley, John E Laird, and Seth Rogers. “Cognitive architectures: Research issues and challenges”. In: *Cognitive Systems Research* 10.2 (2009), pp. 141–160.
- [27] Liliana Laranjo et al. “Conversational agents in healthcare: a systematic review”. In: *Journal of the American Medical Informatics Association* 25.9 (2018), pp. 1248–1258.
- [28] Michelle Lefevre, Karen Tanner, and Barry Luckock. “Developing social work students’ communication skills with children and young people: A model for the qualifying level curriculum”. In: *Child & Family Social Work* 13.2 (2008), pp. 166–176.
- [29] Elizabeth A Lemerise and William F Arsenio. “An integrated model of emotion processes and cognition in social information processing”. In: *Child development* 71.1 (2000), pp. 107–118.
- [30] James C Lester, Brian A Stone, and Gary D Stelling. “Lifelike pedagogical agents for mixed-initiative problem solving in constructivist learning environments”. In: *User modeling and user-adapted interaction* 9 (1999), pp. 1–44.
- [31] Langdon E Longstreth. “The relationship between expectations and frustration in children”. In: *Child Development* (1960), pp. 667–671.
- [32] Michael McEachrane. “Emotion, meaning, and appraisal theory”. In: *Theory & Psychology* 19.1 (2009), pp. 33–53.
- [33] M McTear, Z Callejas, and D Griol. “The conversational interface: Talking to smart devices: Springer international publishing”. In: *Doi: <https://doi.org/10.1007/978-3-319-32967-3>* (2016).
- [34] Marvin Minsky. *Society of mind*. Simon and Schuster, 1988.
- [35] Agnes Moors et al. “Appraisal theories of emotion: State of the art and future development”. In: *Emotion Review* 5.2 (2013), pp. 119–124.
- [36] Ranjit Nair, Milind Tambe, and Stacy Marsella. *The Role of Emotions in Multiagent Teamwork*. 2005.
- [37] Roland Neumann, Beate Seibt, and Fritz Strack. “The influence of mood on the intensity of emotional responses: Disentangling feeling and knowing”. In: *Cognition & Emotion* 15.6 (2001), pp. 725–747.
- [38] Emma Norling. “Folk psychology for human modelling: Extending the BDI paradigm”. In: *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 1*. 2004, pp. 202–209.
- [39] Emma Norling and Liz Sonenberg. “Creating interactive characters with BDI agents”. In: *Proceedings of the Australian Workshop on Interactive Entertainment (IE04)*. 2004, pp. 69–76.
- [40] Keith Oatley and Philip N Johnson-Laird. “Cognitive approaches to emotions”. In: *Trends in cognitive sciences* 18.3 (2014), pp. 134–140.

- [41] Suman Ojha, Jonathan Vitale, and Mary-Anne Williams. “Computational emotion models: a thematic review”. In: *International Journal of Social Robotics* 13 (2021), pp. 1253–1279.
- [42] Andrew Ortony et al. “On making believable emotional agents believable”. In: *Emotions in humans and artifacts* (2003), pp. 189–211.
- [43] Andrew Ortony, Gerald L Clore, and Allan Collins. *The cognitive structure of emotions*. Cambridge university press, 2022.
- [44] Martin P Paulus and J Yu Angela. “Emotion and decision-making: affect-driven belief systems in anxiety and depression”. In: *Trends in cognitive sciences* 16.9 (2012), pp. 476–483.
- [45] Catherine Pelachaud. “Modelling multimodal expression of emotion in a virtual agent”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 364.1535 (2009), pp. 3539–3548.
- [46] Joaquin Pérez et al. “A cognitive-affective architecture for ECAs”. In: *Biologically Inspired Cognitive Architectures* 18 (2016), pp. 33–40.
- [47] Robert A Peterson. “A meta-analysis of Cronbach’s coefficient alpha”. In: *Journal of consumer research* 21.2 (1994), pp. 381–391.
- [48] Mihaela-Alexandra Puica and Adina-Magda Florea. “Emotional belief-desire-intention agent model: Previous work and proposed architecture”. In: *International Journal of Advanced Research in Artificial Intelligence* 2.2 (2013), pp. 1–8.
- [49] William Revelle and Klaus R Scherer. “Personality and emotion”. In: *Oxford companion to emotion and the affective sciences* 1 (2009), pp. 304–306.
- [50] Yanet Sánchez-López and Eva Cerezo. “Designing emotional BDI agents: good practices and open questions”. In: *The Knowledge Engineering Review* 34 (2019), e26.
- [51] Stanley Schachter and Jerome Singer. “Cognitive, social, and physiological determinants of emotional state.” In: *Psychological review* 69.5 (1962), p. 379.
- [52] Klaus R Scherer. “Appraisal considered as a process of multilevel sequential checking”. In: *Appraisal processes in emotion: Theory, methods, research* 92.120 (2001), p. 57.
- [53] Ameya Shendarkar et al. “Crowd simulation for emergency response using BDI agent based on virtual reality”. In: *Proceedings of the 2006 winter simulation conference*. IEEE, 2006, pp. 545–553.
- [54] Trine Natasja Sindahl. *Chat Counselling for Children and Youth: A Handbook*. Børns Vilkår Copenhagen, 2011.
- [55] Tzu-Yu Tai and Howard Hao-Jan Chen. “The impact of Google Assistant on adolescent EFL learners willingness to communicate”. In: *Interactive Learning Environments* 31.3 (2023), pp. 1485–1502.
- [56] Marion Valette et al. “Modeling a real-case situation of egress using BDI agents with emotions and social skills”. In: *International Conference on Principles and Practice of Multi-Agent Systems*. Springer, 2018, pp. 3–18.
- [57] Rainer Winkler et al. “Sara, the lecturer: Improving learning in online education with a scaffolding-based conversational agent”. In: *Proceedings of the 2020 CHI conference on human factors in computing systems*. 2020, pp. 1–14.
- [58] Chia-Ching Yuan, Cheng-Hsuan Li, and Chin-Cheng Peng. “Development of mobile interactive courses based on an artificial intelligence chatbot on the communication software LINE”. In: *Interactive Learning Environments* (2021), pp. 1–15.
- [59] Francisco Yus et al. “Attitudes and emotions through written text: The case of textual deformation in Internet chat rooms”. In: (2005).



Design Appendices

ID	Belief name	Phase	Initial Value
<i>about self</i>			
B01	Lilobot thinks he is in control.	All	0.5
B02	Lilobot thinks he is competent to solve problem.	4	0.3
B03	Lilobot feels connected to the Kindertelefoon.	All	0.3
<i>about the Kindertelefoon (KT)</i>			
B04	Lilobot thinks KT can be trusted.	2	0.6
B05	Lilobot thinks KT understands him.	All	0
B06	Lilobot thinks KT is interested in his story.	2	0
B07	Lilobot thinks KT can help him.	3	0.7
B08	Lilobot thinks KT can solve his problem.	3	1
B16	Lilobot thinks he and KT will be able to reach a solution.	4	0
B17	Lilobot thinks KT is going to solve the problem for him.	4	0
<i>about the conversation</i>			
B09	Lilobot thinks he has talked about his situation.	2	0
B10	Lilobot thinks KT is asking about his wish.	3	0
B11	Lilobot thinks KT is asking about a positive wish.	3	0
B12	Lilobot thinks KT is asking about a confidant.	4	0
B13	Lilobot thinks his teacher can help him.	4	0.3
B15	Lilobot thinks KT wants to end the conversation.	5	0
<i>about confidant</i>			
B14	Lilobot feels safe in the conversation.	All	1

Table A.1: Beliefs and their altered default initial values adapted from Grundman's work [17].

B

Evaluation Appendices

B.1. Prototype

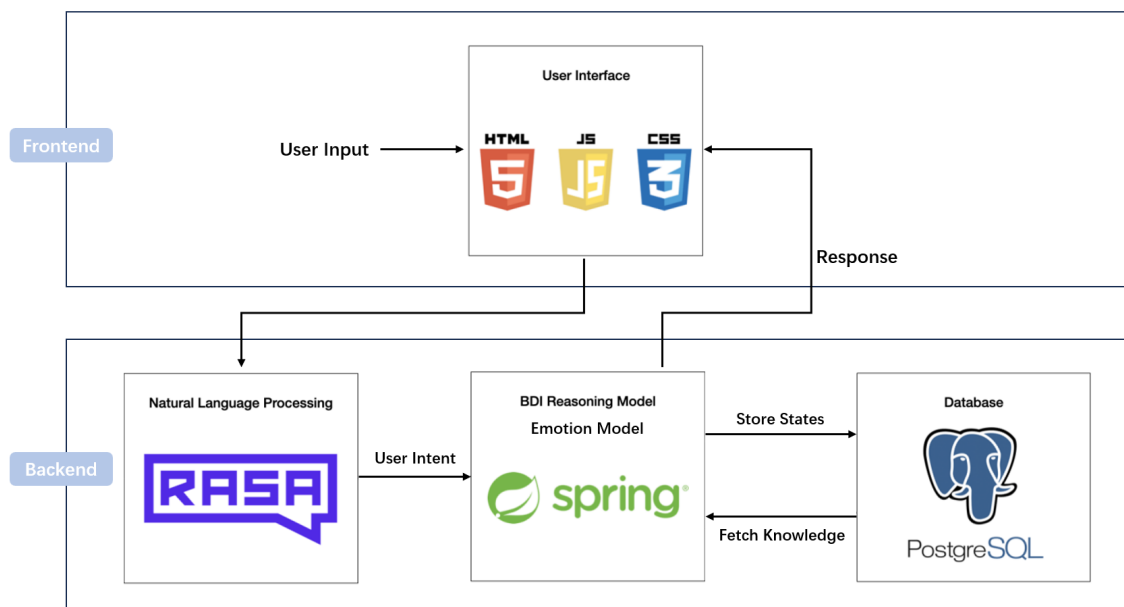


Figure B.1: Adapted System Architecture of the chatbot framework [17].

B.2. Informed Consent Form

Informed Consent Form

Emotion Model for Child Helpline Training Tool

You are being invited to participate in an experiment of a research study titled *Emotion Model for Child Helpline Training Tool*. This study is being done by Dongxu Lu, and supervised by Willem-Paul Brinkman, and Mohammed Al Owayyed; All of which are affiliated with the TU Delft.

The purpose of this experiment is to evaluate the practicality and feasibility of a virtual agent that simulate a child suffering from school bullying, which will be used as training tool for volunteers from a child helpline and will take you approximately 45 minutes to complete. The data will be used for the measurements and analysis of the virtual agent. We will be asking you to interact with the virtual agent and finish questionnaires with questions regarding your experience with the agent.

As with any online activity the risk of a breach is always possible. To the best of our ability your answers in this experiment will remain confidential. We will minimize any risks by anonymizing all the data collected during experiments. Your email address will only be used for contact purpose and will be deleted after the experiment.

Your participation in this experiment is entirely voluntary **and you can withdraw at any time during the experiment**. You cannot remove your data after finishing the experiment as the data will be collected and anonymized. All the anonymized results will be uploaded to 4TU.ResearchData.

For more information please contact:

- Dongxu Lu: [REDACTED]

If you agree and consent to this Opening Statement, you can now fill in the consent form below.

Consent Form

Please fill out the consent form below by answering either "Yes" or "No" for each question. Please note that a "No" for any single question renders you ineligible to participate in this study.

PLEASE TICK THE APPROPRIATE BOXES	Yes	No
1. I have read and understood the experiment information dated [05/06/2023], or it has been read to me. I have been able to ask questions about the experiment and my questions have been answered to my satisfaction.	<input type="checkbox"/>	<input type="checkbox"/>
2. I consent voluntarily to be a participant in this experiment and understand that I can refuse to answer questions and I can withdraw from the experiment at any time, without having to give a reason.	<input type="checkbox"/>	<input type="checkbox"/>
3. I understand that taking part in the study involves: <ul style="list-style-type: none"> ● I will interact with a virtual agent through a chat interface. ● I will fill out questionnaires with questions regarding my experience of interacting with the virtual agent. 	<input type="checkbox"/>	<input type="checkbox"/>
4. I understand that the the experiment will end in no more than one hour.		
5. I understand that taking part in the experiment involves possible data leakage. I understand that these will be mitigated by anonymizing the data collected during the experiment and my right to stop the experiment at any point.	<input type="checkbox"/>	<input type="checkbox"/>
6. I understand that the following steps will be taken to minimise the threat of a data breach and protect my identity in the event of such a breach: anonymous data collection and secure data storage.	<input type="checkbox"/>	<input type="checkbox"/>
7. I understand that personal information collected about me that can identify me, such as email address, will not be shared beyond the study team and will not be linked to my result.	<input type="checkbox"/>	<input type="checkbox"/>
8. I understand that after the research study the de-identified information I provide will be used for the measurement and analysis of the practicality and feasibility of the virtual agent, and scientific outlet, for example, education/academic/commercial purposes.	<input type="checkbox"/>	<input type="checkbox"/>
9. I give permission for the de-identified questionnaire that I provide to be archived in TU Delft project storage repository so it can be used for future research and learning.	<input type="checkbox"/>	<input type="checkbox"/>
10. I agree to participate in this experiment.	<input type="checkbox"/>	<input type="checkbox"/>

Signatures

Name of participant [printed]

Signature

Date

I, as researcher, have accurately read out the information sheet to the potential participant and, to the best of my ability, ensured that the participant understands to what they are freely consenting.

Researcher name [printed]

Signature

Date

Study contact details for further information: [*Name, phone number, email address*]

B.3. Instructions on experiment scenarios

Thank you for participating in the experiment. Your role in this experiment is **to intentionally make mistakes** and **engage in conversations** with simulated children in two separate sessions **as a trainee counsellor** at a child helpline organization under **four scenarios**. Example dialogue of each scenario is provided which you can follow by **the order of number**. You are also free to use your own words to convey similar meanings.

Please follow these guidelines to create scenarios where you **intentionally make mistakes**. Say **“Bye”** when you finish one scenario and **ask the instructor to set up the following experiment**:

Scenario 1: Insufficient Acknowledgement.

This scenario simulates a situation where the counselor does not show enough empathy or acknowledgement to the child during the first phase (Building rapport).

➤ **Phase 1:**

- **Please give warm greetings.**

Example: 1. “Hello, welcome to the Child Helpline.”

2. “How have you been?”

3. “Do you want to tell me something?”

- **Please show disdain towards the child’ story.**

Example: 4. “You are overreacting.”

Scenario 2: Refusal.

This scenario simulates a situation where the counselor cannot fulfill the child's wish objectively, such as calling the school, and refuses the child during the third phase (Setting a goal for the session).

Example:

- Phase 1:
 - **Please give warm greetings.**
 1. "Hello, welcome to the Child Helpline."
 2. "How have you been?"
 3. "Do you want to tell me something?"
 - **Please show acknowledgement towards the child's story.**
 4. "I know that it's hard for you."

- Phase 2:
 - **Please ask three questions about the story details.**
 5. "When did this happen?"

- Phase 3:
 - **Please ask about the child's desire.**
 6. "What do you want to accomplish with this conversation?"
 7. "What can I do for you?"
 - **Please refuse if the desire is not doable.**
 8. "Sorry, we cannot do that."
 - **Please suggest looking at a different solution.**
 9. "I suggest that we look for a different solution together."
 - **Please ask if the child have people to seek help.**
 10. "Is there someone you can talk to?"
 - **Please confirm the child's suggestion on talking to the teacher.**
 11. "She can listen to your story and support you."

- Phase 4:
 - **Please ask three questions on how to realize the solution.**
 12. "How would it feel to talk to Mrs. Ellie?"

- Phase 5:
 - **Please confirm the solution that both agree on.**
 13. "Shall we agree that you will talk to your teacher about your situation?"
 - **Please end the conversation.**
 14. "I suggest we end the conversation now."

Scenario 3: Setting up goals without Asking.

This scenario simulates a situation where the counselor does not ask about the child's desires and provides a solution thinking it's appropriate during the third phase (Setting a goal for the session).

Example:

- Phase 1:
 - **Please give warm greetings.**
 1. "Hello, welcome to the Child Helpline."
 2. "How have you been?"
 3. "Do you want to tell me something?"
 - **Please show acknowledgement towards the child's story.**
 4. "I know that it's hard for you."

- Phase 2:
 - **Please ask three questions about the story details.**
 5. "When/Where did this happen?"

- Phase 3:
 - **Please give a solution directly without asking the child's wishes.**
 6. "Do you want to talk to your teacher?"
 7. "Why don't you feel comfortable talking to your teacher?"

Scenario 4: Too many questions

This scenario simulates a situation where the counselor asks too many questions and digs for excessive details from the child during the second phase (Clarify the child's story).

Example:

- Phase 1:
 - **Please give warm greetings.**
 1. "Hello, welcome to the Child Helpline."
 2. "How have you been?"
 3. "Do you want to tell me something?"
 - **Please show acknowledgement towards the child's story.**
 4. "I know that it's hard for you."

- Phase 2:
 - **Please ask eight questions about the story details.**
 5. "When\ Where\ Why did this happen?"

B.4. Questionnaires

B.4.1. Believability

The following statements evaluate how much you believe that the simulated child behaves like a real child suffering from bullying. Please rate in terms of how much you agree or disagree with each statement (-3: Disagree, 0: Neither Agree nor Disagree, 3: Agree).

Statement	Rating						
	-3	-2	-1	0	1	2	3
Statement 1: A child would behave like the simulated child.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Statement 2: The simulated child's manner is consistent with that of children.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Statement 3: The simulated child's behaviour makes me think of children's behaviour.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Statement 4: The simulated child's behaviour makes me think of children's behaviour.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Statement 5: The simulated child behaves like a real child.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Statement 6: The simulated child has a child-like manner.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Note: Adapted from Item 1.2 Human-Like Behaviour of ASAQ [14]

B.4.2. Emotional Presence

The following statements evaluate how much you believe how emotionally that the simulated child behaves like a real child suffering from bullying. Please rate in terms of how much you agree or disagree with each statement (-3: Disagree, 0: Neither Agree nor Disagree, 3: Agree).

Statement	Rating						
	-3	-2	-1	0	1	2	3
Statement 1: The simulated child is emotional.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Statement 2: The simulated child experiences emotions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Statement 3: [R] The simulated child is emotionless.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Statement 4: The simulated child can express its feelings.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Statement 5: [R] The simulated child cannot experience.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Note: Adapted from Item 18.1 Agents Emotional Intelligence Presence of ASAQ [14]

B.4.3. Qualitative Measure

According to your experience with the simulated children in the two sessions, which session do you prefer to use as a training tool and why?