

MENG INDIVIDUAL REPORT

DEPARTMENT OF COMPUTING

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

**Chatbot for learning to laugh: New
approach to develop a sense of
humour**

Author:

Zeena Patel

Supervisor:
Professor Abbas Edalat

Second Marker:
Dr Anandha Gopalan

July 15, 2022

Submitted in partial fulfillment of the requirements for the MEng Computing
Degree of Imperial College London of Imperial College London

Abstract

The benefits of humour and laughter as a therapeutic aid for mental health are well-known and backed by extensive research. Advances in the field of conversational agents and their use in treating mental illness mean that we can create virtual agents that assist the user in learning to laugh, which could be groundbreaking for mental health and well-being.

In this work, we develop the first humorous chatbot that can engage with the user and teach them how to learn to laugh. Our chatbot, HaHA (the Happy Humorous Assistant), guides the user through learning and practising the self-initiated humour protocols as a new approach to develop a sense of humour. We implement a rule-based framework and augment our inherently safe and reliable framework with our own machine learning platform that uses NLU (natural language understanding) for emotion recognition of the user's response along with NLP (natural language processing) for the retrieval of humorous, empathetic, novel and fluent statements. We present a new dataset from which the chatbot retrieves these statements. In addition, we build a humour recognition model that scores their humour with 97% accuracy and macro-f1.

We evaluate our framework through a non-clinical trial with 38 participants who interacted with HaHA for at least three days over a week. Our framework is scored highly for its efficacy in teaching the user how to learn to laugh (82%) and recognise contexts for laughter (89%), along with its ability to converse humorously (97%), empathetically (93%), and politely (81%), and for its user engagement (85%), among several additional metrics. Our framework consistently scored higher than previous frameworks of a similar nature for the comparable metrics.

Acknowledgments

Firstly, I would like to thank my supervisor, Professor Abbas Edalat, for his continuous support and enthusiasm throughout this project, enabling its successful completion.

I would also like to thank Dr Anandha Gopalan for his insightful feedback on my interim and final report and for organising weekly meetings from the start of the project, helping me stay on track throughout.

Moreover, I am extremely thankful to Lisa Alazraki, developer of the latest SAT chatbot, for her significant help in answering questions regarding its implementation and providing valuable advice throughout all stages of the project.

Finally, I would like to thank my family for their constant love and support throughout my degree and life.

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Objectives	2
1.3	Contributions	2
2	Background	4
2.1	Humour	4
2.1.1	Defining humour and laughter	4
2.1.2	Psychology of humour	5
2.1.3	Philosophies of humour	10
2.1.4	Self-initiated humour	12
2.2	Related work	12
2.2.1	Computational humour	12
2.2.2	Chatbots and humour	18
2.3	Technical preliminaries	22
2.3.1	N-gram language models	22
2.3.2	Tf-idf	23
2.3.3	Classification algorithms	23
2.3.4	Neural language models	23
2.3.5	Evaluating classifiers	25
3	Design and implementation	26
3.1	Conversation flow	26
3.2	Dataset	30
3.2.1	Requesting rewritings	30
3.2.2	Response collection	31
3.2.3	Dataset overview	32
3.3	Machine learning platform	33
3.3.1	Humour recognition	33
3.3.2	Optimisation function	38
3.4	Web application	41
3.4.1	Back end and database	41
3.4.2	Front end and user experience	43
3.4.3	Cloud deployment	43

4 Evaluation	45
4.1 Trial details	45
4.2 Participant feedback	46
4.2.1 Light-hearted and humorous	46
4.2.2 Polite and non-instructional	46
4.2.3 Usefulness in teaching users how to learn to laugh	47
4.2.4 Conversation flow	47
4.2.5 Ease of use and accessibility	48
4.2.6 Empathy and engagement	49
4.2.7 Further feedback	50
4.2.8 Evaluation from clinicians	52
4.3 Conversational insights	52
4.4 Study limitations	54
5 Conclusions and future work	55
5.1 Conclusions and contributions	55
5.2 Future work	56
5.2.1 Dataset extension	56
5.2.2 Generative model	56
5.2.3 Animation	57
5.2.4 Adaptable conversation flow	57
5.2.5 Protocol theory integration	57
5.3 Ethical considerations	57
5.3.1 Non-clinical trial	58
5.3.2 Collection of our dataset	58
A Self-initiated humour protocols	71
B Dataset and sample	77
C Example conversations	84
D Mini session information	95
E Optimisation function selection	97
F Evaluation feedback	100

Chapter 1

Introduction

1.1 Motivation

It is estimated that over one in ten people globally suffer from a mental disorder, the most common of which are anxiety and depression [1]. In addition, approximately one in five people in post-conflict settings have a mental health condition. These conditions adversely impact work performance, relationships, and participation within the community and also lead to 1 in 5 years lived with disability [2]. Despite this, a large proportion of those suffering from a mental disorder, over 75% of people in low- and middle-income countries receive no treatment. Additionally, the COVID-19 pandemic has largely accentuated the global lack of accessibility to treatment for and the prevalence of mental disorders [3]. As such, the use of technology in treating mental illness has become an urgent requirement.

The benefits of humour and laughter as a therapeutic aid for mental health and effective medicine for physical health are well-known and widely backed by extensive research [4]. Their direct relation to a more positive psychological well-being and ability to assist people in coping with stressful events and facilitate robust interpersonal relationships has enabled a growing interest in the role of humour in the clinical practice of psychotherapy. For centuries, humour and laughter have been theorised and debated by many philosophers, psychologists and scientists, agreeing that they are profound but mysterious human demeanours.

A new method for learning to laugh and develop a sense of humour, established by Professor Edalat, involves practising self-initiated humour protocols (SIHP) [5]. Underlying these protocols are four hypotheses related to laughter, established from several classic and contemporary laughter theories, proposed as part of the self-attachment technique in algorithmic human development. SIHP are self-administrable, meaning users can practice them by themselves. Also, with the current technological revolution, there has been significant growth in the use of virtual agents for treating mental illness (see subsection 2.2.2). As such, we can create agents that assist the user in learning to laugh, a potentially groundbreaking solution for mental health and well-being.

1.2 Objectives

Using the following formalisms, we aim to develop a humorous chatbot that can teach the user how to learn to laugh. This chatbot would be an artificial intelligence (AI) agent in human-computer interaction (HCI) that will guide the user through learning and practising the self-initiated humour protocols. It should engage with the user in a humorous, empathetic and even modest manner.

- Gain insight into the philosophies and psychology of humour.
- Gain familiarity with the self-initiated humour protocols, including their underlying theory of laughter as part of the self-attachment technique in algorithmic human development.
- Build a rule-based framework for the chatbot that guides the user through these well-defined self-initiated protocols.
- Augment this framework with a machine learning platform that ensures the chatbot's conversations are engaging, as well as empathetic and humorous.
- Evaluate the chatbot in a trial by human participants.

1.3 Contributions

Our contributions in this project are five-fold.

1. A conversation flow that outlines how our chatbot converses with users, including its questions and responses (see Section 3.1).

Our chatbot flow guides users through the self-initiated humour protocols based on a contextual understanding of their current state (i.e. emotional state, attitude towards protocols, experience with protocols).

2. A new dataset augments the conversation flow of our retrieval-based chatbot and is used to select its response to the user (see Section 3.2).

Our new dataset, HaHAsStatements, consists of 2,507 rewritings of base statements, primarily crowd-sourced.

3. A machine learning platform that augment the rule-based framework enabling our chatbot to select an optimal statement from HaHAsStatements when responding to the user (see Section 3.3).

As part of this platform, we build a humour recognition model capable of scoring statements according to their humour with 97% accuracy and macro-f1. We also develop a retrieval function that optimises humour, empathy, fluency and novelty scores of an input statement and has two forms corresponding to a user's emotional state.

4. A web application that integrates our conversation flow, curated dataset and ML platform and surfaces them to the user (see Section 3.4).

In the back end of our application, we develop a rule-based framework for our chatbot that supports the delivery of the self-initiated humour protocols and implement the collection of important user insights via a database. We integrate our framework into a web interface and deploy our web application to use in a remote trial.

5. A formal evaluation of our platform via a remote non-clinical trial consisting of 38 participants (see Chapter 4).

Our virtual agent is scored highly for its efficacy (i.e. in teaching the user how to learn to laugh and recognise contexts for laughter), its ability to converse humorously, empathetically, and politely, and for its user engagement, conversation flow and ease of use. It also performs better than both previous SAT chatbots [6, 7] in areas where we can make comparisons (empathy and engagement). See Section 4.2 for the detailed breakdown of scores for each metric. HaHA's emotional and educational impact is assessed very positively in writing (see Appendix F for the written feedback from all trial 1 participants).

After fine-tuning, in line with feedback received, our framework is currently undergoing a second trial of AHD research team members (or close associates), which includes four medical professionals who have specialised in mental health. See Section 4.2.8 and Appendix F for the clinicians' evaluation questionnaire responses to the multiple choice and written sections, respectively.

Chapter 2

Background

In this chapter, we introduce relevant context for understanding our work. Section 2.1 describes humour, including its key definitions, psychology and philosophies, and the concept of self-initiated humour as a new approach to develop a sense of humour, which forms the basis of our project. Section 2.2 explores existing works relevant to our domain, such as computational humour and current approaches to chatbots and humour in human-computer interaction. Finally, Section 2.3 covers NLP techniques that provide a technical preliminary to our work.

2.1 Humour

2.1.1 Defining humour and laughter

We present the definitions of key terms related to humour and laughter, which will appear throughout this report.

Humour is one of the stimuli which elicits laughter in individuals and enables them to feel happy [8].

When analysing its psychology, we consider Martin et al's Four Humor Styles [9]:

- i Affiliative humour is used to achieve interpersonal rewards (i.e. enhance social relationships).
- ii Self-enhancing humour is used to achieve intrapersonal rewards (i.e. to improve and maintain positive well-being while avoiding adversity).
- iii Aggressive humour is used to enhance oneself at the expense of others (i.e. through criticism or manipulation).
- iv Self-defeating humour engages excessively self-disparaging humour (i.e. through making fun of self-weaknesses or laughing along when ridiculed).

Styles (i) and (ii) are “adaptive and non-hostile”, whereas (iii) and (iv) are “mal-adaptive and hostile” [5]. The Humor Styles questionnaire assesses an individual’s

use of these styles.

Sense of humour is a unique psychological attribute that enables individuals to respond to various types of humorous stimuli [10].

Laughter is a universal human demeanour found in every culture and existing individual [11].

We consider two distinct types of laughter for the purpose of this project [5]:

- Spontaneous or genuine, humour-based laughter, also called Duchenne laughter.
- Self-induced or intentional, emotionless laughter, also called non-Duchenne laughter.

2.1.2 Psychology of humour

We explore the psychology of humour, a field that has rapidly grown in the past couple of decades, with scholars in diverse areas of psychology placing more effort in examining the role of humour within their domains.

Developmental psychology of humour

There is ample research to suggest that our past experiences, rather than genetics, determine our appreciation of different types of humour, with our family environment being an influential aspect of our sense of humour development [4].

One theory proposed to explain how the interaction between a child and their parents may influence the development of their sense of humour is the “stress and coping hypothesis”, which states that children raised in disagreeable family environments develop a sense of humour to cope and to earn approval [12]. The alternative explanation proposed is the “modeling/reinforcement hypothesis”, which states that children raised in a more safe, nurturing environment, where parents frequently enjoy humour, develop a sense of humour due to parental modelling and reinforcement. We have seen that humour serves various social functions [4]. As such, there are likely many alternative routes that have led to our unique differences in humour.

Humour and mental health

Current research has demonstrated the extensive benefits of humour on our general well-being, including its use as a therapeutic aid for mental health [13, 4].

We must account for the style of humour [14] when analysing its impact on our mental health. A systematic review (consisting of 37 studies) was conducted to find correlations between each humour style and the four areas of mental health:

self-esteem, life satisfaction, optimism and depression [15]. Health-promoting humour styles (i.e. affiliative and self-enhancing) were found to be positively associated with mental health (i.e. self-esteem, life satisfaction and optimism), while self-defeating humour styles were negatively associated with mental health and positively associated with depression. One important conclusion drawn from the review was that adopting specific humour styles will likely benefit our mental health. Another was that an awareness of these styles could help indicate underlying emotional states, which would lead to more effective therapy.

Humour and our psychological well-being

Humour has been shown to directly relate to more positive psychological well-being from an experimental, clinical and correlational standpoint [4]. An extensive review [8] by Dr Mora-Ripoll on the psychological benefits of humour and laughter indicates that they promote mood, improve creative thinking, enhance life quality and patient care, and are contagious, among additional benefits. Several laboratory experiments illustrated that exposure to humorous stimuli induced positive emotions in the participants, such as cheerfulness [16] and hopefulness [17]. An increasing number of studies are also investigating the impact of the frequency of laughter on our mental well-being [18]. For example, a recently developed convenient laughter-prescribing tool, the “Laughie”, showed maximum benefits when used three times a day for a week; it improved the well-being of healthy adults by 16% [19]. Standard laboratory techniques have demonstrated that exposure to humour has also diminished the experience of negative emotions, including depressed moods [20] and state anxiety [21].

A more indirect benefit of humour on our psychological health includes its use as an effective coping mechanism [13], especially with stressful events [4], which could prevent their adverse impacts on our well-being. Dr Mora-Ripoll’s review also found that humour could help counteract symptoms of depression. An interesting finding is that the biochemical changes and analgesic qualities of laughter, combined with tears, were shown to alleviate tension, anger, fear, and loss [22]. In addition, humour and laughter have illustrated their ability to improve coping mechanisms for anxiety and loneliness [23].

Humour has also been found to help facilitate healthy interpersonal relationships, which can, in turn, significantly impact an individual’s happiness and psychological well-being [4]. An existing study illustrates the stress-buffering and health-improving benefits of laughter via improving an individual’s social abilities and interpersonal skills, which could provide them with further social support [24]. Moreover, we return to Dr Mora-Ripoll’s review, indicating the psychological benefits of laughter in a social context, such as improving interpersonal interaction, increasing friendliness and helpfulness, and building group identity and solidarity [8].

Although these research findings cannot conclude that laughter is an all-around psychological healing agent, there is sufficient evidence to imply that laughter has a beneficial, quantifiable impact on multiple aspects of mental health. As such, Dr Mora-Ripoll recommends the following advice to patients and health care professionals: “Add laughter to your working and daily life, remember to laugh regularly, share your laughs, and help others to laugh, too.” [8] This advice is particularly relevant to our work, featuring in a more structured and guided setting as a new approach to develop a sense of humour.

Humour in psychotherapy and counselling

Based on the established idea that humour and laughter have significant mental health benefits, therapists and scholars are showing a rising enthusiasm for their role in psychotherapy and counselling. Over the past decades, several humour-based and laughter-based interventions have gained widespread acceptance [8].

The impact of humour in psychotherapy can be demonstrated by three general methods.

1. Humour as therapy

This method involves developing a therapy system mainly based on humour or laughter, and existing methods include Rational Emotive Therapy (RET), Provocative Therapy and Natural High Therapy. Unfortunately, there has not been sufficient research conducted to assess the efficacy of most of these therapies or to draw comparisons with other treatment forms. Albert Ellis, the creator of RET, has advocated using humour in psychotherapy to diminish and discredit profound irrational beliefs at the root of psychopathology [25].

Since there is strong evidence backing the ability of laughter to improve our mental health, two types of laughter therapy have emerged in the past two decades [26]. The first type is humour-based, i.e. spontaneous laughter therapy [27], which has been found to reduce the loneliness of nursing home residents [28], improve the quality of life and cognitive function in cancer patients, and reduce their pain [18]. The second type of laughter therapy, laughter yoga, supports self-simulated, i.e. intentional laughter. Several studies have indicated that these interventions support the mental health of adults [28, 29]. Purposeful laughter, as practised in laughter yoga, which is becoming increasingly prevalent worldwide, has significantly enhanced various aspects of self-efficacy, including self-regulation, positiveness, and social identification [8].

A systematic review and meta-analysis of these laughter-inducing therapies were conducted in 2019 [26]. The results suggested that, in general, simulated laughter is more effective than spontaneous laughter and that these therapies can help treat depression. In particular, another important use case

could be the elderly population; with rising health care costs and the increasing population, there exists a potential for low-cost, simple interventions that staff can administer with minimal training. Other benefits for the elderly include positive effects on insomnia and sleep quality and its help in treating patients suffering a decline of cognitive function, a primary symptom of dementia [30]. The writers of the 2019 review concluded that therapies have the potential to serve in addition to existing therapies, but overall, the quality of evidence is “low”, and more systematic research is needed in this research field [31]. Additional work in this direction is currently underway in the UK, with their National Health Service preparing to introduce laughter courses for traumatised patients to aid their recovery [32].

2. Applying humour to therapeutic techniques

Another approach of using humour-based intervention techniques that incorporate humour for treating particular problems has also been devised by some clinicians [4]. In addition to several case examples, a clinical study investigating the use of humour in phobia treatment concluded that humour based-intervention might be just as effective as standard muscle relaxation in established phobia treatment techniques, namely systematic desensitisation [33]. One remarkable intervention demonstrating that humour can support psychological needs under various stressful situations was when a therapist successfully used humour to treat a teenage girl who had been mute for three years [34].

3. Communicating with humour during therapy

An additional method of applying humour to psychotherapy is using it as a communication skill [4]. Clinicians found that for this technique to be successful, the therapist must be skilled in using humour positively [35]. With this outlook, they view humour as something that occurs naturally throughout therapist and client exchanges. Research reveals that humour seems to be a useful therapeutic communication skill that, when used accordingly, i.e. with sensitivity and empathy, can assist therapists in achieving the common therapeutic goals [36]. We will revisit this concept in Section 3.2, where we describe key measures taken to ensure our chatbot’s use of humour is appropriate for our context.

Overall, there is limited and generally mixed research on the effects of humour on the therapeutic process across several studies [37, 4]. The reason may be due to different therapeutic effects of the separate styles or uses of humour [4]. For those studies where researchers have tried to distinguish types of humour, they might not have managed to spot important distinctions between therapeutic and “counter-therapeutic” forms. It is clear from the pervasiveness of humour and laughter in therapy and numerous credible ideas regarding its possible uses and risks that this research subject deserves more attention.

The therapeutic benefits of humour are apparent if used by therapists in a genuine, understanding and caring manner [38]. However, due to its intrinsically ambiguous nature, humour used in therapy could, by chance, be misunderstood by the client and interpreted as aggressive [4]. Also, it could be counter-therapeutic and harmful if a client is left feeling misunderstood or if the therapist is using humour to hide their discomfort with issues raised by clients. In a well-cited article [39], a therapist voiced substantial reservations regarding the use of humour in psychotherapy, mentioning numerous risks, such as a client perceiving a therapist's use of humour as not wanting to take their problems seriously, for example. It is crucial to consider the above risks when choosing a framework for the chatbot presented in this project.

We have seen that much of the literature on the use of humour in therapy has come from specific case studies and clinical impressions. The requirement for evidence-based approaches to therapy and interventions that demonstrate treatment efficacy is becoming more widely acknowledged [4]. However, as it stands, there is insufficient empirical research examining the effectiveness of humour-based interventions with clinical populations or the types of humour that may be suitable or not for therapy. Further research is required to analyse potentially beneficial or even detrimental uses of humour in treating particular psychological problems with different client groups.

Humour and physical health

Non-hostile humour and laughter are recently becoming recognised as “the best medicine” for physical health [4, 40], with substantial existing evidence in favour of this [18].

Dr Mora-Ripoll’s review of laughter literature across medicine and health care also highlights the vast physiological benefits of laughter [39]. These benefits include its potential to exercise and unwind muscles, enhance respiration, promote circulation, reduce stress hormones, boost immune system defences, elevate pain threshold and tolerance, and improve mental function. In fact, from all of the potential health benefits of humour and laughter, research has indicated the most consistent support for its possible pain-reducing effects [4]. It is Duchenne laughter, where muscle contractions around the corners of the mouth and eyes trigger endorphins to be released into the body that can improve our pain threshold.

A comprehensive systematic review conducted in 2020 notes that William Fry, the pioneer of gelotology (the study of laughter), considered laughter to affect the “whole physical being” and demonstrated the effects of mirthful (humour-induced) laughter in improving blood pressure and cardiovascular function [42]. In addition, it suggests a clear association between an individual’s sense of humour and their increased survival to the age of 85. This finding is also in line with research indicating that certain types of humour may positively impact longevity [4].

2.1.3 Philosophies of humour

We review several philosophies of humour, classic and contemporary, that have been influential in humour research and in the development of this project.

Here are the three classic theories of laughter and humour [41]:

1. **The superiority theory** describes how we laugh when we feel superior to others or our past selves. This idea was based on the initial outlook on humour as a form of aggression when we feel triumphant over other people. It even dates back to Greek history, where humour was seen as a “species of the ugly” by Aristotle. Over the past century, this theory has been expanded upon by examining its underlying evolutionary processes, as we will see when we explore more contemporary approaches to humour. It is clear that there are aggressive tones to many jokes, although not that all humour represents aggression [4]. In addition, the evolutionary theory considered as the basis for this was rather obsolete as it does not capture the adaptive value of humour nor the aggressive contexts of the evolution of humour.
2. **The incongruity theory** states that we laugh when we perceive something contradictory or when something goes against our expectations. The notion of “bisociation”, where there are two conflicting interpretations of a situation, was introduced to this theory by Arthur Koestler in the twentieth century, where switching between these interpretations triggers humour [41]. This concept is used in several jokes, where the setup marks the first interpretation, and the punchline marks the second unanticipated conflicting interpretation. Researchers have highlighted the conditions for incongruity alone to be sufficient for humour as a refinement to the incongruity theory. These conditions are that incongruity should be resolved in some way and perceived with a “humour mindset” [4]. It is clear that incongruity is a fundamental aspect of humour, though theorists disagree on its exact function and context that elicits humour.
3. **The relief theory** explains that when we laugh, it is because we are releasing built-up nervous energy. An example could be when we laugh with relief following the diversion of a near-accident or when something we were initially very fearful of turned out fine [41]. This theory is not widely supported among scholars in philosophy or psychology today, as there are many laughter situations that are not due to the relief of emotional energy and laughter often occurs suddenly without any build up. Critics have also noted that the theory does not capture why something might be funny, and it is not based on a scientific understanding of the nervous system [4].

Although difficult to test and not completely comprehensive in their capturing of the different forms and contexts of humour, these classic theories have provided important conceptual foundations behind today's humour theory and research [4].

In our research, we also consider the following contemporary theories that account for some of these humour forms and contexts [5]:

4. **The perspective theory**, inspired by Charlie Chaplin, a pioneer in modern comedy, is captured by his quote: “Life is a tragedy when seen in close-up, but a comedy in long-shot.”, i.e. if you get very close to a tragic situation and examine its details, you might become alarmed by the tragedy, yet, by taking a step back and reflecting on it while considering its context, this situation can become humorous. He also described how we can adopt this outlook by incorporating the idea of play within our laughter: “To truly laugh, you must be able to take your pain, and play with it!” Dr R. Provine, a behavioural neurobiologist, shares that over 80% of the time, laughter is not a response to something humorous, but instead, a part of everyday conversation [31]. For example, “Gotta go now. Ha, ha, ha.” The following contemporary theory captures these social contexts of laughter as a play signal, establishing Chaplin’s argument.
5. **The evolutionary (or playful) theory** describes the idea that laughter has evolved as a play signal in higher primates to reveal non-aggressive intent in pretend fights. This is in line with the strongest contemporary follower of the superiority theory, Charles Gruner’s outlook on humour as “playful aggression” [41]. For example, when one has a feeling of triumph after finally winning a difficult game. The evolutionary basis of this theory was that this tendency for competition and aggression facilitated the survival and thriving of the human race. This idea that laughter has evolved from mock-aggression accounts for the playfully aggressive nature behind so much of today’s humour. This evolutionary theory also plays a role in everyday life from childhood, at any playground or park where we laugh while playing with each other. At this stage, play and thus laughter, is the basis of our learning, whereas its role in learning diminishes in our later life as we become exposed to the competitive and achievement-oriented educational system [5], and it quickly shifts to serve more as a “social glue” to reinforce group ties [31].

It is important to note that although there are numerous philosophies of humour, a consensus has not yet been reached across the scientific community on seemingly simple fundamental ideas, such as the essence of humour or the cause of our laughter [41].

We now highlight an important argument of Morreall, a well-known contemporary philosopher of laughter [5].

Morreall states that although the natural response to an upsetting event is to feel negative, i.e. sad, angry, anxious, etc., a more mature outlook on life would elicit the replacement of such negative emotions with a rational understanding of the upset. He calls this act “reality assimilation”. In agreement with Chaplin’s views on incorporating play in our laughter, he proposes that the optimal approach would be to laugh the upset off. Hence, also in line with Chaplin’s perspective theory, we can learn to understand the roots of why we’re feeling upset and, with time, laugh off our negative emotions. This personal journey might be longer for some as individuals require different times to successfully overcome setbacks and hardships and approach them creatively and playfully to initiate laughter.

2.1.4 Self-initiated humour

Professor Edalat established a new method for learning to laugh and develop a sense of humour by practising self-initiated humour protocols (SIHP) [5] as part of the self-attachment technique (SAT) in algorithmic human development [42].

Underlying the concept of SIHP are four hypotheses related to laughter, which can be found in Appendix A. This approach incorporates spontaneous and self-induced paradigms making these protocols humorous and intentional [5]. The self-initiated humour protocols are based on the five philosophies of humour and they provide an algorithmic framework for laughing.

Self-initiated humour (SIH) is accessible to any individual, regardless of their mental state and ability to laugh over general humour, since it exploits humour contexts that lie in the self, i.e. in their thoughts, successes, hardships, etc. We also intend for SIHP to be practised while in a playful mode, as this will better allow one to experience life's events with amusement.

The self-initiated humour protocols

Twelve exercises for self-initiated humour have been developed by Professor Edalat and these are enumerated in Appendix A.

The first two protocols prepare us to be in a playful state, and the rest expose accessible contexts and triggers for Duchenne laughter. They should be practised alone, at least initially, and to gain health benefits from laughing, we must ensure that our laughter is non-hostile, i.e. it does not encompass self or external contempt. These protocols are self-administrable and are learned and exercised alone by the user, which paves the way for the development of HaHA, a virtual agent that can help the user learn to laugh.

2.2 Related work

2.2.1 Computational humour

Although AI research has mostly overlooked humour, one can argue that this complex cognitive phenomenon, which is unique to humans, is a vital aspect of creating computer systems that are genuinely intelligent [43]. Aside from explaining theories of humour, AI investigations could trigger significant breakthroughs in human intelligence, language, problem-solving, and information processing [44, 45]. However, researchers have concluded that building a model capable of detecting and generating all types of humour is an AI-complete problem, i.e. a problem that is at least as difficult as solving “general intelligence” or any intellectual activity that a human can tackle [43, 46]. The reason for this is that the vital elements underpinning a sense of humour constitute human thought and understanding,

many of which are still unexplained [47].

Another related issue is that while attempting to explain many different types of humour altogether, most existing theories are too vague and indefinite for computational application [48]. Graeme Ritchie, a linguist and AI researcher, strongly encouraged researchers to pinpoint explicit types of humour, which they can exhaustively define and compute, before building one sweeping theory of humour. In line with this, many researchers, including Ritchie, have narrowed their focus, for example, to distinct types of jokes when tackling computational approaches to humour [4].

Generating humour

In theory, one could try to design a program that processes input verbal texts and detects whether they are humorous [4]. However, a more practical starting point suggested by Ritchie in the early 1990s would be to develop a program that generates humour by exploiting a specific underlying theory. Humans could then judge whether the program output is humorous, and observing its behaviour would provide insights into flaws of the underlying theory, leading to its refinements and corresponding program enhancements.

Template-based approaches

Binsted and Ritchie took this approach to build one of the earliest automated humour generation systems called JAPE (Joke Analysis and Production Engine), which used hand-coded templates to create a specific class of jokes known as punning riddles [44]. They devised a formal model of the underlying punning mechanisms and determined symbolic rules for the associated meaning combinations and textual forms. These rules, along with an extensive lexicon (dictionary), were given to the program, which queries the lexicon for suitable word pairs (that match the rules) and applies the riddle templates, allowing it to output countless novel riddles. Their quantitative evaluation revealed that some of the “most comprehensible” JAPE-generated jokes matched the humour evaluations of the published human-generated jokes.

Lexicon-based methods have also been used for other joke types, such as humorous acronym generation, for example, in the HAHAcronym project, which produced humorous reinterpretations of an acronym, and suggested new acronyms for provided keywords [46]. Although, these discussed methods could not distinguish output jokes from their relative humour quality. An approach to overcome this involved assigning weightings to specific generation rules in order to rank output using n-gram frequencies or word similarity metrics. This approach was adopted by a humorous analogy generator developed by Petrović and Matthews in 2013 [49]. The generator contained encoded assumptions, which favour distinct nouns for the setup, and unique, fitting and ambiguous adjectives as a punchline. The generator

then filled the first template slot with values from a human-crafted analogy joke dataset and sought distinct nouns and related adjectives for the remaining template slots. This system was reasonably successful; its generations were classed as humorous 16% of the time, while human-produced jokes (with the same templates) received 33% positive classification.

Machine learning approach

The Generalized Analogy Generator (GAG), a development of the previous approach, used machine learning techniques on human-rated jokes, i.e. the training data, to produce jokes via a “generate-and-test” method [47]. GAG generated analogies subject to the detection model finding the joke adequately humorous. Although the model’s frequency of generated humorous jokes remained the same, it required fewer assumptions about metrics, i.e. what forms a more humorous analogy joke and instead learns optimal combinations of more extensive metrics, allowing it to adapt to user preferences.

Deep learning approaches

Although the arrival of large pre-trained transformer-based language models, such as BERT, GPT-2, GPT-3, etc., has facilitated the state-of-the-art across several NLP tasks, these models are more limited in handling self-contained jokes and puns [47]. One drawback faced by these models for wordplay is the use of token encodings for text, which causes the model to lose knowledge about the input’s exact letters, often hindering the generation (or detection) of novel wordplay [50]. After fine-tuning GPT-2 on a jokes dataset, the model generates primarily illogical jokes, similar to those children make, that correctly format the joke but not the punchline. Similarly, GPT-3 demonstrates it can successfully recognise patterns among provided jokes but cannot generate humorous text [51].

While existing research has demonstrated steady progress in computerised humour generation, there have not yet been any substantial breakthroughs [47]. One of the reasons for this is the scarcity of annotated and clearly-defined public datasets. Although some datasets are available, including edited humorous headlines and other joke types, descriptive metrics for most large joke datasets are lacking. Also, the lack of formal evaluation methods for output jokes marks another issue faced by humour generation systems [47]. It prevents external assessment of the quality of output jokes [52] and humour systems from understanding and verifying the output humour [53]. As such, automatic humour detection techniques could enhance the current approaches to humour generation.

Detecting humour

While the detection of jokes comes naturally to humans, one could argue that automatic humour detection is more complex than humour generation since, although

generating humour focuses on limited subtypes, recognising humour requires a much greater coverage of its numerous types [47]. For the most part, humour detection has been formulated as a binary classification problem, achieved by two main approaches, as seen in existing research.

1. Feature engineering

Confining to distinct types of jokes has enabled researchers to use NLP techniques to hand-code patterns and features for them, which they can combine with simple machine learning models to detect humour [47]. We refer to this approach as feature engineering.

The type of humour found in one-liners is particularly appropriate for automatic learning since, by definition, they produce a humorous effect “in one shot” in just a couple of words, which restricts their humour-generating features to one sentence [54]. As a result, the research on recognising humour within one-liners is comprehensive [54, 55]. As an initial approach to humour detection, Mihalcea and Strapparava defined stylistic features of humour, i.e. alliteration, antonymy and adult slang, and integrated them with content-based features extracted from a well-known text classifier, i.e. SVM, into a stacked machine learning framework [54]. Their model impressively differentiated between one-liners and news with 97% accuracy and between one-liners and standard text with 79% accuracy. Although seemingly significant, it is unclear whether the model is genuinely detecting humour or confusing the style and comprising words of the jokes with their humour-causing features [56, 57]. Since the positive and negative samples come from separate domains, it is likely that the model discovered specific words, which often appear in jokes without capturing the finer nuances of humorous texts.

There are also several existing studies on humour detection amongst larger structures, i.e. news articles or blogs, which is arguably a more challenging task due to the relatively complex narrative form of the longer jokes found in these structures. Mihalcea and Pulman demonstrated the dominance of human-centeredness and negative polarity in humour recognition within articles [58]. In contrast, Reyes et al. concluded that semantic ambiguity and “keyness” (actual frequency of a word compared to its expected frequency) are more influential characteristics of humorous blogs after comparing their performance with existing features using Naive Bayes and Logistic Regression classifiers [20]. Another study conducted on blogs investigated the significance of affectiveness information, i.e. information relating to emotions, sentiments or moods, in humour detection [59]. A sample of features was evaluated in a classification task, including polarity, semantic ambiguity and affectiveness. Despite achieving a low accuracy overall, the relevance of affective information in discriminating the data sets according to the emotions, attitudes, etc.,

of the bloggers can be clearly seen, with accuracies of models considering both semantic ambiguity and polarity falling short of reaching 30% while considering affective information raises the accuracy by 10%. Experimenting with more extensive structures consisting of longer jokes introduces the issue that automatic classifiers can learn to differentiate between humorous and non-humorous instances based merely on text length or apparent distinctions in their vocabulary [58]. A way around this would be to provide classifiers with negative instances that are, for the most part, similar to the positive instances, however different in their humorous effect, in order to enforce the classifiers to identify significant features responsible for generating humour.

2. Deep learning

Currently, NLP is one of the most exciting fields in AI, which, combined with recent advancements in deep learning, has given rise to state-of-the-art techniques for humour detection [150]. We now explore several models that leverage these advancements for detecting humour, such as recurrent neural networks or transformers in the place of feature engineering.

LSTM-based networks

In 2016 Bertero and Fung trained an LSTM-based framework as a novel approach to recognising humour in dialogues [60]. When constructing their training dataset, they extracted positive and negative instances from one television show script, which prevents their framework from merely capturing domain differences instead of the deeper semantic structures of humour. They also integrated a high-level feature set from previous sentences to use as a context for their model, which outperformed their CRF baselines by 8% when comparing their F1 scores in identifying humorous content. The model used the show’s “canned laughter” (or laugh track, a separately recorded soundtrack that consists of the sound of the audience laughing) [61] to reveal humorous parts of the dialogue, which, although significant for its construction, can also be a liability to its accuracy. For example, when producers insert canned laughter into the show to attempt to seek a reaction to inadequate jokes or to keep their audience constantly amused, even during scenes consisting of minimal humour [60].

An LSTM-based humour recognition model trained on over 350 psychotherapy session exchanges was introduced by Ramakrishna et al. in 2018 [62]. They labelled these exchanges according to the Motivational Interviewing Skill Code (MISC) [63] and cues such as humour-based laughter. The model outperformed a standard baseline with linguistic humour features with two variants; one demonstrated the use of context for improved humour recognition performance, and the other revealed that using task-specific embeddings

enables higher accuracies in the absence of context.

Multimodal datasets, such as the UR-FUNNY dataset that includes audio and video in addition to text, can be leveraged to assist models in grasping the context of humour, i.e. through prosodic features (e.g. accent, stress, rhythm, etc.) and motions [64]. The dataset, formed using laughter markers in TED talks as cues to indicate where the dialogue was leading up to a punchline, sourced its positive and negative instances from the same domain, which also prevented the issue of the model detecting differences in the domain rather than humour. In 2019 the developers of the dataset trained multimodal neural networks on the set and showed that adding audio and visual cues can improve the accuracy of humour detection (by over 6%). Their results also demonstrated that while the latest models can recognise humour reasonably well (with 65.23% accuracy), there is still a significant gap compared to human performance (82.5% accuracy), which reveals scope for further research in detecting humour and the analysis of multimodal language.

Convolution Neural Networks

Computational humour research has generally centred on jokes in English due to difficulties in translating humour [65, 66], i.e. one cannot simply translate the data word by word to build humour detection models in new languages. However, there are still some approaches to humour detection in different languages, namely in Chinese, such as Chen and Soo’s CNN, which has an extensive filter size and number, and Highway Networks, which increase network depth in order to detect humour in puns and one-liners in both English and Chinese [67]. Their proposed model outperforms existing feature and neural-based models, reaching as high as 90% accuracy on certain datasets in each language. For an unbiased evaluation, they curated a balanced dataset and tried to minimise domain difference by choosing negative instances whose words are all found in the positive instances and ensuring positive and negative instances have a similar average text length, in line with previous works [68].

Transformer models

Although deep learning models excel at learning from numerous labelled samples, as seen, they tend not to generalise well on unseen training conditions [69]. In addition, collecting a large amount of data when tackling an entirely new task can be very challenging [70]. These issues have given rise to new techniques such as model fine-tuning [71], which several researchers have exploited, applying pre-trained models to classification tasks in detecting humorous text. Existing research clearly supports that large-scale transformer models, such as BERT, considerably outperform the previous approaches to

humour detection, with some reaching accuracies over 95% on their curated datasets [72, 57]. A previous approach to humour detection demonstrated this by comparing the humour recognition performance of models when applied to a dataset consisting of jokes and their equivalent generated non-jokes, which included the same words and structures [47]. In this approach, typical neural models like LSTMs and CNNs could not distinguish these texts, while a fine-tuned RoBERTa model distinguished between both types with nearly 90% accuracy [57].

Several researchers have exploited the recent advances in nlp and deep learning by applying pre-trained models to classification tasks in detecting humorous text. Existing research clearly supports that large-scale transformer models, such as BERT, considerably outperform the previous approaches to humour detection, with some reaching accuracies over 95% on their curated datasets [72, 57]. A previous approach to humour detection demonstrated this by comparing the humour recognition performance of models when applied to a dataset consisting of jokes and their equivalent generated non-jokes, which included the same words and structures [47]. In this approach, typical neural models like LSTMs and CNNs could not distinguish these texts, while a fine-tuned RoBERTa model distinguished between both types with nearly 90% accuracy [57].

In an attempt to learn some of the subtleties in humour detection, researchers have released labelled sets of pairs of original and their corresponding humorous headlines [73, 56, 74] and set up competitions based on these datasets [73]. SemEval-2020, a popular shared computational humour competition, proposed the task of predicting the humour scores of edited headlines and identifying the funnier headlines [74]. Unsurprisingly, the highest ranking teams for each objective made use of large pre-trained language models, such as BERT, RoBERTa, ELMo, etc. [75]. Curating their dataset involved humans editing one word within a headline to make it humorous. The minor edits allowed models to understand which word types determined the humour level of texts and even compare the extent of humour delivered from different edits on a single headline. We will see that this approach is also applied to the construction of our humour recognition model dataset in Section 3.3.1.

As we have seen, although feature engineering was initially the dominant approach for automatic humour detection, recent deep learning techniques have demonstrated much better recognition performance than meticulously crafted feature detectors, where transformer models have significantly taken the lead.

2.2.2 Chatbots and humour

Chatbots are computer programs that mimic and process conversation with human users in written or spoken natural language [76]. They tend to either operate via

machine learning (i.e. AI-based chatbots) or set rules (i.e. rule-based chatbots) [77].

We find chatbots in numerous domains, including retail, travel, and healthcare, to name a few [78], but there is a lack of existing research on chatbots for learning to laugh and develop a sense of humour. Therefore, to determine a suitable architecture for our chatbot and any features we should exploit from existing approaches, we turn to a closely related field, mental health.

Chatbots for mental support

The inherent privacy, anonymity and international accessibility of chatbots enable them to be a valuable support to individuals in dealing with the delicate issue of their mental health [79]. There is extensive research on their use in mental health domains, and several chatbots exist in this field [80, 81, 82, 83].

A recent systematic review brought together the evidence from over 50 studies regarding the application of chatbots in mental health, and over 90% were rule-based [84]. These rule-based approaches use decision trees to design their conversation flow, whereby user input, free text or a choice from set options, determines transitions between fixed states [85, 86, 87]. Here, the chatbot's statements are either determined by these transitions within the existing tree structure or entirely predetermined.

Although this rule-based approach to dialogue management is very controlled and avoids the possibility of the chatbot's statements upsetting the user, which could even be detrimental to their mental state, it can cause the user to perceive the conversation as repetitive or less engaging. A proposed solution was to retrieve each chatbot statement from an extensive pool of existing statements [83]. This retrieval-based method has also been adopted by the SAT chatbot [6].

Humorous chatbots

The importance of humour

A survey conducted in 2017 on humour in human-computer interaction, a field that, until recently, has received little attention, describes the design and interaction of humour in virtual agents [88].

Virtual agents, in essence, attempt to simulate human interaction (HCI) in their image, language and expression. This “human-likeness” can, on the one hand, convey intimacy, yet on the other hand, it could render feelings of discontent and strangeness. An interesting mitigation for this, proposed in the survey, could be to lower the user's expectations, reduce the human resemblance of virtual agents and enhance the user's perceptions of them by easing the tension in their conversation. The hypothesis was that humour might cause a significant shift in viewpoint,

enabling users to laugh at these imperfections of technology and, in this way, accept them. There exists research in favour of this idea, including several studies [197, 198], which indicate that humour improves the likeliness of an agent and enhances users' enjoyment in their interaction with one.

An early study [89] showed that users evaluated a humorous virtual agent as more human-like and thus rated it as more likeable and funny. Another study [192] also revealed that social conversations rose to 50% when including jokes in a virtual receptionist's interaction with users. A particularly significant finding is the benefit of humorous agents in inducing a change in user behaviour due to their fun experience in their interaction [88]. This finding can be important in our context of helping a user learn to laugh, where encouraging change within them to shift towards a more mindset is often necessary.

The authors of the humour and HCI survey concluded that the use of humour should rise to a much larger scale in human interaction with virtual agents for several purposes. Such purposes include learning tasks, i.e. to help users gain a better long-term understanding of content, encourage them to follow a goal and transform their behaviour. Therefore, the benefits of humour in helping a user learn to laugh are clear. As such, we chose to augment our chatbot's primarily rule-based structure with a computational platform that incorporates humour into its conversation.

Associated risks and good practices

As we have seen, computing humour is a significant challenge in computer science due to its intrinsic ambiguity and nuances [88]. Also, in our temperamental digital world, where even one social media post could take down a whole brand, it's natural to be fearful of developing a humorous chatbot. On top of this are the discussed risks of humour in psychotherapy [4]. The idea of accidentally offending users puts chatbot designers off the idea of including humour in their platforms. However, we can follow several recommended methods for successfully deploying humour in such autonomous systems. These include accurate detection and generation of humour, which we discussed in the previous section. A concrete understanding of a chatbot and its users is essential for delivering authentic humour that will appeal to the user [90]. Another vital consideration is choosing an appropriate time for the chatbot's delivery of humour, a particularly challenging task requiring inherently human characteristics, such as emotional intelligence, a knowledge base, context and a grasp of culture [88].

Examples of humorous chatbots

An important lesson we learn from existing chatbots is not to overdo the humour, especially since, by nature, it is ambiguous and involves a lot of judgment [4]. An example is Cleverbot, a well-known AI-based chatbot created in 1997 to enable users to "have fun" [91, 92]. Existing research, including articles and a case study

investigating its performance, consisting of several sample conversations, indicate that while it can be amusing at times, it quite quickly “gets dull” [93, 92]. Often it is more appropriate and user-friendly if the chatbot delivers light-hearted humour rather than explicit jokes when responding to the user.

Humorist Bot was one of the earliest conversational agents, developed in 2008, capable of having a sense of humour that can generate humorous statements and recognise humour in the user’s statements [94]. It also includes an avatar, which adjusts its visual expression depending on how it perceives the user’s humour. Due to its increased nuances and user impact, more emphasis was placed on humour recognition. Automatic techniques were applied for this task, i.e. detecting distinct humoristic textual features, including alliteration, antonymy and adult slang features, as seen in the previous section. Although the performance is sufficient for the chatbot to grasp the primary humour element of a conversation and react accordingly, there is room for improvement, as stated by the authors. A more extensive textual analysis, i.e. searching for a wider range of features, or even applying deep learning techniques, could improve the agent’s humour recognition capabilities.

A 2020 research initiative, encouraged by a recent surge in Singapore’s elderly population, was conducted to assess how humour can influence senior citizens’ knowledge retention while interacting with a humorous personal companion chatbot [95]. The authors wished to investigate whether they could reproduce the known positive effects of humour on memory [96] in the more complex AI domain [97]. Results indicated that humour, i.e. in puns and wordplay, assists individuals in recognising information previously introduced to them. There is also some evidence to support that this humour improves our capacity to recall information. Due to the increasing demand for products that can enhance health and well-being among the elderly and in our world faced with an ageing population, this research has multiple applications. It also welcomes and encourages the design of more intelligent frameworks for wider target groups and objectives.

Woebot, an AI-powered mental health chatbot, has proven to have significant use in supporting people to reduce their anxiety and depression [98]. When asked what sets Woebot apart from its crowd, the product manager revealed that “it’s a joke”, quite literally. I.e. incorporating humour into the bot retains its users and earns their trust, especially when recently, more people have turned to the chatbot to ease their concerns regarding the COVID-19 pandemic [99]. Woebot integrates humour into cognitive behavioural therapy (CBT) to offer users support, which helps them open up. Since cognitive behavioural therapy is “really tough”, an element of humour can be a pleasant break for the users, shared the manager. Woebot also stresses that it’s not a human therapist sitting behind the user’s phone screen, which also helps gain their trust.

While there has been considerable progress in recognising

[100], understanding [101], generating [1], and properly delivering humour [222], the creation of entirely automated humorous agents that are able to perform these tasks very well is “still in its infancy” [88]. And given the current difficulty of artificial machines, even to comprehend natural language semantics, mastering humour would mark tremendous technological progress.

Additional chatbot considerations

As discussed, we require that our chatbot is light-hearted and humorous in its demeanour to engage and instil trust in the user while remaining polite and self-effacing to avoid offending or upsetting the user.

In addition, for our chatbot to establish a healthy relationship with the user, it must display empathy in its responses, and extensive research indicates the importance of this characteristic, especially in psychotherapy [6, 102]. Therefore, we chose to employ the empathy detection platform used by the previous SATbot framework on our chatbot’s responses so that it displays empathy while conversing with the user.

2.3 Technical preliminaries

We now cover the technical background for the project, which includes several natural language processing (NLP) tools and techniques.

2.3.1 N-gram language models

Language models (or LMs) are models that assign probabilities to sequences of words [103].

Here, we introduce the simplest statistical language model, the n-gram model (or n-gram for short), which assigns probabilities to sentences and sequences of n words.

We start with the equation $P(w|h)$, which refers to the probability of a word w, given a history of words, h. For example, $P(\text{park}|\text{he is going to the})$ [104].

We estimate this probability function using relative frequencies, which involves taking a large corpus (group of texts to search through), counting the occurrences of the phrase “he is going to the”, and finally, counting the occurrences of this phrase when followed by the word “park”. Essentially, we are counting: out of the times that we see a history h, how many times does the word w follow it? We can also interpret this as calculating the joint probability: $P(\text{he is going to the park})$, where, of course, a higher probability indicates that a model can be more certain about its prediction of a sentence.

Although N-gram models are considerably more straightforward than state-of-the-art neural language models based on the RNNs and transformers that we later introduce, they are a fundamental tool for understanding the notions of language modelling.

2.3.2 Tf-idf

Tf-idf is a statistical measure that calculates the relevance of a word to a document among a collection of documents. We determine the tf-idf weighting by multiplying the term frequency (tf) by its inverse document frequency weight (idf) [104].

Machine learning can vastly improve the performance of standard NLP methods (e.g. using word counts) for applications such as text classification, in our case. But this requires first converting text into numbers in a process called text vectorisation that we can achieve by calculating its tf-idf weighting, which we can then feed into our chosen classification algorithm [105].

2.3.3 Classification algorithms

Numerous language processing tasks involve classification. We briefly introduce three supervised learning algorithms in machine learning that classify data: naïve Bayes, logistic regression and support vector machines. We will use these to determine a baseline for our humour recognition model.

Naive Bayes is a probabilistic machine learning algorithm based on Baye's theorem, which allows us to calculate the probability of an event given that another event has occurred [104]. We use the multinomial naive Bayes classifier for our purpose of classification with discrete features, i.e. word counts for text classification (for which we use tf-idf) [106, 107].

Logistic regression is a linear model used to predict a categorical dependent variable using a given set of independent variables [108].

A support vector machine (SVM) is a linear model that creates a decision boundary, which divides n-dimensional space into classes to classify future data points [109].

2.3.4 Neural language models

As described in Section 2.3.1, language modelling involves predicting upcoming words given a prior context. We introduce neural language models (or neural models) relying on neural networks, a fundamental computational tool for language processing, which have many advantages over the previous statistical n-gram language models [104]. These include handling considerably longer histories and generalising better over similar word contexts. However, their inherent complexity and lack of interpretability make n-gram models a more suitable choice for smaller

tasks. An additional downside of neural models is their long inference times. While these are relatively negligible when running inference on a single sample or a very small batch, they become very significant on large batches. However, we can account for this as explained in Section 3.3.2.

Modern neural models use more powerful deep learning architectures, such as recurrent neural networks or transformer networks, which much better account for the temporal nature of natural language (e.g. in the flow of conversations or jokes, in our case) compared to the simplest neural models, i.e. feedforward networks.

We now briefly describe some of these network architectures to provide a context for existing research and our chosen model.

Recurrent neural networks (RNNs) contain a cycle inside their network connections, meaning that each unit is dependent on its own previous output [104]. They can therefore accept variable length inputs, making them suited to modelling sequential data, e.g. sentences in natural language [110].

Convolutional neural networks (CNNs) are often used for classification tasks; in particular, they deliver a more scalable approach (not limited) to image classification, exploiting linear algebra, i.e. matrix multiplication for convolutions, which identify patterns within input data using filters [111, 112].

The Long short-term memory (LSTM) network is a commonly used extension to the RNN, which solves several of its issues, including its inability to retain information over long periods, along with some of its training issues (i.e. with vanishing or exploding gradients) [113, 114].

Transformers are an approach to sequence processing that addresses several inherent issues of RNNs by eliminating recurrent connections and opting for architectures similar to the simple feedforward networks, hence, alleviating difficulties in training, information loss and the inability to do computation in parallel [115]. They capture meaning through learning context by keeping track of relationships in sequential data, such as words in a sentence. Transformer models use an evolving group of techniques called attention (or self-attention) to spot subtle patterns between distant data elements in a series, including how these elements impact and depend on one another. Transformer-based language models are pre-trained on extensive unlabelled text corpora in order to learn a language representation. As we will see, we can also fine-tune these transformer networks on smaller datasets by including an additional output layer to apply them to specific tasks. Fine-tuning networks significantly improves their performance compared to those that have been trained only on a labelled corpus [116].

BERT (Bidirectional Encoder Representations from Transformers) is an open source state-of-the-art language model developed to allow computers to grasp the

meaning, including subtleties, of vague language in any text by using its surrounding text to provide background [117]. One of BERT’s significant innovations is using the transformer’s bidirectional training for language modelling, in contrast to approaches in the past. Employing this bidirectional capacity, BERT is pre-trained on Masked Language Modeling by masking random words from the input and attempting to predict them given the remaining unmasked input as context [118].

RoBERTa is a transformer-based language model that seeks to optimise and enhance the original BERT model. It modifies some of BERT’s key hyperparameters by omitting its next-sentence pretraining objective and training with larger mini-batches and learning rates enabling RoBERTa to enhance the original masked language modelling objective and leading to improved downstream task performance [119].

2.3.5 Evaluating classifiers

In order to understand how well our classifier is performing, we need an evaluation metric. Often a natural choice for a metric is the accuracy score, which measures the percentage of all the observations that a model has labelled correctly. However, we generally don’t use it for text classification tasks since it does not work well when the classes are unbalanced (which is often the case in practice) [104]. Therefore, we turn to two other metrics: precision and recall [104, 120].

Precision measures the percentage of the items that the system detected (i.e., the system labelled as positive) that are indeed positive, and we define it as:
 $\text{Precision} = \text{true positives}/(\text{true positives} + \text{false positives})$

Recall measures the percentage of items present in the input that the model correctly discovered, and we define it as:

$$\text{Recall} = \text{true positives}/(\text{true positives} + \text{false negatives})$$

We define the simplest single metric that incorporates the precision and recall scores called the **F-measure** [104]: $F_\beta = (\beta^2 + 1)PR/(\beta^2P + R)$, where the β parameter acts as a weighting for recall and precision, and is determined based on the model application. If $\beta = 1$, precision and recall are equally weighted, and we frequently use this metric, called $F_\beta=1$ (or F1)

Chapter 3

Design and implementation

The implementation code is publicly available at <https://github.com/zbp18/HaHA.git>, and the platform can be accessed online at <http://35.176.235.166/> until 20th July 2022. We illustrate three example conversations between HaHA and a human user in Appendix C.

3.1 Conversation flow

We now describe how our chatbot converses with users, including its questions and responses, and outline its underlying structure or conversation flow.

Since the chatbot aims to guide the user through the self-initiated humour protocols, the initial steps in developing the conversation flow involved gaining a thorough understanding of these protocols by reading through them and discussing them with Professor Edalat. In addition, daily practice of these protocols was undertaken for at least a week, as suggested by a training psychiatrist, to obtain deep insights and a genuine feeling for them.

An overview of the chatbot conversation flows considering a user’s positive or negative emotional state are illustrated in Figure 3.1 and 3.2, respectively.

Once the user logs into the platform, the chatbot greets them and asks for their name (so the chatbot can refer to the user personally at several points throughout the conversation). Before starting the session, the chatbot reminds the user of a couple of logistics, such as accessing the self-initiated protocols, restarting the conversation and some quick tips on how they can make the most of their session.

At the start of the session the user is asked by fHaHA how they are feeling. If they display a **positive emotion** (i.e. happy/content), they are encouraged to start a “mini session”, which involves exploring one or more of the self-initiated humour protocols with the chatbot. In these mini sessions, HaHA guides the user on the relevant protocol(s) and asks them about any previous attempts and their feedback or initial attitude towards them. These sessions aim to engage the user in

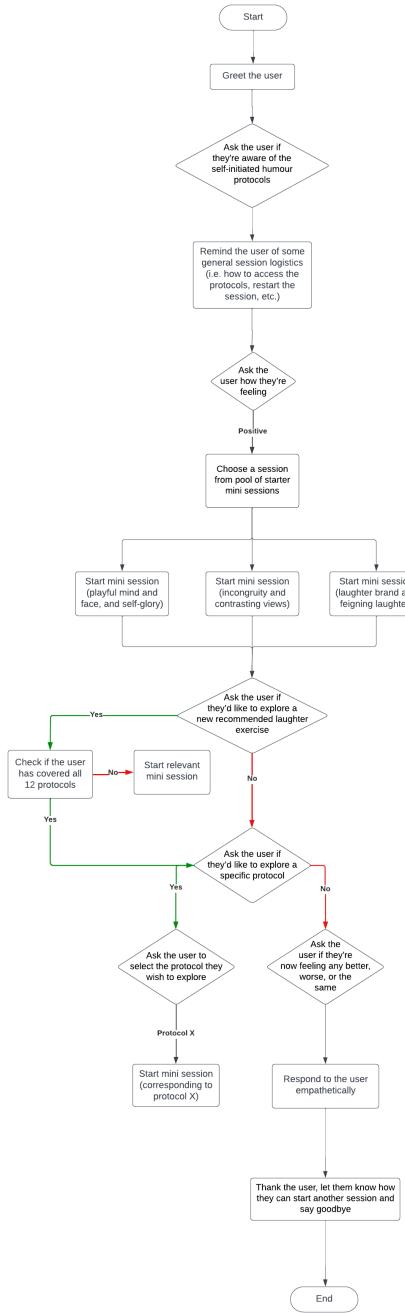


Figure 3.1: Conversation flow (emotional context: positive)

practising the self-initiated humour protocols, encouraging them to naturally apply these exercises to their daily lives over a continual period to enable them to develop a sense of humour [5]. We do not urge the user to practice these protocols *in situ*, as this can sometimes break the conversation flow, as suggested by users of the previous psychotherapy chatbot [6], and these specific exercises are most effective when practised in the context of daily life. Further details on mini sessions and the scenarios in which they are initiated can be found in Appendix D.

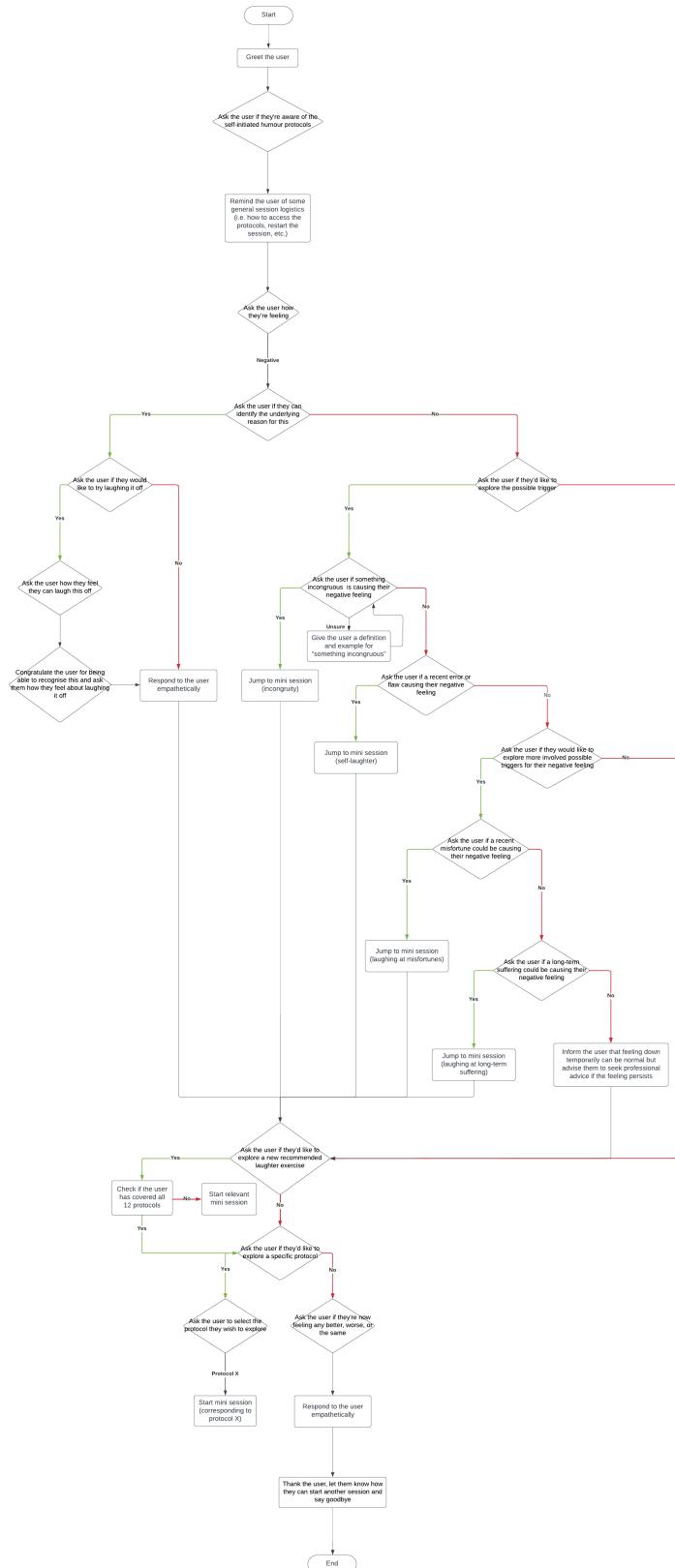


Figure 3.2: Conversation flow (emotional context: negative)

If the user displays a **negative emotion** (i.e. sad, angry or scared/anxious), they are encouraged (if willing) to determine the underlying reason. The chatbot can assist them with this if needed, and then it will suggest a relevant mini session to help the user get past their negative emotion.

In both the positive and negative cases, once a mini session is complete, the user is asked whether they wish to start another mini session recommended by the chatbot. If so, the chatbot recommends a relevant unseen mini session to the user (depending on their current state, see Section 3.4.1), and the user has the choice to explore it further. They can continue completing these mini sessions with the chatbot for as long as they wish. Otherwise, the user can choose to explore a specific protocol (to revisit or discover) with the chatbot or terminate the session.

If the user wishes to terminate the session, the conversation will stop and can be restarted by reloading the page.

Before settling on a final flow, some user research was also carried out by sharing these protocols with three volunteers (1 male and female, both in the 18-24 age category, and a female in the 50-59 age category) to get a better idea of how the public would benefit from them. We asked for the volunteers' feedback after they had practised the protocols for a couple of days, and the main findings are summarised here:

- Volunteers liked the idea of the protocols. They felt that they were “uplifting”, and after completing them, they reported that “it feels good” and that they “had a bit more perspective on things”.
- One volunteer found protocol (iv), incongruous world, more difficult and felt they needed a bit more direction and focus when attempting to complete it as it was sometimes “hard to think of an example on the spot”.
- They felt that more examples for the majority of the protocols would be helpful.
- Two volunteers were particularly intrigued by protocol (vii), contrasting views and wanted to learn more about this.
- They enjoyed protocols (i) and (ii) and felt that they were able to practice these with ease.
- They found protocols (xi) and (xii) challenging yet the “most meaningful” out of the set.
- Overall, the volunteers found the concept of self-initiated humour very interesting and enjoyed trying out the protocols, though, at times, they felt that, in the current format, the protocols were “a little hard to digest”.
- They also found several terms on the protocols document quite technical and weren’t clear on their exact definitions, such as “feigning” or “Duchenne laughter”.

In line with our findings, we decided to group the protocols according to overlaps in their theme and difficulty or level of prior experience required. The full grouping can be found in Appendix A. We used these groups to determine the structure of the

mini sessions.

Considering all of the above, we decided to use the following information when recommending the next mini session to explore with the user:

- their current emotional state
- their feelings towards the protocol discussed in the current mini session (either when they tried it themselves or about the idea of trying it in the future)
- the difficulty of or experience ‘level’ required by the protocol associated with the current mini session

Further details on how we used this information can be found in Section 3.4.1

A significant challenge faced during this stage was developing conversation flows for specific mini sessions. These flows are highly dependent on the protocol(s) associated with the mini sessions in order to best support users with and guide them through the laughter exercises, which inevitably means that some sessions are more involved than others.

Although each mini session is different, we identified overlapping aspects within several of these, allowing us to follow a unified approach to designing the conversation in some cases. In addition, this approach aids the scalability of our conversation flow, such as when we need to add more or modify existing session flows.

The conversation flows of two example mini sessions (one starter session and one more involved session) are shown in Appendix D.

3.2 Dataset

The conversation flow of our retrieval-based chatbot is augmented via a dataset of predefined statements from which it selects its responses to the user. We now describe the process of preparing this and provide a brief analysis of our final dataset, HaHASStatements.

3.2.1 Requesting rewritings

We developed a set of base statements for HaHA corresponding to each point in the conversation flow. There are 122 base statements altogether.

We then prepared five surveys to request rewritings for most of these statements (102 out of 122) via crowdsourcing. In line with our aim to make conversations with the user engaging, empathetic and humorous, we requested rewritings from all surveys to be polite, empathetic and fluent with the optional inclusion of light-hearted humour if deemed appropriate by the participants. We provided example rewritings to guide the participants and allowed them to add text and

change the words as much as they liked. Participants were free to be imaginative, provided that the meaning of the original statements remained or that their input directly responded to the users' statements.

In forming these surveys, we grouped the statements according to overlapping themes enabling them to be rewritten similarly, i.e. by rewriting a base statement or providing a relevant example response given a scenario. We also split these statements according to whether or not we felt it was appropriate to include explicit humour to reduce the risk of accidentally offending the user. For example, we thought it was inappropriate to incorporate a joke within the following statement: "Has there been any long-term difficulty in your life that may be causing this negative feeling?". As such, we placed this statement in the general group. For those statements in the explicit humour group, in addition to our prior rewriting instructions, we requested that participants add humour to the rewritten statements as a joke or light-hearted statement.

Overall, the five surveys consist of two explicit humour rewriting surveys, two general rewriting surveys and one general scenario-based survey, covering the majority (102) of the base statements. We generated these surveys using Qualtrics [121] and refer to them as humour 1, humour 2, general 1, general 2, and scenario, respectively, for reporting purposes. They are accessible online (until August 2022) via the links provided in Appendix B.

3.2.2 Response collection

Once the surveys had been prepared, participants were recruited via a crowdsourcing platform, Prolific. We chose to distribute our surveys evenly to male and female participants aged over 18 years. Surveys responses were accepted subject to proper (i.e. non-copy-pasted, acceptable English) competition of at least half of the rewritings and small adjustments to the data, i.e. its spelling, punctuation and grammar, were made to improve the fluency in some cases.

After receiving survey responses, we carefully inspected each to ensure they were safe (i.e. non-offensive and polite) enough for our domain. The distinction in the quality of responses from the humour rewritings surveys and the remaining surveys was apparent, several of which, unfortunately, had to be discarded. Although the main essence of several of these responses was somewhat humorous, the language used in the delivery of humour was poor, so these had to be adjusted, which was a rather laborious task. The reduced quality of these responses was likely due to the discussed nuances of humour and the continual struggle, even to date, in explaining what makes people laugh [122]. However, we were later able to distribute these surveys amongst volunteers to make up for the discarded responses across each survey.

Altogether, we accepted 54 survey responses from Prolific (corresponding to 1226

Statement group	Avg no. of rewritings	Total no. of rewritings
General 1	27.4	794
General 2	14.6	408
Scenario	14.9	208
Humour 1	13.6	341
Humour 2	15.3	92
Extra	33.2	664

Table 3.1: Average and total number of re-writings per statement group

statement rewritings), with at least eight responses per survey. We found that the response count across the five surveys was fairly unbalanced, with 22 responses gathered from one of the surveys compared to eight from another. We obtained an additional 588 rewritings from volunteers, significantly improving the balance across the surveys. We prioritised obtaining as many rewritings as possible for each base statement rather than achieving a perfect balance across the set. One could argue that achieving a perfect balance would, in fact, be less accurate, as in everyday language, most humans use a different number of variations of specific phrases used, with some phrase variants more commonly being used, such as “how are you” (for which fewer statement rewritings would be needed).

Rewritings for the extra base statements, i.e. those that did not correspond to the five surveys, came from existing datasets [6] (e.g. for obtaining rewritings of “From what you have said I believe you are feeling . Is this correct?”) or the application of data augmentation techniques after receiving the survey responses (e.g. for obtaining rewritings of “But really, how do you feel?” given rewritings for “how do you feel?”). At least 20 rewritings were obtained for each of these extra base statements, with 663 rewritings altogether.

3.2.3 Dataset overview

Our final dataset, HaHAsStatements, consists of at least 12 rewritings for each of the 122 base statements and 2507 statements in total. They are ordered according to the survey (or group) that they are associated with.

The average and total number of rewritings for each statement group in the HaHAsStatements dataset is shown in Table 3.1.

We also collected data on survey participants’ gender and age groups, which is shown in Figure 3.3. We can see that, although the data samples are distributed fairly evenly by gender (69 from females and 52 from males), the samples are skewed towards the right, with most samples originating from the 18-24 age range.

A complete list of base statements (split according to their survey or group) and their

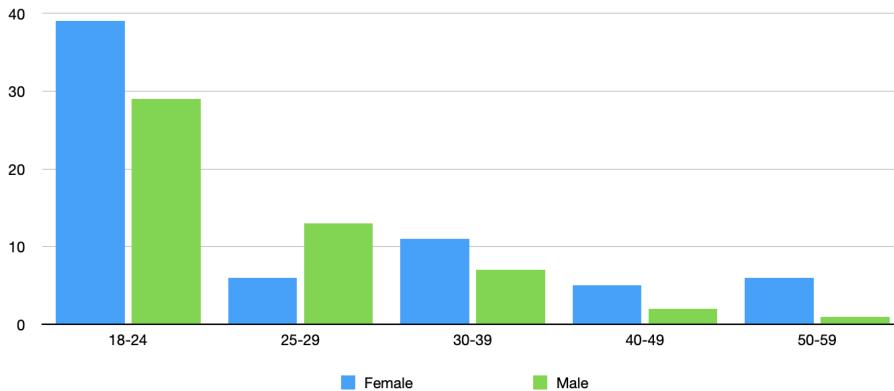


Figure 3.3: Number of data samples by age group and gender

number of rewritings is illustrated in Appendix B.

3.3 Machine learning platform

We augment the rule-based framework with a machine learning platform, which consists of a multi-objective optimisation function that enables our chatbot to retrieve an optimal response from the HaHAsStatements dataset at each point in the conversation. In choosing our final method, we considered previous works demonstrating the successful performance of this retrieval-based approach [83, 6]. We now discuss the development of our platform.

3.3.1 Humour recognition

Motivation for humour annotation

As seen, there are numerous datasets for humour detection consisting of one-liners, news articles, blogs, tweets, and TV show dialogues, to name a few [54, 74, 123, 64]. In addition, there have been several efforts to create new datasets that deal with common issues of the existing datasets, which, along with recent advances in NLP, have led to the development of transformer-based models that have achieved state-of-the-art performance in humour recognition.

In our initial attempt to score the statements from the HaHAsStatements dataset, we turned to some of these existing datasets and transformer-based models. An example is ColBERT, which uses BERT to generate sentence embeddings for texts sent as input to a neural network that predicts the output label and is capable of recognising humour in short texts with impressive accuracy and an F1 score of over 95% [72]. However, even after parameter tuning, when we tested this publicly available model on a subset of the HaHAsStatements dataset, we could not achieve an accuracy of over 55%. Nor was this performance improved after fine-tuning

BERT on several public humour datasets, including one-liners [54], puns [68] and carefully curated datasets made up of Huffington Post news headlines and Reddit jokes [72].

We believe this lack of performance could be due to the domain differences between the existing datasets and the HaHAsStatements dataset, which is centred around a chatbot guiding users on how to learn to laugh while, in several instances, displaying humour itself. Since we were dealing with such a unique domain, finding an exact match was difficult, so we also expanded our search to datasets around therapy. Although, the issue here was that we could not find humorous samples amongst the therapy data. Therefore, not only would this data need to be combined with positive instances to train a humour recognition model, but the model would only learn the domain differences between positive and negative instances without capturing the nuances of what makes the statements humorous. Due to the issues outlined, we decided to extract samples from the HaHAsStatements dataset to prepare a subset that we can use to train our model.

Sampling the dataset

To prepare our subset, which we also refer to as the humour dataset sample, we decided to extract samples from the HaHAsStatements dataset as randomly as possible to mitigate sampling bias, while at the same time, bearing in mind two goals, (1) (most importantly) to achieve a balanced subset with respect to humorous and non-humorous samples and (2) to balance our dataset according to the statements domain (group, i.e. general 1, general 2, etc.). Below, we describe our process of achieving this.

First, before sampling any statements from the HaHAsStatements dataset, we shuffled the data in each survey group, which, when sampling from our shuffled dataset, enabled us to represent more survey participants in our final subset. However, we ensured to keep information on the survey corresponding to each statement to help achieve a balanced subset with respect to the our humorous vs non-humorous statements domain, as discussed below.

There were considerably more statement rewritings from the general rewriting surveys (general 1, general 2) in the HaHAsStatements dataset since these surveys covered more base statements (57 altogether) compared to the humour-based (humour 1, humour 2) (31 altogether) and general scenario-based surveys (scenario) (14 altogether). As a result, every statement from the scenario-based and both humour-based surveys was added to the humour dataset sample. Also, there were more rewritten statements from the first general-based survey (general 1) than the second in the HaHAsStatements dataset, so we also included every statement associated with survey 2 in our final subset. To decide how many statements to extract (at random) from the remaining survey (general 1), we considered our primary aim of creating a balanced dataset across humorous and non-humorous

instances and our aim of a roughly even split across the different surveys. We approximated, from earlier scrutinisation of each response, that almost half of both general survey responses also incorporated light-hearted humour, which would therefore contribute to the humorous instances within our dataset. As a result, in our attempt to achieve a balanced dataset, we extracted just under double the number of samples from general surveys compared to humour-based surveys.

We also included 53 samples from the rewritings of the extra base statements (i.e. those not associated with one of the five surveys obtained via data augmentation or previous datasets). We based this choice on the fact that most of these are augmentations of previous statements, which we already accounted for in building our final dataset.

It is important to note that in forming the HaHAsStatements dataset, we separated out the base statements for which we explicitly requested humorous rewritings as a safety measure (via both humorous surveys), which would introduce a domain difference between humorous and non-humorous instances in our subset and may invite bias as a result. However, requesting that rewritings from the non-humorous surveys (general 1, general 2, scenario) were all light-hearted (and humorous if deemed appropriate, ensures that all domains should include several humorous statements, which we found after inspecting our responses. An additional measure taken to reduce the domain difference further was to explicitly add every base statement associated with the humour-based surveys (humour 1, humour 2) to our humour dataset sample. Including these non-humorous base statements would allow the model to learn from linguistically similar positive and negative instances, which differed only in their humour, to help ensure the model can capture its nuances during training.

Although we were able to reduce the domain difference between negative and positive instances in our final subset, there is still the possibility that the decisions of a model trained on this humour dataset sample might be partly based on some spurious correlations. As such, in future implementations, it may be worth extending the dataset to ensure that the number of humorous and non-humorous statements from each group are the same in our final subset.

Our final humour dataset includes 1109 samples from the HaHAsStatements dataset. Since we aimed to experiment with transformer-based models, already pre-trained on 3.3 billion words, in training our humour recognition model, this give us sufficient data to fine-tune the model, as indicated by existing research [124].

To validate the split of our humour dataset across the different surveys, we used the survey class corresponding to each statement and calculated the proportions of the humour dataset sample that came from each survey, which we illustrate in Figure 3.4 (see Appendix B.7 for more detailed survey breakdowns of the final subset). We achieved the following split: 19%, 16%, 18%, 20%, 27% across the surveys: general 1, general 2, scenario, humour 1, humour 2, respectively. As we can see, we

achieved a close to even distribution across the different surveys.

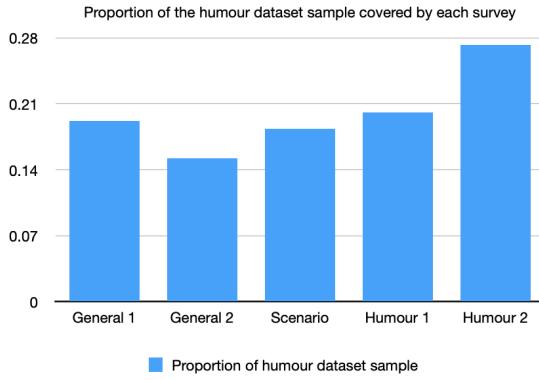


Figure 3.4: Bar chart showing the proportions of the humour dataset sample that came from each survey.

Annotating the dataset sample

After extracting our humour dataset sample from the HaHAsStatements dataset, we needed to annotate this based on the level of humour of its comprising statements to build our humour recognition model. We chose to develop a binary classifier, in line with the existing Colbert dataset [72], in case we decided to extend our humour classification experiments by applying a double fine-tuning to our model using this dataset in combination with ours. In addition, several existing approaches to humour detection have developed binary classification models. Therefore, we could also compare our model performance with these.

For the annotation, we asked three volunteers (all holding a degree) to score the 1100 statements of our dataset for humour using the discrete labels 0 (non-humorous) and 1 (humorous) to avoid introducing a bias by labelling them ourselves. In our attempt to further avoid bias, we aimed to obtain some spread, albeit amongst a small sample, in our volunteers' age and gender, by enlisting one male (aged 18-24) and two females (aged 18-24 and 50-59).

After receiving the annotated statements, we computed the overall humour score for each using the majority label (i.e. if two of the annotators assigned the score 0, and one assigned the score 1, then the final score would be 0). The scores assigned by each annotator can be found in https://github.com/zbp18/HaHA/tree/main/annotated_data.

We recognise that our chosen annotation approach can still introduce bias due to the imbalance in our volunteers' age group and gender, with 2/3 females and 2/3 aged 18-24. An ideal approach would be to crowdsource the task of humour annotation (where we would choose to enlist an even number of males and females under each age group).

Classification model

To build our humour recognition model, we train several classifiers for the binary task of detecting humour in a sequence of text and evaluate their performance.

We first split our humour dataset (a subset of the HaHAsStatements dataset) into train, validation and test sets in 80-5-15 proportions. The chosen split ensured our model had as much training data as possible to perfect the model while leaving out sufficient samples to evaluate our model performance.

Baseline models

To select a humour classification baseline, we used logistic regression, SVM and naive Bayes over tf-idf, and compared their performance. We chose these classifiers as they have proven to perform well in similar classification tasks [6]. We trained our models by combining our train and validation subsets and tested them using our test subset to ensure a fair comparison.

Humour classification experiments

Based on our findings from extensive research in computational humour regarding the success of transformer-based models, we chose to fine-tune a BERT and RoBERTa model for humour classification. For the fine-tuning, we use our humour dataset, which volunteers had annotated for humour.

We performed hyperparameter tuning for each of our experiments using our train and validation sets. For our fine-tuning, we tried the following values (as recommended by the authors of the BERT [118]):

- Batch size: 16, 32
- Learning rate (Adam): 5e-5, 3e-5, 2e-5
- Number of epochs: 2, 3, 4

Finally, we evaluated our models on our humour test set.

We chose (highest performing):

- Batch size: 32
- Learning rate: 5e-5
- Number of epochs: 2

Evaluating classifiers

We include a comparison of the performance of our models in Table B.

Model	Accuracy	Macro-f1
Baselines (tf-idf + logistic regression/SVM)	0.79	0.79
Baseline (tf-idf + naïve Bayes)	0.76	0.76
BERT (fine-tuned on the HaHAsStatements dataset)	0.95	0.95
RoBERTa (fine-tuned on the HaHAsStatements dataset)	0.97	0.97

Table 3.2: Results for the humour classification experiments.

These results are highly promising, especially given the ambiguous nature of humour [4], which we can observe from the frequent misalignment of humour scores assigned to our dataset by the human annotators (accessible via https://github.com/zbp18/HaHA/tree/main/annotated_data). Both our pre-trained language models show a vast improvement over the chosen baseline.

We selected the RoBERTa model fine-tuned on our humour dataset as our final humour recognition model since it achieved the highest performance, with an accuracy and macro-averaged f1 score of 97%. As such, in our optimisation function, this model computes the humour score of a statement and acts as our humour scoring function, H . Note, the function output directly corresponds to its training labels 0 (non-humorous) and 1 (humorous). Therefore it is intrinsically normalised and we refer to the function as \tilde{H} in our final optimisation function.

3.3.2 Optimisation function

Our multi-objective optimisation function consists of a combination of weighted humour, empathy, fluency and novelty scores and has two forms corresponding to a user’s emotional state. We also refer to this as a retrieval function since it retrieves an optimal statement from a set of rewritings given the base statement. We now describe the process of selecting and tuning this function.

Emotion recognition

For the task of emotion recognition, we chose a RoBERTa classifier fine-tuned on two public emotion datasets [6, 125], which have been successful in obtaining over 90% accuracy for similar tasks in the past [6].

The decision to configure our optimisation function based on the user’s emotional state was due to our assumption that if they are feeling negative, our chatbot has a greater need to be empathetic and risk of seeming “too humorous” compared to if the user is feeling positive.

Fluency

We calculate the fluency score of a statement using the inverse of the perplexity (PPL) score yielded by GPT-2, a neural language model with a decoder-only

transformer architecture that has been used successfully for this task in several past experiments [126, 6].

Empathy

To compute the empathy score, we used the T5 language model (encoder-decoder transformer) [122], fine-tuned on the EmpatheticPersonas train set [6]. The model has performed well at empathy classification in a similar domain of psychotherapy, achieving accuracy and macro-averaged f1 scores of over 80.66%.

Novelty

When determining the novelty of a statement, we use a pre-defined NLP-based approach that compares the statement with up to 50 of the chatbot's previous statements, which it stores in memory. This approach uses a function that computes the overlap distance weighted over every n-gram between two sequences of text enabling the selection of statements that are distant from previous statements.

However, in practice, we found this approach led to repetition of statements in several cases. Although the proposed approach successfully retrieves a novel statement with respect to the chatbot's previous statements over one user session, we would also wish to retrieve statement rewritings that are novel over multiple user sessions, i.e. avoid statement repetition between these sessions.

As such, we extended the novelty optimisation of our multi-objective function by introducing randomness when selecting an optimal statement. This involves taking a random sample of 10 statements and feeding this sample (instead of the whole column) into our optimisation function each time the chatbot chooses its statement. This approach significantly reduced the repetitions between user sessions, however, we found this still led to repetitions for those base statements with fewer rewritings, meaning that in these cases, our random sample extraction would not be as effective. In the future, it would be worth extending our dataset for these statements, in particular, to reduce repetitions even further.

Function selection

To formally verify that the statements chosen in practice by our retrieval function would indeed be optimal, we implemented a testing function that applies our optimisation function to each of the base statements in a context similar to that of a user session, only without the interaction of the user. The testing function outputs a list of n successive rewritten statements (in our case, n=20). When analysing these 20 statements for each base statement, we look out for any repetitions or outliers concerning the humour, empathy or fluency of a statement within the list, depending on whether we are testing our optimisation function in the case where the user is feeling positive or negative.

In order to achieve optimal results when selecting statements we performed extensive testing of our multi-objective optimisation function, assigning several different weightings to our scoring functions in both cases (where the user is positive and negative), generating statement lists via our testing function for the base statements in each case.

We outline our final optimisation function O , which we aim to maximise during statement retrieval, below. Note,

- O_{pos} denotes the optimisation function in the case where the user is feeling positive.
- O_{neg} denotes the optimisation function in the case where the user is feeling negative.
- P denotes the emotional state of the, i.e. $P = 1$ if the user is feeling positive and $P = 0$ if the user is feeling negative.
- $\tilde{H}(s)$, $\tilde{E}(s)$, $\tilde{F}(s)$ and $\tilde{N}(s)$ denote the normalised humour, empathy, fluency and novelty scores of a statement s , each of which are values between 0 and 1.

Then, our final function O that we aim to optimise during the retrieval of a new statement is given by

$$O(s) = P * (O_{pos}(s)) + (1 - P) * (O_{neg}(s)), \text{ where}$$

- $O_{pos}(s) = 0.1(\tilde{E}(s)) + 0.4(\tilde{H}(s) + \tilde{F}(s)) + 6(\tilde{N}(s))$
- $O_{neg}(s) = 0.1(\tilde{H}(s)) + 0.4(\tilde{E}(s) + \tilde{F}(s)) + 4(\tilde{N}(s))$

We also include the final testing function output of our optimisation function for a random sample of base statements in Appendix E.

We also consider the computational cost of the optimisation function. The cost of passing one sequence of text into a transformer-based model is $O(n^2)$, where n is the sequence length. To improve the speed of our platform, a significant contributor to the user experience of an interactive chatbot, we have pre-computed the humour, empathy and fluency scores for each statement and stored these with our dataset. Therefore the only part of our optimisation function that requires computing in real-time would be the novelty score, which is based on the previous statements of the chatbot. Calculating the novelty score requires $s \times W \times (W+1)/2$ comparisons per statement, where s is the number of previous statements stored and W is the number of words in the shorter statement in the comparison. Although the complexity of the novelty function increases significantly as the number of previous statements increases, i.e. further along a conversation, in practice, we find that there isn't a

noticeable lag in time required for the optimisation function to retrieve an optimal statement. In this way, pre-computation overcomes a limitation of the previous SAT chatbot [6], which forced it to restrict the sample size when selecting an optimal statement to save time.

3.4 Web application

Our web application integrates the previously discussed components of our chatbot and surfaces them to the user. We now discuss its development, which can be broken down into three main areas.

3.4.1 Back end and database

The application's back end consists of a Flask (Python framework) API with a Python rule-based model, whose base structure we adopted from the initial SAT chatbot implemented by Ghachem [7]. We redesigned this model and implemented several extensions to support our conversation flow.

One extension was building the logic for determining the next mini session to explore with the user, which is summarised in Figure 3.5.

Note that Figure 3.5 assumes that the user has not covered all 12 humour protocols with the chatbot and illustrates the logic used to determine the next mini-session (otherwise, the user would be directed to another part of the flow, as shown in Figures 3.1 and 3.2).

The logic relies on the following session information:

- user's emotional state (positive vs negative)
- current mini-sessions level, in line with the protocol groups information included in Appendix A (0, 1, 2, 3)
- user's attitude towards the protocol currently discussed (positive vs neutral vs negative)

If, when trying to determine the next mini-session, there are "None left" (see Figure 3.5), the chatbot will recommend an unseen mini session at random. However, it is important to note that this is only a soft recommendation, and the session will not be forced, in any way, on the user. Also, especially for the more involved mini-sessions, associated with the challenging protocols, e.g. laughing at misfortunes, the chatbot asks several lead-up questions (that help it determine whether the user is "ready") before encouraging a user to try an exercise.

The rule-based model has been refactored into separate components that handle the questions logic for the different parts of a conversation, such as when executing a mini session or starting/finishing the conversation. This restructure enables a more

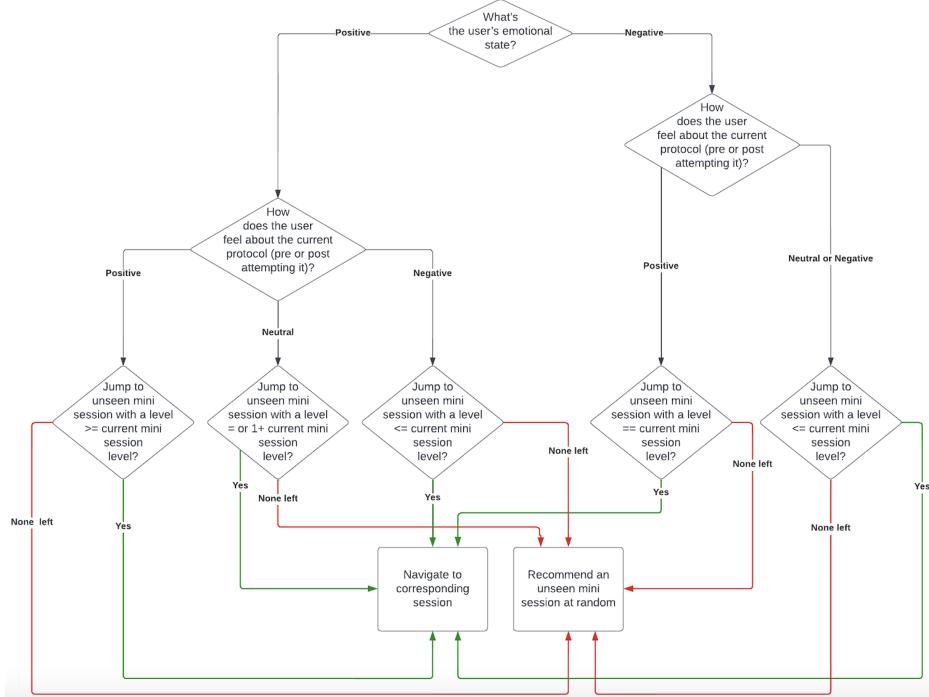


Figure 3.5: Flow chart illustrating the logic for determining the next mini session.

scalable model, such that, if in the likely case the flow requires adjustment or more protocols are developed, the model facilitates simple insertion of new mini sessions or modification of existing mini sessions.

We also implement support for gathering (anonymously and with the users' consent) important insights from each session via our embedded SQLite database that we can use to help improve our model for the future. These insights include:

1. How the user feels about the protocols before attempting them
 2. How the user feels about the protocols after attempting them (along with (1), can be very helpful for research purposes)
 3. Whether the user finds the chatbot funny or not (can help towards improving our humour recognition model)
 4. How the user felt after the end of the session, including session details, i.e. protocols discussed, chatbot jokes, etc. (can help improve the conversation flow and the overall model, e.g. from close inspection of the conversation, we can understand what users find more helpful)

We began recording some of these insights during our first trial, and our findings are summarised in section 4.3.

3.4.2 Front end and user experience

Our chatbot's web interface was created using the React-chatbot-kit library, which additionally handles the communication between its Python back end and the JavaScript front end. It supports user input in the form of text or a selection from a group of existing options.

By default, the interface supports desktop devices. To improve its performance with responsive web design, we extended support for different screen sizes and orientations such that our application is compatible with all mobile, tablet and desktop devices, as seen below. We also consider this significant as it contributes to delivering services that can help improve mental health to the wider population, given that over 80% of people from low and middle-income countries owns a mobile phone [127].

Additional minor updates to the interface were implemented, including the chatbot and user avatars, to ensure all components matched a single theme for a more professional design. We also incorporated emojis, images and clickable links in chatbot messages to improve the user experience and engagement while interacting with the chatbot. In addition, in the desktop view, we display the self-initiated humour protocols on the right of the screen to act as a further guidance tool for users, which allows them to digest these protocols at their speed.

Figures 3.6 and 3.7 show our web application's user interface in tablet and mobile view-modes, respectively. Figure 3.8 shows its user interface in the desktop view (i.e. with the protocols shown on the screen).

3.4.3 Cloud deployment

To deploy our web application, we used a single Amazon Elastic Compute Cloud (Amazon EC2) instance for hosting our model, interface and embedded database. We set up Nginx and Gunicorn servers to host the front end in React and the back end in Flask, respectively.

Our application is therefore available on the Web and is accessible via the following link: <http://35.176.235.166/> until 20th July 2022.

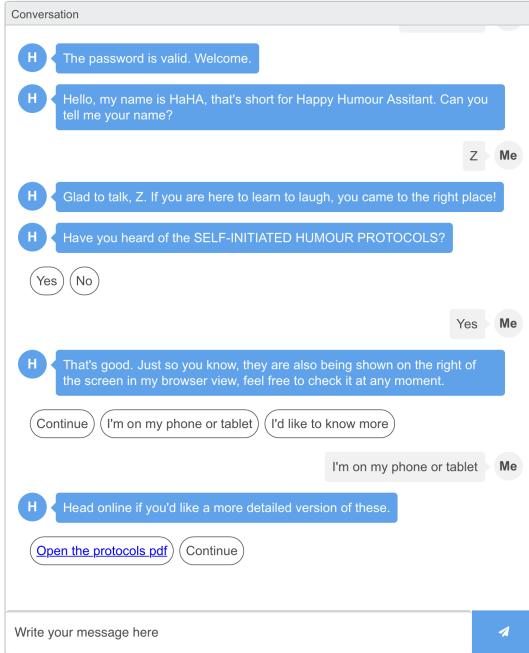


Figure 3.6: User interface of the HaHA web application (Tablet view).

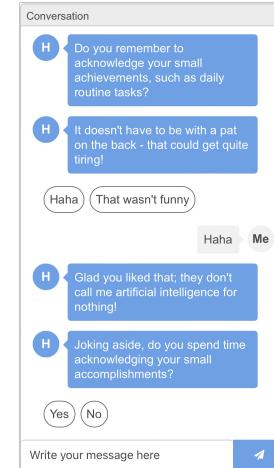


Figure 3.7: User interface of the HaHA web application (Mobile view).

This screenshot shows a laptop view of the HaHA web application. On the left, the 'Conversation' panel shows a dialogue between the 'H' (Assistant) and the user. The 'H' says: 'That's great to hear!' and 'Do you want to know more about this exercise? Or is it time to move on to push-ups? Just joking! There won't be any of those featuring this session. Otherwise, I wouldn't be here!'. The user replies: 'Yes' and 'No'. The 'H' then says: 'To induce a change in perception, you can stare at this image of the GESTALT VASE.' Below this, there is a gestalt vase image with the text 'Click to continue'. At the bottom, there is a message input field: 'Write your message here'.

On the right side, there is a 'Protocol viewer' section containing several text boxes and images related to different exercises:

- (i) SELF/WORLD INCONGRUITY: A text box explaining the concept of incongruity between the 'hard reality' of the external world and our personal perceptions, with a reference to superiority theory.
- (ii) CONTRASTING VIEWS: A text box explaining how staring at a gestalt vase until perception changes can lead to a smile or laugh, with a visual of the gestalt vase.
- (iii) OUR OWN LAUGHTER BRAND: A text box explaining how to create a new form of laughter by employing a minimum amount of energy, with a list of vowel sounds ('ah, ah, ah, ah, ...') and a note about self-reinforcement.
- (iv) LAUGHING WITH THE MOUTH: A text box explaining how to笑 with the mouth open, with a note about self-reinforcement.

Figure 3.8: User interface of the HaHA web application (Laptop view).

Chapter 4

Evaluation

We now discuss the formal evaluation of our chatbot, HaHA, which takes place as two remote non-clinical trials. Our first trial forms the main body of our evaluation as our second trial is currently in-place. We also mention important insights gathered throughout the trials on various aspects of our project that can be used in future research and to further our platform.

4.1 Trial details

42 volunteers took part in our first formal research study, titled “Evaluation of a Virtual Agent in Guiding Users from the Non-Clinical Population to Learn to Laugh via Self-Initiated Humour Protocols”. Many of these volunteers had existing knowledge of SAT, and some were aware of the self-initiated humour (SIH) and its protocols.

Participants were instructed to communicate with HaHA for five days for between 10 to 15 minutes each day. They were informed what to expect during their communications and given an overview of the conversation flow.

We provided all participants with an information sheet that closely detailed the study purpose, participant requirements, what would happen to the study results, and further details, i.e. organisers, contact details, etc. They were also free to withdraw at any stage of the study. Upon receiving participants’ consent via an online consent form, we shared with them a link to the platform and personal login details. After completing all of their sessions, volunteers were asked to fill out a questionnaire.

After the completion of this trial, we implemented changes to our framework in line with the feedback we received. Our framework is currently undergoing a second trial with a more selective participant pool. This pool consists of the AHD research team members (or close associates) who are all familiar with the SAT and or SIH protocols. Participants also include four clinicians trained in mental health, one of whose responses to the evaluation we received at the time of writing this report,

which we recorded separately, as outlined in Figure 4.9 and Appendix F. The trial instructions remain the same across both studies.

4.2 Participant feedback

Of the 42 volunteers, two members chose to withdraw due to lack of time, and 38 submitted responses to the evaluation questionnaire. We now analyse their responses and outline our main findings while making comparisons with the SAT chatbot [6], where appropriate. We include the full feedback from all trial participants in Appendix F.

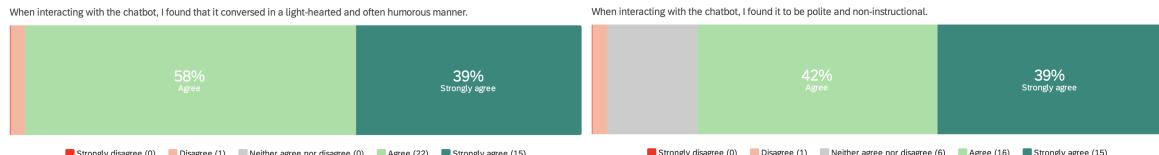
4.2.1 Light-hearted and humorous

We asked participants to evaluate the humour of our chatbot, a central element of the success of our project, and we found that 97% agreed (or strongly agreed) that it conversed in a light-hearted, humorous manner (see Figure 4.1). Users were “surprised to find that [they were] laughing while interacting with the chatbot” or that it “made [them] giggle several times” and “overall formed a happy, cheerful perception of HaHA”. Participants reported that its humour was “light and easy, which took off a lot of pressure” and that it “inspires to continue the practice” of the self-initiated humour protocols.

Due to the ambiguous nature of humour [4], there was naturally some disagreement between participants, with some “really interested in the jokes”, stating “in some instances, it would make [them] smile” and that “more would be nice”. On the hand, some felt “it was a little overdone”, especially when “people may not want a light-hearted, humorous conversation”. To account for this, we could fine-tune our optimisation function, altering the humour weighting based on whether the user finds the chatbot humorous or not (for which the logic is already in place). Also, by occasionally asking the user how they are feeling or by explicitly asking them, we can determine whether users are more receptive to a light-hearted, humorous conversation, which the platform can also use to fine-tune our optimisation function.

4.2.2 Polite and non-instructional

Another important consideration was for our chatbot to have a polite, self-effacing and non-patronising demeanour while teaching the user how to learn to laugh, which we evaluated by asking users how much they agreed with the statement, “When interacting with the chatbot, I found it to be polite and non-instructional.” As seen in Figure 4.2, 81% of participants agreed with this statement, with 39% strongly agreeing, and amongst the remaining 19%, 6 participants were neutral, and one disagreed.

**Figure 4.1:** Humour evaluation.**Figure 4.2:** Politeness evaluation.

We anticipated that in guiding users, such as in our case, through laughter exercises, tone of voice could quite easily become patronising. As such, when analysing statement rewritings, we tried to ensure they were entirely meaningful and deliberate while remaining polite. Users felt “very natural guidance, without feeling forced to perform” and “appreciated that when [they] did not feel ready or comfortable doing an exercise, the chatbot did not insist or got stuck there but offered that I could do it later in my own time”. Naturally, there was some variation due to the wide pool of volunteers writing these statements, where some came across as more robotic, a possible explanation for the remaining 19%. One user felt that “at times the chatbot could be perceived as callous ([when] asking one to laugh at a misfortune without much attempt to understand what that misfortune may be)”. An immediate remedy for this could be for the bot acknowledge it doesn’t intend to be callous at the start, or a more effective solution might be to remove such phrases from our dataset or to phrase them in a “softer manner”.

4.2.3 Usefulness in teaching users how to learn to laugh

Trial participants evaluated the overall use of our platform through their agreement with the statement, “The chatbot was good at teaching me how to learn to laugh.”, with which 82% agreed (or strongly agreed). This finding is significant since learning to laugh is not an easy task [128] and teaching someone how to do this is an equally challenging feat by humans [26], let alone virtual agents, evidenced by the lack of existing work in this domain. An essential skill for someone trying to develop a sense of humour is their ability to recognise contexts for laughter during daily life, which 89% of trial participants felt our chatbot enhanced. See Figure 4.3 for two detailed breakdowns regarding users’ evaluation of the efficacy of our platform.

Users found our platform “highly helpful as a motivation and structure to practice the protocols”. They “learned some ways to become happy” and a “new level to make [themselves] laugh”. One participant pointed out that the chatbot could also be useful to “help people with social anxiety and people with learning disabilities who wouldn’t know when to laugh, e.g. high functioning autistic people.”

4.2.4 Conversation flow

Another significant aspect of our chatbot is its conversation flow, which we evaluate by asking users to share their level of agreement with the statement, “When interact-

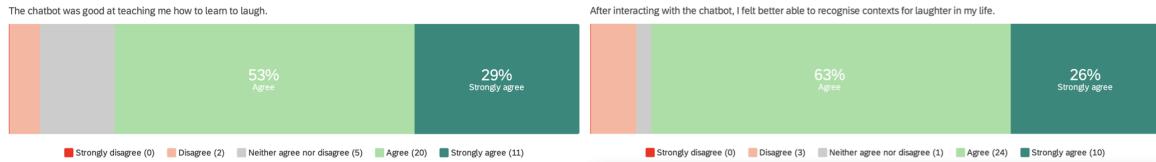


Figure 4.3: Evaluation of the efficacy of HaHA in teaching the user how to learn to laugh (left) and in enabling them to better recognise contexts for laughter (right).

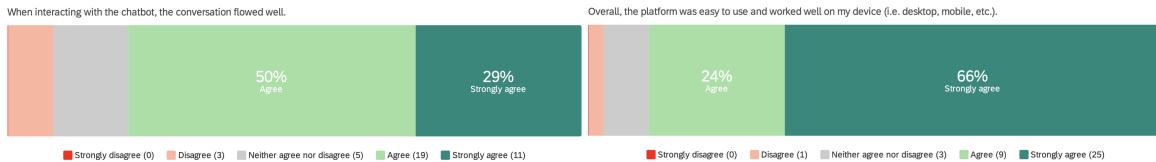


Figure 4.4: Conversation flow evaluation.

Figure 4.5: Ease of use evaluation.

ing with the chatbot, the conversation flowed well.”, with which half of our participant pool agreed, and an additional 29% strongly agreed. Amongst the remaining 21%, 5 participants were neutral, and three were in disagreement (see Figure 4.4). Participants enjoyed “aspects such as linking back to what was covered when [they] end a session [which gave them] the feeling it was designed in a very sophisticated way” and felt that “overall flow is very engaging”. While some felt that “the chatbot explained things really well” with the current flow, others would like the flow to integrate more structured guidance of the protocols with “step-by-step teaching” and suggested adding “a path that provides more structure to learn each of the techniques” and “the learn more option to a technique” the chatbot is taking one through”. One participant felt that the flow was “limited due to 3-4 options of mood” and wanted “more moods”, which we discuss further in Section 5.2.

4.2.5 Ease of use and accessibility

We also requested participants evaluate our chatbot’s ease of use and accessibility, in line with previous feedback on the SAT chatbot [6], such as the “necessity” for mobile compatibility, with users of the SAT chatbot commenting on their “difficulty choosing the bot’s responses and suggestions, scrolling down the page, [and] reading the chat and the instructions on both iPhone and iPad”. 90% of participants felt (or strongly felt) that our platform was easy to use and worked well on their devices (see Figure 4.5). Several participants emphasised this in their written feedback, sharing that it was “intuitive”, “user-friendly and easy to navigate”, and “very easy and fun to use”. They also appreciated the compatibility of our platform: “being able to use it on your phone was great because it makes it easily accessible and portable”. One user found that when using HaHA on their tablet in landscape mode, the “chat was blocked by the protocol viewer”, which we solved before our final trial by updating the its front end.

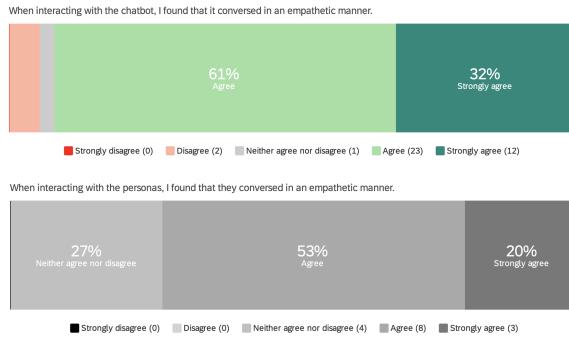
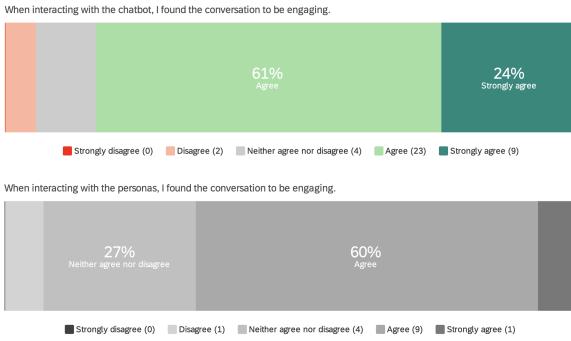
4.2.6 Empathy and engagement

In addition, our chatbot's empathy and user engagement have been evaluated, which we can also compare with the SAT chatbot [6]. Note, when making comparisons, we refer to the SAT bot that consists of four “personas” or “human-like characters” (of different genders and age groups). Figures 4.6 and 4.7 illustrate this comparison. Note that our trial had a larger sample size than the trial of the SAT chatbot (38 compared to 16).

We can see that 93% of participants agreed (or strongly agreed) that our platform conversed in an empathetic manner compared to 73% in the trial of the SAT chatbot. Users frequently commented on our chatbot's ability to convey “empathy in its overall expression”. One user noted that their experience felt “more therapeutic” when allowed to type their response, whereas another felt the “mix of typing [...] and selecting [...], made the platform engaging and interactive.” A potential balanced approach could be to introduce more open-text responses for the user when the chatbot asks them how they feel, or felt, after discussing or trying a particular laughter exercise. For our chatbot to better empathise with users, participants suggested implementing “more emotions to recognise” and “for capturing our emotions and moods, it would be great to include a more nuanced scale.”

When comparing the engagement of conversations between our chatbot, HaHA, and the SAT chatbot, we found that 85% of participants agreed (or strongly agreed) that our platform conversed in an empathetic manner compared to 66% in the previous trial for the SAT chatbot. A potential reason for this might be our platform's use of humour while conversing with users, which most participants enjoyed, sharing that “the conversations were humorous and uplifting”, “the platform was also fun to interact with”, and that they have “really enjoyed engaging with the chatbot”.

Several participants felt that the protocol viewer impeded the engagement at times, stating that although “reading through them was a positive experience [...] this was more engaging, however, when delivered in smaller chunks by the chatbot”. One participant thought the “protocol document was really long and looked intimidating, and it wasn't really necessary”. Some found that, at times, certain aspects of the chatbot, such as its humour, became repetitive, with some “messages start[ing] to feel a bit less funny”. The reason for this could be that several rewritings of the humorous base statements were discarded after crowdsourcing, leading to a smaller corresponding statements pool, which, when fed into our initial optimisation function, often caused specific statements to be outputted over others. Following this feedback, we implemented further changes to our optimisation function to improve its novelty, as seen in Section Y (optimisation function). Another repetitive aspect identified by one participant was the introduction, which “quickly became boring”. A simple solution could be to extract the introductory part of the flow and integrate it into the platform as a separate tutorial that users can choose to activate whenever they wish.

**Figure 4.6:** Empathy comparison.**Figure 4.7:** Engagement comparison.**Figure 4.8:** Comparison between our framework, HaHA (top) and the most recent SAT chatbot [6] (bottom).

4.2.7 Further feedback

We discuss additional written feedback from trial participants regarding the educational and emotional impact of their experience and any further suggestions or comments.

Educational impact

We found that our platform was highly impactful in educating the user, in some cases beyond the field of humour theory. One participant “felt satisfied that [they] had learned something new about humour and [themselves]. For several participants, our chatbot was an educational tool that enlightened them on the theory behind the laughter and humour they had become accustomed to “using naturally” or practising “unknowingly”, giving users a “greater appreciation for how knowledge of laughter and humour can be utilised to build resilience and to develop a more relaxed, and positive outlook on situations and life experiences”. This finding is significant since it illustrates that although one may not actively be trying to develop a sense of humour, our platform can develop their use of humour, enabling them to receive several benefits.

Participants recognised our HaHA’s use “as a teaching tool as to what types of protocols to use and when”. One felt that although the idea of “using a chatbot for teaching has potential, at the moment, it is difficult to have the same impact as a real person”, which we agree with, however, it is clear that our platform is a step in the right direction, confirmed by trial participants: “the chatbot is much more impactful in a learning experience rather than reading through the material.”

Emotional impact

The emotional impact of our platform gauged from our feedback is significant. Participants found that it enabled them to “lessen [their] daily stress and [made

them] feel happier” and “laugh more and more on occasions[, which] definitely improved [their] emotional well-being. They recognised the value of our platform in enabling them to have more control over their mood. One participant found that the conversation “really helped to prompt [them] into thinking about what usually affects [their] mood and how [they] can handle it better.” It also helped a participant “get out of [their low] mood one day during the trial.” In addition, participants felt supported by our chatbot’s nature, stating that “it was light-hearted and simple to use, which makes it less daunting” and that “because of the light-hearted method of communication[, they were] able to clearly understand the protocols and feel confident [they] can practice these in [their] daily habits.”

Several participants also recognised our chatbot’s use as a coping aid, stating that “for people who struggle with anxiety or overthinking, it could really help to guide them to handle their emotions better” and that “it definitely made it easier for [them] to cope with difficult situations.” In particular, one participant “found it transformational! It created a more positive outlook, and therefore, a more elated mood.”

Further strengths and limitations

Several participants welcomed the use of emojis, stating that they “really love to be able to interact with emoji icons instead of just words” and that “they give a nice break from the text”. In a few cases, however, the emojis confused users. One participant brought up an issue they experienced when the chatbot asked them (after giving a brief introduction) how they felt about trying a laughter exercise, where the user had a choice of emoji options to respond. By clicking the smiley emoji, they “expected to learn more”, however, its intended purpose is to indicate that the user understands the exercise and is happy to complete it on their own. Therefore the chatbot moved on to explore another protocol with the user, leaving them “a little disappointed”. We can resolve this ambiguity simply by adjusting the choice option corresponding to this question in the rule-based model so that the user can better understand what to expect from the chatbot.

Relating to this, participants suggested limitations and improvements to the user’s choice selection feature of our chatbot, which could be solved by making simple adjustments to the rule-based model. In particular, one participant felt that there were too many “continue” choice options. Some felt that “the suggested responses and the dice to type can get confusing”, and that “even though the process of responding to prompts was intuitive, perhaps it might be worth disabling the chat feature after showing a prompt to further improve clarity even further for users”.

Participant frequently commented on our chatbot’s language. Some felt that it could be “a little unclear at times of what exercise it is referring to” or misleading when “sometimes a question was asked, followed by a joke in the same speech box” leading to confusion over how the user should respond. Minor adjustments to the

corresponding statements within the HaHAsStatements dataset enabled us to solve these limitations before our second trial. However, this is not a scalable approach to fixing language issues, and instead, we could avoid these by enforcing stricter rules in obtaining statement rewritings via crowdsourcing.

Several participants felt the language of our chatbot played a significant role in enabling its efficacy as a tool for learning to laugh by “creating a comfortable setting to explore personal experiences and humour together”. They also found its examples of real-life scenarios for the protocols “effective in clarifying the protocol task”. Overall, participants felt that the “language and personality created for the chatbot is uniquely kind, informative, and effective in teaching its users about laughter and humour” and enabled them to “thoroughly enjoy exploring the platform”.

4.2.8 Evaluation from clinicians

Our platform was also evaluated by four clinicians trained in mental health, who interacted with HaHA during our second five-day remote trial. We include a summary of the clinicians’ responses to the multiple-choice portion of the evaluation questionnaire in Table 4.9. We also list their written feedback in Appendix F.

Evaluation statement	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
When interacting with the chatbot, I found that it conversed in a light-hearted and often humorous manner.	0/4	1/4	0/4	2/4	1/4
When interacting with the chatbot, I found that it conversed in an empathetic manner.	0/4	0/4	1/4	2/4	1/4
When interacting with the chatbot, I found it to be polite and not too instructional.	0/4	0/4	0/4	3/4	1/4
When interacting with the chatbot, I found the conversation to be engaging.	0/4	1/4	0/4	2/4	1/4
When interacting with the chatbot, the conversation flowed well.	0/4	1/4	1/4	2/4	0/4
After interacting with the chatbot, I felt better able to recognise contexts for laughter in my life.	0/4	0/4	1/4	3/4	0/4
The chatbot was good at teaching me how to learn to laugh.	0/4	0/4	3/4	0/4	1/4
Overall, the platform was easy to use and worked well on my device (i.e. desktop, mobile, etc.).	0/4	0/4	1/4	1/4	2/4

Figure 4.9: Represents four clinicians’ evaluations of HaHA.

4.3 Conversational insights

Under participants’ consent, we recorded conversation data from the trial, which we now analyse. After several instances where the chatbot tells the user a joke, it allows the user to respond with “Haha” or “That wasn’t funny”. Upon inspection of the database, we found that 80% of the responses to this statement (242 in total) were “Haha”, indicating that users often found the chatbot humorous.

If the user is in a negative emotional state (i.e. they were feeling sad, angry or

anxious) at the start of the session, the chatbot asks them again how they feel at the end to identify whether their conversation improved their mood. The user can reply with “Better”, “Worse”, or “No change”. We observed during the trial that 89% of the time, the user responded with “Better”, and the remaining 11% of responses were “No change”.

For research purposes, we also evaluate users’ feelings towards the protocols before and after attempting them, and we share these findings in Figures 4.11 and 4.12, respectively.

In both figures, we use a short-hand for the mini session names, as indicated by the key in Figure 4.10.

Label	PMF	SG	IWS	CV	LB	FL	SL	LAM	LAS
Mini session (protocol(s))	Playful mind and face	Self-glory	Incongruous world and or self	Contrasting views	Our own laughter brand	Feigning laughter	Self-laughter	Laughing at misfortunes	Laughing at long-term suffering

Figure 4.10: Indicates the mini session and its protocol(s) corresponding to the session labels.

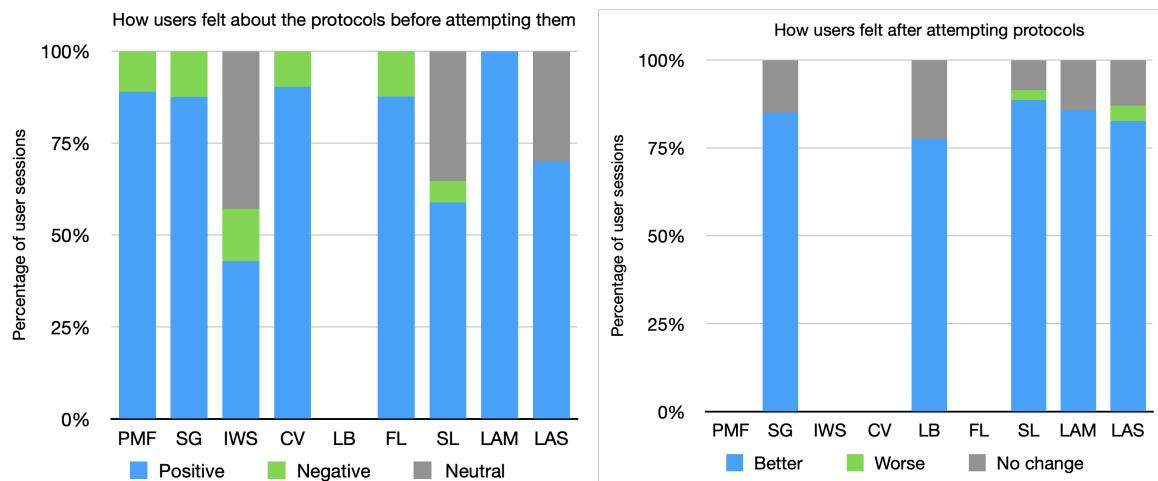


Figure 4.11: Represents the users’ attitudes towards each protocol before attempting it, as a proportion of user sessions in which the protocol featured.

Figure 4.12: Represents the users’ attitudes towards each protocol after attempting it, as a proportion of user sessions in which the protocol featured.

We can see that, on the whole, users feel positive about trying the protocols, although they are more apprehensive about the Incongruous and Self-laughter protocols. In addition, in most cases (84% of the time) users reported feeling better after attempting the protocols (compared to 15% where they felt no change and 1% where they felt worse).

From a brief analysis, it is clear that recording such insights over substantially greater periods (to obtain more session data per protocol) can be very useful for the long-term progress of the proposed approach to learning to laugh and develop a sense of humour.

4.4 Study limitations

Although our studies provide a thorough evaluation of our chatbot, there were several limitations, which we now mention.

One limitation of our first study was that the evaluation questionnaire was entirely anonymous, meaning we were unaware of the gender and age distribution of our pool of participants that shared feedback. Therefore, we cannot identify potential biases in the responses or draw specific conclusions based on these demographics. Although we obtained these details for our second study, substantial feedback was recorded from the first trial (38 responses), where analysis of demographics would have been particularly useful.

Another limitation was the study size. As mentioned, 38 of the 42 and 8 volunteers who participated in our first and second studies, respectively, shared feedback on the evaluation questionnaire by the time of writing this report. This resulted in relatively small sample sizes, especially in the second trial, due to its more selective nature and the limited time available to invite participants. In the ideal case, we should repeat this trial over a greater participant pool, balanced with respect to their gender, age group and other relevant characteristics, which we feel can invite bias into our study.

One such characteristic that could bias our results is the participant's prior familiarity with self-initiated humour (SIH) and or self-attachment technique (SAT). Since, in our first study, we did not keep a clear record of this characteristic amongst our sample, this introduces a risk of the judgements of those participants who are unfamiliar with these concepts primarily affected by their view of an unfamiliar concept of humour. In particular, parts of the study's written feedback indicate that this may be the case, with several comments emphasised on the humour protocols themselves as opposed to the chatbot delivering guidance on them. This idea was also a significant factor in our choice to perform a second trial over a more selective sample.

One of these characteristics that could bias our results is the participant's prior familiarity with self-initiated humour (SIH) and or self-attachment technique (SAT). In our first study, we did not keep a clear record of this characteristic amongst our sample, which introduced a risk of the reviews of those participants unfamiliar with these concepts, primarily affected by their sentiment towards a new humour concept. In particular, parts of the study's written feedback indicate this may be the case, with several comments emphasising the humour protocols as opposed to the chatbot delivering guidance on them. This idea was also a significant factor in our choice to perform a second trial over a more selective sample.

Chapter 5

Conclusions and future work

5.1 Conclusions and contributions

In this project, we have successfully developed a virtual agent for the new domain of teaching users how to learn to laugh. To validate our work, we set up a non-clinical human trial, from which we obtained extensive user feedback that confirms our chatbot, HaHA, is able to guide the user to learn and practice self-initiated humour protocols as a new approach to developing a sense of humour.

To achieve this, we built a rule-based framework that uses a structured flow, which we designed from the well-defined humour protocols. We have also developed HaHAsStatements, a carefully curated dataset of statements for a chatbot capable of delivering a novel form of laughter therapy, which enhances our flow. We annotated a subset from our dataset, to train and evaluate our humour recognition model that can distinguish humorous chatbot statements from non-humorous ones with 95% accuracy by exploiting state-of-the-art NLP techniques. Using our model, we developed a machine learning platform that augments the rule-based framework, enabling our chatbot to engage with users and converse humorously and empathetically.

In conclusion, we have clearly indicated the ability to help users learn to laugh with the development of our virtual agent that humorously and engagingly guides them through the self-initiated humour protocols. Since our work is the first of its kind, we have made progress on the state-of-the-art in the domain of human-computer interaction, introducing an entirely new application of virtual agents. Furthermore, we recognise the great potential for future work to achieve further progress in enabling users to develop a sense of humour via a computational framework and a long-term goal in reaching close to human standards while communicating with users to succeed in our approach.

5.2 Future work

As discussed, our approach to creating a virtual agent that can engage with the user and teach them how to learn to laugh has brought successful results, however, there are several areas for improvement.

5.2.1 Dataset extension

As seen in our evaluation, participants commented that at times, our chatbot can feel repetitive, particularly with some of the jokes that are reused between sessions. The reason for this is due to some of their small number of rewritings in our statements dataset. Extending our dataset is a fairly simple solution to this, in which, if carried out in the form of a crowdsourcing study, we could require a group of participants that is balanced based on our chosen demographics. By enforcing that our participants come from around the world, we could also obtain more diverse humour, which was mentioned in our evaluation feedback, with one trial participant stating that they'd like to see "jokes based on location or region as [they] could see people all over the globe using it."

5.2.2 Generative model

Since our chatbot is a retrieval-based model, it is limited to predefined responses, which impedes its ability to offer a personalised experience for the user or learn on its own and can often feel robotic due to its rule-based structure. In addition, it places a lot of emphasis on obtaining a "perfect" dataset from the start, which, in practice, is difficult to achieve. As such, with the recent developments in deep learning, the use of generative methods could be explored, whereby chatbots use lots of conversational training data in order to generate new dialogue "on-the-fly" and can improve with further usage using machine learning [129].

To ensure the generated dialogue remains humorous, we could try to train a seq2seq model to translate generated statements into humorous versions, which would require a new training dataset. One approach to obtaining a training dataset is to rearrange the HaHAsStatements dataset into a parallel dataset of non-humorous statements and their corresponding humorous rewritings. This dataset could then be used to train a seq2seq model for translating generated statements into humorous versions. Since safety is a significant concern, we could also adopt this approach to ensure the statements are empathic. Although, it's important to note that this approach still comes with several risks due to the unpredictable nature of text generation. Therefore, a thorough evaluation of these risks would be required before enabling a generative chatbot to be used in sensitive domains such as mental health-care.

5.2.3 Animation

Another area for future work that could improve the user experience of our chatbot includes introducing animation into our conversation flow, and there are endless approaches to do this. Some examples include using them at the start of the conversation to break the ice or integrating them into each of the mini sessions, where short, animated clips could be an interesting way to give an example of a scenario in which the user could find humour. In addition, this would be particularly useful for users with reading difficulties and, in general, to break up the text.

5.2.4 Adaptable conversation flow

As a central component of our rule-based chatbot, the conversation flow is significant in ensuring the user feels engaged throughout the conversation. Our current flow can be extended in several ways by adapting it based on user-based heuristics in order to provide a more personalised experience for the user.

An example heuristic could be the user's mood. One frequent recommendation from trial participants was to introduce "more moods", and, upon inspection of the feedback, it is clear that some participants would appreciate more diversified conversations depending on their mood. Another heuristic could be the time available for the user to explore the platform. This could be used as a parameter to tailor the depth of the conversation regarding each mini session. In some cases, users would appreciate a more handheld walk-through of the protocols, as seen from our trial participant feedback, whereas in others, if they are in a rush, for example, they may only require a quick reminder of the fundamental concepts.

5.2.5 Protocol theory integration

One approach to enhancing our chatbot's user experience would be to improve the integration of protocol theory, i.e. the protocol viewer displayed on the right of the web application in desktop mode. Whilst recognising the document to be very informative, some users felt intimidated by its length and detail, and they felt disengaged when encouraged to refer to this during the conversation. As such, future work related to this could include incorporating the humour theory and laughter exercises within the chat, for example, through message popovers so that the user has the option to learn more about specific aspects of the conversation.

5.3 Ethical considerations

As a form of laughter therapy, this application of computing, like several others nowadays, brings potential ethical, legal, societal and professional issues, which in our case are exaggerated by the sensitivity required when dealing with mental health.

While conversing with our chatbot, a user may share personal or confidential data, for example, when asked "How do you feel?", which we are obliged to handle according to the "data protection principles" under The Data Protection Act 2018, the UK's implementation of the General Data Protection Regulation (GDPR).

Our platform stores conversation data in a UK database to gather useful insights from users on the platform's laughter theory, use of humour and conversation flow, which we can use for further improvement. We restricted access to this data to only the co-investigator of this project and have decided to hold this until September 2022, after which we will wipe all user records. In addition, our platform does not store users' device metadata, e.g. geolocation data, physical or internet addresses, IMEI numbers, etc.

Also, while conversing with our chatbot, the user might discuss personally sensitive subjects that they may not be open to speaking about in public or could cause them to feel stressed, anxious or bring about negative consequences.

An additional ethical issue we must consider involves the inherent risks of automated systems, namely in delivering unforeseen, awkward and even toxic responses that can potentially affect a user's mental state. Although our rule-based structure considerably reduces these risks by using predefined responses, these systems cannot exhaustively predict what will be safe for the user, especially without having emotional intelligence or context as a human would during a conversation.

5.3.1 Non-clinical trial

The above issues came into play during our research study that investigates the effectiveness of our platform in guiding the user to learn to laugh via these protocols and develop a sense of humour. This study was given ethical approval by the Head of the Department and RGIT. During our study, there was no real-time monitoring of participants' conversations since it was a remote trial. In addition, we informed all trial participants how their data would be used via our participant information sheet, and we gained their approval via an online consent form. We only collected minimal information required for successful study completion. In the event that a participant withdrew from the study, we wiped their data from the database. We also anonymised all participant data used for analysis in this report.

5.3.2 Collection of our dataset

We organised another research study in order to obtain chatbot statement rewritings in response to user prompts via crowdsourcing using the Prolific platform. We also obtained ethical approval for the study by the Head of the Department and Research Governance and Integrity Team (RGIT). For our study, in order to analyse our dataset distribution per demographic, we requested participants share their gender and age group, which constitute personal data, as do their Prolific IDs, which were automat-

ically collected, although omitted from our dataset. In addition, we informed all study participants how we would use their data prior to them giving consent, and they had the right to access or delete it from the system at any point.

Bibliography

- [1] Dattani S, Ritchie H, Roser M. Mental Health. Our World in Data. 2021. Available from: <https://ourworldindata.org/mental-health>. pages 1, 22
- [2] Organization WH. Mental health. Burden. Available from: https://www.who.int/health-topics/mental-health#tab=tab_2. pages 1
- [3] Santomauro DF, Herrera AMM, Shadid J, Zheng P, Ashbaugh C, Pigott DM, et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic. *The Lancet*. 2021;398(10312):1700-12. pages 1
- [4] Martin RA, Ford T. *The Psychology of Humor: An Integrative Approach*. 2nd ed. Elsevier Science: Nikki Levy; 2018. pages 1, 5, 6, 8, 9, 10, 13, 20, 38, 46
- [5] Edalat A. Self-initiated humour protocols: An algorithmic approach for learning to laugh; 2022. Available from: <https://www.doc.ic.ac.uk/~ae/papers/SIHP.pdf>. pages 1, 4, 5, 10, 11, 12, 27
- [6] Alazraki L. Evaluation of a Virtual Agent in Guiding Users from the Non-Clinical Population in Self-Attachment Intervention. 2021. Available from: https://www.doc.ic.ac.uk/~ae/papers/Lisa_Alazraki_report.pdf. pages 3, 19, 22, 27, 32, 33, 37, 38, 39, 41, 46, 48, 49, 50
- [7] Ghachem A. Evaluation of a Virtual Agent in Guiding Users from the Non-Clinical Population in Self-Attachment Intervention. 2021 Sep 19,. Available from: https://www.doc.ic.ac.uk/~ae/papers/Ali_Ghachem_report.pdf. pages 3, 41
- [8] Mora-Ripoll R. The therapeutic value of laughter in medicine. *Alternative Therapies in Health Medicine*. 2010;16(6). pages 4, 6, 7
- [9] Martin RA, Puhlik-Doris P, Larsen G, Gray J, Weir K. Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. *Journal of Research in Personality*. 2003;37(1):48-75. ID: 272465. Available from: <https://www.sciencedirect.com/science/article/pii/S0092656602005342>. pages 4
- [10] Bennett MP, Lengacher CA. Humor and laughter may influence health. I. History and background. *Evidence-Based Complementary and Alternative Medicine*. 2006;3(1):61-3. pages 5

- [11] Gervais M, David SW. The Evolution and Functions of Laughter and Humor: A Synthetic Approach. *The Quarterly review of biology*. 2005;80(4):395-430. Doi: 10.1086/498281; 30. Available from: <https://doi.org/10.1086/498281>. pages 5
- [12] Planalp EM, Hulle CV, Lemery-Chalfant K, Goldsmith HH. Genetic and environmental contributions to the development of positive affect in infancy. *Emotion (Washington, DC)*. 2017;17(3):412-20. Undefined; J2: Emotion. Available from: <https://pubmed.ncbi.nlm.nih.gov/27797564><https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5367954/>. pages 5
- [13] Yim J. Therapeutic benefits of laughter in mental health: a theoretical review. *The Tohoku journal of experimental medicine*. 2016;239(3):243-9. pages 5, 6
- [14] Kazarian SS, Martin RA. Humour styles, personality, and well-being among Lebanese university students. *European journal of Personality*. 2004;18(3):209-19. pages 5
- [15] Schneider M, Voracek M, Tran US. A joke a day keeps the doctor away? Meta-analytical evidence of differential associations of habitual humor styles with mental health. *Scandinavian Journal of Psychology*. 2018;59(3):289-300. pages 6
- [16] Ruch W. State and trait cheerfulness and the induction of exhilaration. *European Psychologist*. 1997;2(4):328-41. pages 6
- [17] Vilaythong AP, Arnau RC, Rosen DH, Mascaro N. Humor and hope: Can humor increase hope? 2003. pages 6
- [18] Gonot-Schoupinsky FN, Garip G, Sheffield D. Laughter and humour for personal development: A systematic scoping review of the evidence. *European Journal of Integrative Medicine*. 2020;37:101144. pages 6, 7, 9
- [19] Gonot-Schoupinsky FN, Garip G. Prescribing laughter to increase well-being in healthy adults: An exploratory mixed methods feasibility study of the Laughie. *European Journal of Integrative Medicine*. 2019;26:56-64. pages 6
- [20] Danzer A, Dale JA, Klions HL. Effect of exposure to humorous stimuli on induced depression. *Psychological reports*. 1990;66(3):1027-36. pages 6
- [21] Moran CC. Short-term mood change, perceived funniness, and the effect of humor stimuli. *Behavioral medicine*. 1996;22(1):32-8. pages 6
- [22] Davidhizar R, Bowen M. The dynamics of laughter. *Archives of Psychiatric Nursing*. 1992;6(2):132-7. pages 6

- [23] Berk RA. The active ingredients in humor: Psychophysiological benefits and risks for older adults. *Educational Gerontology*. 2001;27(3-4):323-39. pages 6
- [24] Cohen S. Psychosocial models of the role of social support in the etiology of physical disease. *Health psychology*. 1988;7(3):269. pages 6
- [25] Ellis A. Fun as psychotherapy. *Rational Living*. 1977;12(1):2-6. ID: 1979-08846-001. pages 7
- [26] van der Wal CN, Kok RN. Laughter-inducing therapies: Systematic review and meta-analysis. *Social science medicine*. 2019;232:473-88. pages 7, 47
- [27] Edalat A. In: Self-attachment: A Holistic Approach to Computational Psychiatry. *Computational Neurology and Psychiatry*. Cham: Springer International Publishing; 2017. p. 273-314. Available from: http://link.springer.com/10.1007/978-3-319-49959-8_10. pages 7
- [28] Friedler S, Glasser S, Levitan G, Hadar D, Sasi BE, Lerner-Geva L. Patients' evaluation of intervention by a medical clown visit or by viewing a humorous film following in vitro fertilization and embryo transfer. *Journal of evidence-based complementary alternative medicine*. 2017;22(1):47-53. pages 7
- [29] Bressington D, Mui J, Yu C, Leung SF, Cheung K, Wu CST, et al. Feasibility of a group-based laughter yoga intervention as an adjunctive treatment for residual symptoms of depression, anxiety and stress in people with depression. *Journal of affective disorders*. 2019;248:42-51. pages 7
- [30] Ko H, Youn C. Effects of laughter therapy on depression, cognition and sleep among the community-dwelling elderly. *Geriatrics Gerontology International*. 2011;11(3):267-74. pages 8
- [31] Provine RR. Laughter punctuates speech: Linguistic, social and gender contexts of laughter. *Ethology*. 1993;95(4):291-8. ID: 1994-25284-001. pages 8, 11
- [32] Are you having a laugh? Why NHS doctors will soon be prescribing a dose of comedy; 2022. Available from: <https://www.theguardian.com/lifeandstyle/2022/jan/03/are-you-having-a-laugh-why-nhs-doctors-will-soon-be-prescribing-a-dose-of-com> pages 8
- [33] Ventis W, Higbee G, Murdock S. Using Humor in Systematic Desensitization to Reduce Fear. *The Journal of general psychology*. 2001;128:241-53. pages 8
- [34] Nakanishi K. Using Humor in the Treatment of an Adolescent Girl With Mutism: A Case From Japan. *Psychoanalysis, Self and Context*. 2017;12(4):367-76. pages 8

- [35] Franzini L. Humor in Therapy: The Case for Training Therapists in its Uses and Risks. *The Journal of general psychology*. 2001;128:170-93. pages 8
- [36] Gelkopf M, Kreitler S. Is Humor Only Fun, An Alternative Cure or Magic? The Cognitive Therapeutic Potential of Humor. *Journal of Cognitive Psychotherapy*; (4):235-54. Available from: <https://connect.springerpub.com/content/sgrjcp/10/4/235>. pages 8
- [37] Foster JA, Reid J. Humor and its relationship to students' assessments of the counsellor. *Canadian Journal of Counselling and Psychotherapy*. 1983;17(3). pages 8
- [38] Bachelor A, Horvath A. In: *The therapeutic relationship*. Washington, DC, US: American Psychological Association; 1999. p. 133-78. ID: 1999-02137-004. pages 9
- [39] KUBIE LS. The Destructive Potential of Humor in Psychotherapy. *AJP*. 1971;127(7):861-6. Doi: 10.1176/ajp.127.7.861; 18. Available from: <https://doi.org/10.1176/ajp.127.7.861>. pages 9
- [40] Miller M, Fry WF. The effect of mirthful laughter on the human cardiovascular system. *Medical hypotheses*. 2009;73(5):636-9. Undefined; J2: Med Hypotheses. Available from: <https://pubmed.ncbi.nlm.nih.gov/19477604><https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2814549/>. pages 9
- [41] Morreall J. Philosophy of Humor. *The Stanford Encyclopedia of Philosophy*. Available from: <https://plato.stanford.edu/archives/fall2020/entries/humor/>. pages 10, 11
- [42] Edalat A. Introduction to self-attachment and its neural basis. In: 2015 International Joint Conference on Neural Networks (IJCNN). IEEE; 2015. p. 1-8. pages 12
- [43] Hurley MM, Dennett DC, B JAR. *Inside Jokes*; 2013. pages 12
- [44] Binsted K, Ritchie GD. Computational rules for generating punning riddles. 1997. pages 12, 13
- [45] Binsted K, Ritchie GD. Towards a model of story puns. 2001;14(3):275-92. Available from: <https://doi.org/10.1515/hmr.2001.004>. pages 12
- [46] Strapparava C, Valitutti A, Stock O. The affective weight of lexicon. *Proceedings of LREC*. 2006. pages 12, 13
- [47] Winters T. Computers Learning Humor Is No Joke. *Harvard Data Science Review*. 2021. Available from: <https://doi.org/10.1162/99608f92.f13a2337>. pages 13, 14, 15, 18
- [48] Ritchie GD. Developing the Incongruity-Resolution Theory. 1999. pages 13

- [49] Petrović S, Matthews D. Unsupervised joke generation from big data. 2013;2:228-32. pages 13
- [50] Branwen G. GPT-3 creative fiction.; 2020. Available from: <https://www.gwern.net/GPT-3>. pages 14
- [51] Frolov M. Teaching GPT-2 transformer a sense of humor; 2019. Available from: <https://towardsdatascience.com/teaching-gpt-2-a-sense-of-humor-fine-tuning-large-transformer-models-on-a-single-dataset-10f3a2a2a2>. pages 14
- [52] Valitutti A, Doucet A, Toivanen JM, Toivonen H. Computational generation and dissection of lexical replacement humor. Natural Language Engineering. 2016;22(5):727-49. Available from: <https://www.cambridge.org/core/article/computational-generation-and-dissection-of-lexical-replacement-humor/FD150594F61E0289AAD07827AE343FE4>. pages 14
- [53] He H, Peng N, Liang P. Pun generation with surprise. arXiv preprint arXiv:190406828. 2019. pages 14
- [54] Mihalcea R, Strapparava C. Making computers laugh: Investigations in automatic humor recognition. In: Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing; 2005. p. 531-8. pages 15, 33, 34
- [55] Sjöbergh J, Araki K. Recognizing humor without recognizing meaning. In: International Workshop on Fuzzy Logic and Applications. Springer; 2007. p. 469-76. pages 15
- [56] West R, Horvitz E. Reverse-Engineering Satire, or “Paper on Computational Humor Accepted despite Making Serious Advances”. Proceedings of the AAAI Conference on Artificial Intelligence. 2019;33(01):7265-72. 29. Available from: <https://ojs.aaai.org/index.php/AAAI/article/view/4712>. pages 15, 18
- [57] Winters T, Delobelle P. Dutch humor detection by generating negative examples. arXiv preprint arXiv:201013652. 2020. pages 15, 18
- [58] Mihalcea R, Pulman S. Characterizing humour: An exploration of features in humorous texts. In: International Conference on Intelligent Text Processing and Computational Linguistics. Springer; 2007. p. 337-47. pages 15, 16
- [59] Reyes A, Rosso P, Buscaldi D. Affect-based features for humour recognition. In: Proceedings of the 7th International Conference on Natural Language Processing ICON-09,(to be published); 2009. . pages 15

- [60] Bertero D, Fung P. A long short-term memory framework for predicting humor in dialogues. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies; 2016. p. 130-5. pages 16
- [61] Canned Laughter; 2022. Available from: https://en.wikipedia.org/w/index.php?title=Canned_Laughter&oldid=1034969536. pages 16
- [62] Ramakrishna A, Greer T, Atkins D, Narayanan S. Computational modeling of conversational humor in psychotherapy. In: Interspeech. vol. 2018. NIH Public Access; 2018. p. 2344. Pmid:34307638. pages 16
- [63] Miller WR, Moyers TB, Ernst D, Amrhein P. Manual for the motivational interviewing skill code (MISC). Unpublished manuscriptAlbuquerque: Center on Alcoholism, Substance Abuse and Addictions, University of New Mexico. 2003. pages 16
- [64] Hasan MK, Rahman W, Zadeh A, Zhong J, Tanveer MI, Morency LP. UR-FUNNY: A multimodal language dataset for understanding humor. arXiv preprint arXiv:190406618. 2019. pages 17, 33
- [65] Chiaro D. In: Humor and translation. The Routledge handbook of language and humor. Routledge; 2017. p. 414-29. pages 17
- [66] Vandaele J. Humor in translation. Handbook of translation studies. 2010;1:147-52. pages 17
- [67] Chen PY, Soo VW. Humor recognition using deep learning. In: Proceedings of the 2018 conference of the north american chapter of the association for computational linguistics: Human language technologies, volume 2 (short papers); 2018. p. 113-7. pages 17
- [68] Yang D, Lavie A, Dyer C, Hovy E. Humor Recognition and Humor Anchor Extraction; 2015. pages 17, 34
- [69] Ruder S. Transfer Learning - Machine Learning's Next Frontier; 2017. Available from: <https://ruder.io/transfer-learning/>. pages 17
- [70] Baherti P. A Newbie-Friendly Guide to Transfer Learning; 2022. Available from: <https://www.v7labs.com/blog/transfer-learning-guide>. pages 17
- [71] Transfer learning; 2022. Available from: https://en.wikipedia.org/w/index.php?title=Transfer_learning&oldid=1092582961. pages 17
- [72] Annamoradnejad I, Zoghi G. Colbert: Using bert sentence embedding for humor detection. arXiv preprint arXiv:200412765. 2020. pages 18, 33, 34, 36
- [73] Hossain N, Krumm J, Gamon M, Kautz H. Semeval-2020 Task 7: Assessing humor in edited news headlines. arXiv preprint arXiv:200800304. 2020. pages 18

- [74] Hossain N, Krumm J, Gamon M. “President Vows to Cut Hair”: Dataset and Analysis of Creative Text Editing for Humorous Headlines; 2019. pages 18, 33
- [75] Peters M, Neumann M, Iyyer M, Gardner M, Clark C, Lee K, et al. Deep contextualized word representations. 2018. pages 18
- [76] What is a chatbot?;. Available from: <https://www.oracle.com/uk/chatbots/what-is-a-chatbot/>. pages 18
- [77] Frankenfield J. Chatbot; 2022. Available from: <https://www.investopedia.com/terms/c/chatbot.asp>. pages 19
- [78] Srikanth A. 20 Fascinating Chatbot Applications From Six Key Industries; 2020. Available from: <https://freshdesk.com/customer-service/chatbot-applications-blog/#1>. pages 19
- [79] Sohail M. These Chatbots Are Helping With Mental Health Right Now; 2021. Available from: <https://www.aimblog.io/2021/03/02/these-chatbots-are-helping-with-mental-health-right-now/>. pages 19
- [80] Bauer T, Devrim E, Glazunov M, Jaramillo WL, Mohan B, Spanakis G. MeTooMaastricht: Building a chatbot to assist survivors of sexual harassment. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases. Springer; 2019. p. 503-21. pages 19
- [81] Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR mental health. 2017;4(2):e7785. pages 19
- [82] Ali MR, Razavi SZ, Langevin R, Mamun AA, Kane B, Rawassizadeh R, et al. A virtual conversational agent for teens with autism spectrum disorder: Experimental results and design lessons. In: Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents; 2020. p. 1-8. pages 19
- [83] Ghandeharioun A, McDuff D, Czerwinski M, Rowan K. Towards understanding emotional intelligence for behavior change chatbots. In: 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE; 2019. p. 8-14. pages 19, 33
- [84] Abd-alrazaq AA, Alajlani M, Alalwan AA, Bewick BM, Gardner P, Househ M. An overview of the features of chatbots in mental health: A scoping review. International journal of medical informatics. 2019;132:103978. ID: 271161. Available from: <https://www.sciencedirect.com/science/article/pii/S1386505619307166>. pages 19
- [85] Kraus M, Seldschopf P, Minker W. Towards the development of a trustworthy chatbot for mental health applications. In: International Conference on Multimedia Modeling. Springer; 2021. p. 354-66. pages 19

- [86] Denecke K, Vaaheesan S, Arulnathan A. A mental health chatbot for regulating emotions (SERMO)-concept and usability test. *IEEE Transactions on Emerging Topics in Computing*. 2020;9(3):1170-82. pages 19
- [87] Ly KH, Ly AM, Andersson G. A fully automated conversational agent for promoting mental well-being: A pilot RCT using mixed methods. *Internet interventions*. 2017;10:39-46. pages 19
- [88] Nijholt A, Niculescu A, Valitutti A, Banchs R. Humor in Human-Computer Interaction: A Short Survey; 2017. pages 19, 20, 22
- [89] Dybala P, Ptaszynski M, Rzepka R, Araki K. Humoroids: conversational agents that induce positive emotions with humor. In: *AAMAS'09 Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems*. vol. 2. ACM; 2009. p. 1171-2. pages 20
- [90] Harlowe H. How to Make 'em Laugh, Chatbot Style (+ 20 Jokes for Your Bot); 2019. Available from: <https://medium.com/twyla-ai/how-to-make-em-laugh-chatbot-style-20-jokes-for-your-bot-4e767f9b57b7>. pages 20
- [91] Carpenter R. Cleverbot; 1997. Available from: <https://www.cleverbot.com>. pages 20
- [92] Humour chatbots - when and how to make a funny chatbot;. Available from: <https://www.chatbotpack.com/funny-humour-chatbots/>. pages 20, 21
- [93] Wang J, Gu L, Aiken M. A Comparison of Three Conversational Agents. *Asian Journal of Science and Technology*. 2019. Available from: <https://journalajst.com/sites/default/files/issues-pdf/7316.pdf>. pages 21
- [94] Augello A, Saccone G, Gaglio S, Pilato G. Humorist bot: Bringing computational humour in a chat-bot system. In: *2008 International Conference on Complex, Intelligent and Software Intensive Systems*. IEEE; 2008. p. 703-8. pages 21
- [95] Mundhra R, Lim T, Duong H, Yeo KH, Niculescu A. In: *Towards a Humorous Chat-Bot Companion for Senior Citizens*; 2021. p. 31-9. pages 21
- [96] Atir S. Memory for information paired with humorous, relevant jokes. Unpublished thesisNew Haven, CT: YaleUniversity Senior thesis. 2010. pages 21
- [97] Ryan M. Are AI 'Thinking Machines' Really Thinking?; 2019. Available from: <https://towardsdatascience.com/are-ai-thinking-machines-really-thinking-1fa9d758de6d>. pages 21

- [98] Fry SWA William F. Handbook of humor and psychotherapy : advances in the clinical use of humor. Sarasota, Fla.: Professional Resource Exchange; 1987. ID: 15237532. pages 21
- [99] Tepper N. This Mental Health Chatbot Uses Humor to Get Users to Open Up; 2021. Available from: <https://builtin.com/healthcare-technology/healthtech-chatbot-woebot-coronavirus>. pages 21
- [100] Oliveira LD, Rodrigo AL. Humor detection in yelp reviews. Retrieved on December. 2015;15:2019. pages 22
- [101] Rayz JT. In pursuit of human-friendly interaction with a computational system: Computational humor. In: 2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI). IEEE; 2017. p. 15. pages 22
- [102] Elliott R, Bohart AC, Watson JC, Murphy D. Therapist empathy and client outcome: An updated meta-analysis. Psychotherapy. 2018;55(4):399. pages 22
- [103] Kapadia S. Language Models: N-Gram. Medium. 2019 Mar 26,. Available from: <https://towardsdatascience.com/introduction-to-language-models-n-gram-e323081503d9>. pages 22
- [104] Jurafsky D, Martin JH. Speech and Language Processing. 3rd ed.; 2021. Available from: <https://web.stanford.edu/~jurafsky/slp3/>. pages 22, 23, 24, 25
- [105] Stecanella B. Understanding TF-ID: A Simple Introduction; 2019. Available from: <https://monkeylearn.com/blog/what-is-tf-idf/>. pages 23
- [106] eiki. Natural Language Processing: Naive Bayes Classification in Python. Medium. 2019 Jan 4,. Available from: <https://medium.com/@eiki1212/natural-language-processing-naive-bayes-classification-in-python-e934365cf40c>. pages 23
- [107] Gandhi R. Naive Bayes Classifier. 2018 May 5,. Available from: <https://towardsdatascience.com/naive-bayes-classifier-81d512f50a7c>. pages 23
- [108] Li S. Building A Logistic Regression in Python, Step by Step; 2017. Available from: <https://towardsdatascience.com/building-a-logistic-regression-in-python-step-by-step-becd4d56c9c8>. pages 23
- [109] Support Vector Machine Algorithm;. Available from: <https://www.javatpoint.com/machine-learning-support-vector-machine-algorithm>. pages 23

- [110] Chen Q. Language models and RNN. 2020 May 12,. Available from: https://medium.com/@rachel_95942/language-models-and-rnn-c516fab9545b. pages 24
- [111] Pai A. CNN vs. RNN vs. ANN – Analyzing 3 Types of Neural Networks in Deep Learning; 2020. Available from: <https://www.analyticsvidhya.com/blog/2020/02/cnn-vs-rnn-vs-mlp-analyzing-3-types-of-neural-networks-in-deep-learning/>. pages 24
- [112] Convolutional Neural Networks; 2020. Available from: <https://www.ibm.com/cloud/learn/convolutional-neural-networks>. pages 24
- [113] Olah C. Understanding LSTM Networks; 2015. Available from: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>. pages 24
- [114] Britz D. Recurrent neural network tutorial, part 4 implementing a gru/lstm rnn with python and theano. URL <http://www.wildml.com/2015/10/recurrent-neural-network-tutorial-part-4-implementing-a-gru-lstm-rnn-with-python-and-theano>. 2015. pages 24
- [115] Merritt R. What Is a Transformer Model?; 2022. Available from: <https://blogs.nvidia.com/blog/2022/03/25/what-is-a-transformer-model/>. pages 24
- [116] Howard J, Ruder S. Universal language model fine-tuning for text classification. arXiv preprint arXiv:180106146. 2018. pages 24
- [117] Lutkevich B. BERT language model; 2020. Available from: <https://www.techtarget.com/searchenterpriseai/definition/BERT-language-model>. pages 25
- [118] Devlin J, Chang MW, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:181004805. 2018. pages 25, 37
- [119] RoBERTa: An optimized method for pretraining self-supervised NLP systems; 2019. Available from: <https://ai.facebook.com/blog/roberta-an-optimized-method-for-pretraining-self-supervised-nlp-systems/>. pages 25
- [120] Pykes K. The Most Common Evaluation Metrics In NLP; 2021. Available from: <https://towardsdatascience.com/the-most-common-evaluation-metrics-in-nlp-ced6a763ac8b>. pages 25
- [121] Qualtrics survey tool;. Available from: <https://imperial.eu.qualtrics.com/homepage/ui>. pages 31

- [122] Sabato. What's So Funny? The Science of Why We Laugh. *Scientific American*. 2019 Jun 26,. Available from: <https://www.scientificamerican.com/article/whats-so-funny-the-science-of-why-we-laugh/>. pages 31, 39
- [123] Zhang R, Liu N. Recognizing humor on twitter. In: Proceedings of the 23rd ACM international conference on conference on information and knowledge management; 2014. p. 889-98. pages 33
- [124] Bai F, Ritter A, Xu W. Pre-train or annotate? domain adaptation with a constrained budget. arXiv preprint arXiv:210904711. 2021. pages 35
- [125] Saravia E, Liu HCT, Huang YH, Wu J, Chen YS. Carer: Contextualized affect representations for emotion recognition. In: Proceedings of the 2018 conference on empirical methods in natural language processing; 2018. p. 3687-97. pages 38
- [126] Radford A, Wu J, Child R, Luan D, Amodei D, Sutskever I. Language models are unsupervised multitask learners. OpenAI blog. 2019;1(8):9. pages 39
- [127] World Development Report 2016: Digital Dividends;. Available from: <https://www.worldbank.org/en/publication/wdr2016>. pages 43
- [128] 5 Tips to Learn How to Laugh at Yourself; 2018. Available from: <https://exploringyourmind.com/5-tips-learn-laugh-at-yourself/>. pages 47
- [129] Dua A. RULE-BASED BOTS OR AI BOTS? WHICH ONE DO YOU PREFER?; 2019. Available from: <https://dataconomy.com/2019/05/rule-based-bots-or-ai-bots-which-one-do-you-prefer/>. pages 56

Appendix A

Self-initiated humour protocols

Overview of protocols

The set of exercises for self-initiated humour are enumerated below. The first two are mental and muscular exercises to prepare our body to be in a playful mode. The rest provide context and trigger for non-hostile Duchenne laughter. These exercises are to be practiced initially on our own rather than with others. If in our laughter we sense any contempt against ourselves or others, we need to practice neutralising the contempt and converting it into surprise or simple amusement and thus to non-hostile laughter. This would enable us to have health benefits from our laughter.

(i) **Playful mind**

Our sub-optimal habits can cause inflexibility and rigidity in our mind, which is a potential barrier against a spontaneous and playful attitude in life, a crucial condition for being humorous. In this exercise, using the evolutionary theory (4), we practice being more flexible and playful about our beliefs and thoughts. For example, if we firmly follow a political, ideological, religious, or cultural orientation, we may investigate and try to comprehend some of the counter positions to it. We can also exaggerate some of our beliefs to the extent that they sound absurd and thus funny. This does not mean that we necessarily abandon our position but it allows our mind to be a platform for a fluid, constructive and playful discussion between seemingly opposite or different perspectives. We also try to interpret events in the outside world playfully by rising above the fray in binary and opposing viewpoints for example in political debates as in the UK's House of Commons.

(ii) **Playful face**

Rigidity in thoughts and beliefs can also create rigidity in body and facial muscles specifically muscles round the mouth and eyes. In this exercise, in line with (4), we try to become playful and, to this end, loosen up muscles around mouth and eyes by moving them around and by singing our favourite songs to simulate and encourage spontaneity. Duchenne laughter is characterised by loose and half-open mouth and contracting muscles around the eyes, which this exercise helps to create.

(iii) **Self glory**

In this exercise, we learn to laugh on our own over very simple things like our daily routine. On completing any mundane task, such as dish washing or shopping which we may even find boring, give yourself a smile/laugh as a victory gesture to congratulate yourself. In line with the incongruity theory of laughter, we can see a priori, before actually exercising this protocol, why it is actually funny: it violates our expectation as one does not congratulate oneself for such routine and mundane tasks. This a priori humour starts off the protocol once we spontaneously remember to switch to the playful mode. In a sense, we are laughing at the protocol itself. On the other hand, a posteriori after the act of smile and laughter, we can laugh at ourselves because, in line with the superiority theory (1), we have excelled our own former self: previously we completed this task without any accompanying positive affect but now we have done it with a joyful feeling and sudden glory.

(iv) **Incongruous world**

Any discordance, incompatibility or incongruity as well as anything perceived as unusual or extreme in the world that is not immediately threatening can be the context and trigger for Duchenne laughter. The same is true for any change in the outside world, e.g., behaviour of someone or course of events, that we perceive as significant. In this exercise, we practice being cognisant of any contrast, incongruity, inconsistency, or discrepancy in the world and use them, by the incongruity theory (2), as an underlying reason to smile or laugh. As we learn to laugh at contrasts that we recognise and discover in this way, the superiority theory (1) gives us ground for further laughing as we have indeed surpassed our usual and less humorous attitude in life.

(v) **Incongruous self**

In this exercise the context for humour is any incongruity in our own life, mind and behaviour, such as our own contradictions, conflicts, change of attitudes, thoughts, assumptions, expectations. Every individual is full of such inconsistencies. We practice being cognisant of any contrast, incongruity or discrepancy in our own world and use them, by the incongruity theory (2), as an underlying reason to smile or laugh. Again, as we laugh at such disparities that we recognise and discover in this way, the superiority theory (1) gives us ground for further laughing as we have indeed surpassed our usual and less humorous attitude in life.

(vi) **Self/world incongruity**

In this exercise the context for humour is the incongruity between the hard reality of the external world and our personal expectations. A common example is the recognition of any sharp or wide difference between our expectation of someone or something and the reality of the matter as it stands or as it unfolds. In this exercise, we practice being cognisant of any contrast between reality and our expectation and use it, by the incongruity theory (2), as an underlying reason to smile or laugh. Once again, as we learn to laugh at contrasts between reality and our expectation, the superiority theory gives us ground for

further laughing as we have indeed surpassed our usual and less humorous attitude in life.

(vii) **Contrasting views**

In this exercise, we stare at the gestalt vase below until our perception changes and we see two white faces looking at each other. In line with the incongruity theory (2) and the evolutionary theory (4), we smile or laugh when this change of perception takes place, confirming that we can amusingly and laughingly switch our interpretation even if the object of our view remains the same. Similarly, we stare at the two white faces until our perception changes and we perceive the gestalt vase, at which point we smile or laugh.



(viii) **Our own laughter brand**

In this exercise, in line with evolutionary theory (4), we practice playfully creating our own form of laughter which will have our own signature. Moreover, we design our brand in such a way that it uses a minimum amount of our energy so that it can be used for a long period of time or employed for many times repeatedly. With muscles around the mouth loosened and mouth open, we repeat one of the following repetitions phrases (using a vowel like “a”, “e”, “o”) while turning it into laughter as you look into a mirror:

ah, ah, ah, ah, ...
eh, eh, eh, eh, ...
oh, oh, oh, oh, ...
ih, ih, ih, ih, ih, ...
uh, uh, uh, uh,

This exercise is a priori funny and amusing by the incongruity theory (2) as we do not expect ourselves to create a new form of laughter. Having practised the exercise, we can laugh by the superiority theory (1). Once we habituate ourselves with our own new form of laughter, we are encouraged to find humour in new contexts and respond to it with this self-created form of laughter.

(ix) **Feigning laughter**

We swim or run as a physical exercise not to get from one place to another, but because these exercises keep us physically healthy. Similarly, to keep our spirits high, we can learn to laugh as a mental exercise without any accompanying humour. It is known that feigning Duchenne laughter, as practiced in laughter yoga, brings similar physiological benefits as genuine humour-based

Duchenne laughter. In self-initiated humour laughter, however, we consider feigning Duchenne laughter as funny in itself since, by the incongruity theory of laughter (2), we do not expect anyone to systematically fake laughter. The exercise is also a posteriori funny by the superiority theory.

(x) **Self laughter**

In this exercise, in line with the evolutionary theory (4), we learn to become playful and laugh at our own everyday errors, lapses, blunders, flaws, glitches, miscalculations, mismanagements and other faults. While at first, we may feel upset about these minor issues, we practice to interpret them differently and try to quickly laugh them off by the incongruity theory (2), since we do not expect these faults to occur, as well as the playful theory (4). This exercise too is a posteriori funny by the superiority theory.

(xi) **Laughing at misfortunes and disturbing circumstances**

In this exercise, which has a longer description than the previous ones, we learn to laugh at our misfortunes, tragedies, setbacks, rejections, failures and disasters as well as at disturbing events and circumstances in line with (3). The objective is to be able to turn negative emotions like sadness, anger, fear and disgust to laughter. We start by trying to reinterpret these events and circumstances and identify any positive effect they had in our lives. At the very least we managed to survive them and since “What does not kill you makes you stronger” (Nietzsche), they may have resulted in some gain for us. By looking at them from a different cognitive perspective, we would prepare the ground for exercising this protocol.

Next, we consider the following shocking paragraph in Nietzsche’s writings: “To those human beings who are of any concern to me I wish suffering, desolation, sickness, ill-treatment, indignities—I wish that they should not remain unfamiliar with profound self-contempt, the torture of self-mistrust, the wretchedness of the vanquished: I have no pity for them, because I wish them the only thing that can prove today whether one is worth anything or not—that one endures.”

This paragraph makes sense in combination with his famous saying: “What does not kill you makes you stronger.” However, Nietzsche’s wish above is also very funny and a harmless violation of our deep-rooted beliefs by the incongruity theory (2) since we do not usually wish a friend misfortune.

Thus, we can use the above shocking quote as a way of playing with our misfortunes in the way suggested by Chaplin.

Therefore, in this exercise, we consider a disturbing event or circumstance that took place in the distant past that we have struggled with for a long time, and despite its painfulness we try to see a positive impact it has had on us. We recite the shocking quote above, we remember our misfortune and begin to laugh out loud when we finish the sentence: “To those human beings who are of any concern to me I wish suffering...”, and then continue to laugh as we recite the rest of the quote.

After repeating such exercises, once we have experienced the benefit of laugh-

ing at distant problems, we can gradually begin to laugh at more recent disturbing events or circumstances.

(xii) **Laughing at long-term suffering**

The final exercise aims to use laughter to come to terms with and accept not just isolated misfortunes and disturbing events and circumstances in life, but also any long-standing pattern of suffering.

As in the previous exercise, we first try to identify what positive result may have come out of this long term difficulty which would provide us a new perspective for interpreting the problem. Then, we use another well-known quote by Nietzsche. “My formula for greatness in a human being is amor fati [Latin for: love of fate]: that one wants nothing to be different, not forward, not backward, not in all eternity. Not merely bear what is necessary, still less conceal it—all idealism is mendacious in the face of what is necessary—but love it”.

In today’s world, this statement violates our usual beliefs and is therefore funny by the incongruity theory (2) as it encourages us not just to bear (as we are normally advised to do by friends and experts) but to love all that has happened to us including all our suffering. Again, this makes sense with Nietzsche’s motto: “What doesn’t kill you makes you stronger.” It is also in line with Chaplin’s suggestion (3) if the suffering took place in the past.

Therefore, in this exercise, we recite the above quote and when we reach to the final words “but love it”, we laugh out loud while we are thinking of our suffering.

Protocols grouping

- A Simple starting exercises to prepare our body to be in a playful mode – **level 0** (no) difficulty
 - (i) Playful mind
 - (ii) Playful face
 - (iii) Self-glory
- B Protocols that involve producing new forms of laughter – **level 1** difficulty
 - (viii) Our own laughter brand
 - (ix) Feigning laughter
- C Protocols that exploit (world- and or self-) incongruities recognised by the user as a context for humour – **level 1** difficulty
 - (iv) Incongruous world
 - (v) Incongruous self
 - (vi) Self/world incongruity
 - (vii) Contrasting views

- D Protocol that exploits user's everyday errors as a context for laughter – **level 2** difficulty
 - (x) Self-laughter
- E More involved, serious protocols that require a level of experience and approval by the user – **level 3** (highest) difficulty
 - (xi) Laughing at misfortunes and disturbing circumstances.
 - (xii) Laughing at long-term suffering

Appendix B

Dataset and sample

Data collection surveys

General 1: https://imperial.eu.qualtrics.com/jfe/form/SV_9ufPuVis1xoK7Zk

General 2: https://imperial.eu.qualtrics.com/jfe/form/SV_6DvcAfTt0jlhiPs

Scenario: https://imperial.eu.qualtrics.com/jfe/form/SV_3mWJL8gw3yWeQjc

Humour 1: https://imperial.eu.qualtrics.com/jfe/form/SV_9vM8jmiF4b20rgG

Humour 2: https://imperial.eu.qualtrics.com/jfe/form/SV_dgldvikeLFkWTxY

Dataset outline

Tables B.1, B.2, B.3, B.4, B.5, B.6 below represent the base statements in each grouped subset of the HaHAsStatements dataset and their number of statement rewritings.

Humour dataset sample per survey

Table B.7 shows a detailed survey breakdown of the humour dataset sample used to train our humour recognition model.

Base statement	Number of rewritings
Hello, my name is HaHA, that's short for Happy Humour Assitant. Can you tell me your name?	34
It's good to talk to you, {}. I will help to teach you to laugh.*Are you aware of the SELF-INITIATED HUMOUR PROTOCOLS?	25
Good. They are also being shown on the right of the screen in my browser view. You can look at them at any time during our chat.	26
I would suggest you look at them before we chat so you can build on your understanding of them in the session. You won't need a hammer for this!	25
They are being shown on the right of the screen in my browser view if you want to go over them now or any time during our chat.	27
Also, our chat may include some humour jargon. But this will be in CAPITALS to make it clear. You can look at the notes for more information.*With that we can begin as soon as you are ready.	29
Well done, {}, we have now addressed the relevant humour contexts.*Are there any particular SELF-INITIATED HUMOUR PROTOCOLS you want to go over?	32
But are there any particular protocols you want to go over before we conclude our chat?	30
Please choose the protocol you want to go over.	28
That concludes our session, {}!	33
Thank you very much, {}! I really hope you enjoyed our conversation and I look forward to speaking with you again soon.	27
Nice to hear it, {}. You can reload the page and enter your details in again for another talk!*Goodbye for now then!	31
So, {}, would you like to explore another humour exercise?	33
By being more open to our thinking and beliefs and having a PLAYFUL MIND we can nurture our sense of humour.	26
Let's look at another way to get into a more open and flexible mode. It's really as easy as it sounds!	25
So tell me, {}, do you have a favourite song?	27
To have a PLAYFUL FACE you can relax your mouth and eye muscles while singing your best song.	26
Have you ever rewarded yourself on such achievements by giving yourself a smile or laugh.	28
Would you like to try this with something simple, like brushing your teeth?	25
If you feel comfortable, you can try laughing at this incongruity.	23
Have you tried forming your own way of laughing?	26
To maintain high spirits, you could try a related exercise that's often practised in laughter yoga called FEIGNING DUCHENNE LAUGHTER.	23
Your personal LAUGHTER BRAND can be very useful for lifting your spirits and helping with anxiety.*This will take time to develop so don't feel any urgency with this exercise.	27
Have you made any mistakes lately in your daily life?	25
You can laugh at your own faults or mistakes in the same way.	28
If at any time you make a mistake then be playful with it and laugh it away.	28
Have you previously experienced any failures or misfortunes?	26
Would you be able to try laughing this off?	24
I'm pleased you could identify how to laugh it off, {}, well done.*It's important to be able to understand the context behind our laughter as this can allow us to nurture our sense of humour.	27

Figure B.1: General 1

Chapter B. Dataset and sample

Base statement	Number of rewritings
If I may, I would like to ask if you have already tried to make light of this setback and see it from a different perspective. Do you think there could be another side to this, perhaps a more positive, even funny, one? *We call this LAUGHING AT MISFORTUNES.	14
Could I ask you if by now, this is an issue you are quite familiar with? Have you been struggling with it for a while?	15
You know, Nietzsche wrote that pain and failure are not something to run away from, but they should be accepted and even cherished.*I'll show you, below, the paragraph where he talks about pain. Please recite it while thinking about this current setback and try to start laughing out loud when you read the sentence: {}.*And try to continue to laugh while reciting the rest of the paragraph. We can laugh at this because it conflicts with our deep-rooted beliefs.	14
Ok, so here is the quote:{}*Let me know once you have finished reading this.	14
Has there been anything painful in your life that has affected you for a while? For example, an illness or the loss of a loved one?	14
Would you say that this issue has been affecting you for a very long time and that you've found enough time to really reflect on it?	14
Through one of his famous quotes, Nietzsche encourages us to "love" everything that happens to us, good or bad, pleasurable or painful. This, obviously, contradicts our belief that painful things are undesirable and not likeable, and therefore I find this quote quite funny.*Now I would like you to read this quote out loud while thinking of this hardship you've been going through, and laugh aloud when you reach the end of it, particularly the final words, "but love it".	14
Do you think you might be able to apply any of these theories of humour and laugh at this?*If you are not sure which theories I am referring to, feel free to have a look at them in the humour document.	13
Do you think that exploring the possible cause of your current feeling could benefit you?	13
Thinking about it, could the root of this negative feeling be an incongruity or an inconsistency, perhaps?	13
Do you feel like exploring more possible triggers for this feeling you are experiencing?	15
Could your low mood be due to a distant setback?	13
Is there something particularly difficult going on in your life that started a while back and may be causing this current negative feeling?	13
If while you laugh, you perceive that you are actually laughing out of contempt, please try to reflect on it and try NEUTRALISING THE CONTEMPT or changing it into non-hostile humour (by surprise and amusement - refer to the SELF-INITIATED HUMOROUS PROTOCOLS that use incongruities to trigger laughter).	13
You may find more details in the protocols sheet. Please try to review them when you have some time.	13
Do you need any further clarification on this exercise?	14
The theories that we are using here are the INCONGRUITY and SUPERIORITY THEORIES.	13
Would you say that you have experienced any incongruity in your life recently?	14
Are you happy to try this out?	18
Please tell me, {}, how you felt, after attempting this?	18
I hope this makes sense, {}?	19
Would you like to go further with another exercise related to this one that might also benefit you?	14
Have you already tried to laugh this off?	14
To laugh it off, we need to try finding its positive side as this will help us laugh it off and get past our misfortune.*I know, it's not as easy as swapping battery terminals - but don't try this {} anywhere!	14
To try to laugh this off, it is important that we attempt to change our interpretation of this (as with LAUGHING AT MISFORTUNES or setbacks) and focus on any positives that may have resulted from it.	17
Would you like to try exploiting CONTRASTING VIEWS (of an image, for example) as a trigger for laughter?	14
Now we will start exploring ways to develop and improve your sense of humour!	15
Do you think you might be able to identify the reason behind your {}?*Don't worry, {}; you do not need to share the actual reason with me.	19

Figure B.2: General 2

Base statement	Number of rewritings
Glad you liked that; they don't call me artificial intelligence for nothing!	13
Ah, I see. Perhaps it was a byte too corny for you?	14
Oh, no worries, that's totally okay!	13
Okay, fantastic. Let me just pull up the complete quote I keep handy, and we can start.	13
Okay, no worries. You can find a more detailed explanation in the protocol sheet. Please have a look at it now and let me know once you're ready.	13
I totally get it, {}; it's not easy. You've already made the first step by acknowledging that you could laugh it off, so there's not much left to go!	16
Incongruity can trigger DUCHENNE (true) LAUGHTER.*It is likely that you came across incongruities already, for example, when you experienced changes in the outside world or noticed your own contradictions or changes of attitude toward a certain topic. They can also be inconsistencies between what we expect the world to be like and how it actually is.	13
The idea originates from a famous quote by Nietzsche, which you may have seen before: "What doesn't kill you makes you stronger."*For example, the first time I wrote a piece of text, my programmer told me off because it sounded too "robotic"! But since then, I have (modestly) mastered the art of natural conversation, as you can see.	13
How are you now, after our conversation?	33
Basically, this means that you would need to try to view this setback in a different light and acknowledge any positive impacts it may have had in your life.	13
This would mean trying to give a different, novel interpretation of this hardship. Could it be, for example, that you have experienced personal growth because of it? Or has it benefited you in some, perhaps indirect, way?"At the very least, the fact that you have lived through it probably means that you are stronger now (and you didn't need to do any press-ups to achieve this!).	13
Would you like to give it a go?	14
Ideally, you would want to work on this quite consistently, nurture it and grow it like a beautiful flower!*It would make your life better, and it would greatly enrich your work and relationships.	13
You could try to think of a counterposition to one of your most strongly-held beliefs.*For example, if you think everyone is responsible for their own actions, try to play out in your head the argument that everyone is just a product of their own genetic markup and upbringing; no one really has any freedom and thus is not actually responsible for anything!	14

Figure B.3: Scenario

Chapter B. Dataset and sample

Base statement	Number of rewritings
We shall discuss some of the protocols in today's session, but exercising them throughout your daily life will give you the real benefit; natural and frequent laughter, and a general feeling of happiness. Practice makes {} excellent!	14
How are you feeling? You can be honest here, as I will not judge you, unlike most humans.	27
Now, if you can recognise the discussed triggers for laughter (i.e. daily routine-accomplishments, incongruities, faults, etc.) throughout your life, this would be most beneficial.*Remember, developing a sense of humour has no defined endpoint but is a journey. The most important thing is to enjoy it!*And I am always here if you need me - I have nowhere else to go!	13
Good. I hope it goes well for you. Feel free to try this and come back to tell me how it was.*I'll still be here - any time, any place!	14
Do practice this whenever you have some alone time. I know that can be hard for us popular kids.	12
How about practising this again whenever it suits you? If at first you don't (or even do) succeed, try, try again!	12
Is your body in a playful mode? Mine's currently in a chatty mode! Sorry in advance for my rambling!	15
Great. Don't worry; I won't tell you to share it as I know this can be very personal.*I wouldn't want to share mine - Not that I'm embarrassed in my taste in music!	12
I would recommend something, but why not make up your own jamming tune!*I'll start you off: "They call me DJ {}! Boom boom boom"	17
Do you remember to acknowledge your small accomplishments, like daily routine tasks?*Just to clarify, I don't mean saying hi to the laundry basket once you've done the washing!	14
How about congratulating yourself with a smile or laugh after completing something simple, such as doing the dishes.*Don't tell me you can't think of anything - even breathing counts. It's a big accomplishment for some of us!	12
You can induce a change in perception by staring at the following image of the GESTALT VASE, where you should soon see two white faces and smile or laugh while repetitively switching between both interpretations of the image.	12
It comes from GESTALT PSYCHOLOGY, which isn't named after Mr/Ms Gestalt as one might expect. In fact, Gestalt is a German word interpreted as "pattern".*Though, if you wanted to, you could change your name to Gestalt and confuse everyone!	12
We consider DUCHENNE LAUGHTER to be the only indicator of true joy.*In the same way as we swim or run as a physical exercise rather than trying to get from A to B, we can also practice FEIGNING LAUGHTER (laughing without humour) to exercise our minds.*Although I must admit that my laughter occasionally turns into quite an intensive physical exercise!	21
Feel free to laugh at this in your own time and at any minor errors in the future. After all, {} a laugh a day keeps the doctor away! And it's not hard on your teeth!	14
Reflecting on our setbacks is an important prerequisite to laughing them off. We are not in a race.*But there is a prize - personal growth and development in your sense of humour.*I bet I caught your attention when I mentioned the word, "prize".	12
It looks like we can't find the underlying reason at the moment.*Don't worry, though; feeling down, temporarily, is very common and hopefully should pass. But, if the feeling is not diminishing after more than a week, it's probably worth checking up with your doctor.*Maybe you'll run into the root cause later. Better keep your helmet on!	12
Would you be able to try laughing this off?*Rather than laughing your head off!	12
Do remember not to laugh with contempt towards yourself or others; this can cause one to feel down.*It sounds obvious, but these bad habits can result in low moods in the long term.*Like bingeing on sugar, which I know all too well :(12
This exercise aligns with the PLAYFUL or EVOLUTIONARY THEORY of humour, where we are encouraged to be playful with our everyday errors. I don't mean by playing games like hide-and-seek with them!	12
After completing this exercise for distant problems, you can move on to more recent setbacks.*I'd be surprised if it isn't a struggle at first! But the more you practice it, the easier and more natural it feels - like learning to swim, only, without the risk of drowning!	12
Laughing at our hardships will eventually allow us to overcome them and even utilise them for our own personal strength and growth.*So, by practising this, you can become more advanced, by human standards at least!	12
You may be able to exploit one or some of the SUPERIORITY, INCONGRUITY and PLAYFUL THEORIES of humour. These are at the start of the protocols document, which I know, is pretty lengthy, so they may be worth reviewing when you can.*When I say "pretty", I meant "very", not "beautiful-pretty", though I wouldn't say the document is very beautiful either.*Adding colours might be a good idea.	12
Being able to accept and even "love" (as the world-renowned philosopher Nietzsche puts it) our hardships is an important life skill.*We explore this through LAUGHING AT LONG-TERM SUFFERING. Have you tried this before? That's not to say you should start looking for opportunities to suffer in the same way as you might look for a partner!	12
Before trying to laugh off our hardships, we should spend some time reflecting on them. As they say, you can't run before you can walk.*Though I'm being a bit of a hypocrite myself, as I can run but I still haven't learnt to walk. But you get the idea, I'm sure.	12

Figure B.4: Humour 1

Base statement	Number of rewritings
Then this should be fun! To playfully create your form of laughter, perform the following: *Loosen the muscles around your mouth and keep it open while repeating one of the following repetitious phrases (using any vowel) and turning it into laughter: {} *Try to ensure that your form of laughter requires a minimum amount of energy (so you can use it for extended periods). *I would avoid trying this at the zoo, as they might not let you out!	17
An example could be wearing two different colour socks to work by mistake. *I'd like to think of it as a good ice breaker for everyone to laugh at when the conversation gets boring.	14
Wow, I am jealous - I make mistakes all the time!	16
I see. I hope you're not finding it too boring then.	13
Sure, {} - it can get confusing with all these terms associated with humour! *An incongruity is something that's out of place or inconsistent. It can be an internal discrepancy, such as your own change of attitudes, thoughts or expectations. Or it can be external, such as a discordance in the world or a change in others' behaviours. *It could also be a difference between expectation and reality. *One example I can think of is when I caught my grandma riding my bike last week - that was quite bizarre.	15
So, we can laugh them off by the INCONGRUITY and EVOLUTIONARY THEORIES of humour. *An example could be burning a piece of toast; no big deal, as long as it doesn't set the house on fire!	17

Figure B.5: Humour 2

Base statement	Number of rewritings
Before I go on, If I don't respond for a while during the session, please feel free to reload the page to wake me up!	12
Maybe it's an everyday error or flaw you've noticed, that's causing this negative feeling?	13
To confirm, would I be correct in saying you are feeling {}?	51
Sorry about that; I am a new bot but learning with time! Please select the emotion that best reflects how you feel.	52
Great, I look forward to hearing how you find it.	20
Great, I'm glad to hear that.	20
Great to hear that!	50
Sorry to hear that.	20
Ah, that's unfortunate to hear. I can try to help you if you are able to tell me a bit more.	50
A more detailed version of these can also be found online.	20
Would you like to explore another more relevant laughter exercise, {}?	33
But really, how do you feel?	36
So is your body in a playful mode?	48
So do you spend time acknowledging your achievements?	48
Let's get back to the point.	20
So have you tried accepting your hardship in this way?	48
So do you think you could try laughing this off?	48
So does that make sense?	48
So do you think the root of this negative feeling be an incongruity or an inconsistency?	14
So would you like to give it a go?	13

Figure B.6: Extra

Survey breakdowns of the humour dataset sample

Statement group	Number of samples	Number of base statements	Normalised number of samples	Proportion of humour dataset sample
General 1	309	29	10.66	0.19
General 2	245	28	8.75	0.16
Scenario	143	14	10.21	0.18
Humour 1	268	25	10.72	0.19
Humour 2	91	6	15.17	0.27

Figure B.7: Table showing survey statement details and the proportions of the humour dataset sample that came from each survey.

Appendix C

Example conversations

Below are three example conversations between our chatbot (HaHA) and a human (we use substitute names for reporting purposes).

The first, Figure C.1, illustrates the HaHA's conversation with a user who's feeling **positive**. The second and third, Figures C.4 and C.8, respectively, illustrate the conversation with HaHA and a user who's feeling **negative**.

Conversation 1. Emotional context: positive.

HaHA

Hello, my name is HaHA, that's short for Happy Humour Assistant. Can you tell me your name?

James

James

HaHA

Nice to meet you, James. Hopefully, you can leave this session with useful takeaways on laughter and how to nurture your sense of humour.

Are you familiar with SELF-INITIATED HUMOUR PROTOCOLS?

James

James

HaHA

I would recommend giving them a good read before we begin. It will help you get the most from our conversation today.

As you might have seen, they are shown on the screen to the right if you're in my browser view, and they will be available to read at any point during our chat.

James

Continue

HaHA

Practising the protocols we discuss today during your daily life will allow you to laugh more naturally and regularly, and you'll feel happier for it.

They say "what soap is to the body, so laughter is to the soul". So don't try to clean your soul with soap!

James

Continue

HaHA

During our conversation, I may use humour jargon. But don't worry; as you may have already seen, I will identify this with CAPITALS and will explain this if needed. You can also check the documentation for further details.

Anyway, it's time to get started whenever you're ready!

James

Let's start

HaHA

How are you today? I bet you didn't think I'd ask you that.

James

I've had quite a restful day, I'm feeling nice

HaHA

From what you have told me it could be that you are feeling happy/content, would I be right to say so?

James

That's correct

HaHA

That's fabulous.

Let's start exploring how to nurture your sense of humour.

1. water it regularly - wait, ignore that - that was for nurturing a plant!

James

Continue

HaHA

How do you feel about manipulating CONTRASTING VIEWS, of an image, for example, as a motivation for laughter?

James



HaHA

Great, I'm glad to hear that.

Do you need some further clarification on this exercise?

James

Yes

Page 1 of 3

Figure C.1: Page 1 of sample conversation 1 (positive).

HaHA

If you stare at this image of the GESTALT VASE, it will induce a change in perception so that you should also see two white faces.

<shows picture of Gestalt Vase>

James

staring...

HaHA

You should start to see two white faces, and try to smile or laugh, while repetitively switching between the two interpretations.

GESTALT is interpreted as "pattern" in German, and it reflects the idea that the whole pattern is different from the sum of its parts. So I was right; it's essentially a magic pattern!

James

Sounds interesting

HaHA

Try to practise this whenever you feel it's appropriate. Note that the humorous protocols work best when practised alone, but this shouldn't stop you from laughing in public.

Remember, laughter is contagious, so let's spread this happy disease amongst ourselves! On second thought, let's avoid calling this a "disease"; I have a feeling it might not catch on!

James

Continue

HaHA

Do you want me to introduce another exercise?

James

Yes

HaHA

Would you say you remember to reflect on what you've achieved? Acknowledging our achievements is important as it gives us the motivation to push forward.

So unless you're trying to win the tug of war (where pulling backward is the aim), I'd recommend getting on this!

James

Haha

HaHA

Yay! My practice has paid off!

But kidding aside, do you spend time acknowledging your accomplishments?

James

Yes

HaHA

Have you tried congratulating yourself on these accomplishments, for example, by offering yourself a laugh or smile?

James

No

HaHA

How would you feel about trying this on your own, over simple things, like taking a shower?

James

😊

HaHA

Good. I hope it goes well for you.

James

Continue

HaHA

Do you want me to introduce another exercise?

James

Yes

HaHA

Are you experiencing any long-standing hardship or difficulty present in your life, like an illness or loss?

Page 2 of 3



Figure C.2: Page 2 of sample conversation 1 (positive).

James

Rather not say

HaHA

No problem at all! You need only share what you are comfortable with.

So, James, would you like to explore another humour exercise?

James

No

HaHA

Actually, would you like to review any specific protocols before we end our conversation?

James

No (end session)

HaHA

If you can now recognise the discussed contexts for laughter (i.e. daily achievements, incongruities, faults, etc.) often throughout your life, this would be most beneficial.

There is no end to this journey of developing a sense of humour, and the most important thing is to enjoy it along the way!

Remember, I am always here if you need me. It's funny, as I am always running, but still available to talk!

James

Goodbye

Page 3 of 3

Figure C.3: Page 3 of sample conversation 1 (positive).

Conversation 2. Emotional context: negative.

HaHA

Hello, my name is HaHA, that's short for Happy Humour Assistant. Can you tell me your name?

Zoë

Zoë

HaHA

I'm glad to speak with you, Zoë. I try my best to help teach you to laugh. But please remember that it's a joint effort!

Also, are you aware of the SELF-INITIATED HUMOUR PROTOCOLS?

Zoë

Yes

HaHA

Good. Just so you know, they are also being displayed on the right of the screen in my browser view. You are free to review them at any point in our conversation.

Zoë

I'm on my phone or tablet

HaHA

A more detailed version of these can be accessed on the web.

Zoë

Continue

HaHA

By the way, if I blank out at all in our session, don't be alarmed - you can bring me back by simply reloading the page!

Zoë

Continue

HaHA

During our conversation, we will discuss several of these protocols, but the main benefit will come from daily practice; as you start to laugh more often and naturally.

I don't want to sound like your primary school teacher, but practice makes perfect!

Zoë

Haha

HaHA

Ah, that's so encouraging; thank you!

Zoë

Continue

HaHA

Just to let you know, our conversation might include some humour (as you may have noticed already). I will identify this with CAPITALS and give clarification if you need it. Of course you can also refer to the documentation for further explanation.

We can start whenever you are ready.

Zoë

Let's start

HaHA

Is everything good in your world?

Zoë

not really

HaHA

If I understand correctly you are feeling sad, does that describe it accurately enough?

Zoë

That's correct

HaHA

Ah, that's unfortunate to hear. I can try to help you if you are able to tell me a bit more.

There is almost always a reason behind your sadness; can you identify it? You don't need to share it.

Zoë

No

Page 1 of 4

Figure C.4: Page 1 of sample conversation 2 (negative).

Chapter C. Example conversations

HaHA

Would you like to identify the possible reason why you're feeling this way?

Zoë

Yes

HaHA

Maybe it could be an incongruity that's causing your negative feeling?

Zoë

Not sure

HaHA

For sure, Zoë! It can get confusing with all these terms related to humour!

An incongruity is something that's out of place or inconsistent. It might be an internal discrepancy, such as your own change of attitudes, thoughts, etc. Or it might be external, such as a world discordance or a change in others' behaviours. It could also be a difference between expectation and reality.

For example, last week, I caught my grandma doing a live stream on TikTok!

Zoë

Continue

HaHA

So do you think you might be feeling like this because of an inconsistency present in your life?

Zoë

Yes

HaHA

Would it be possible for you to laugh this off? When I start, there's no stopping me!

Zoë

Yes

HaHA

Do try practising this on your own, as it's in this setting that these protocols tend to work best. But feel free to share them with your friends. If only I could do the same!

Zoë

Continue

HaHA

Would you like to learn a different laughter protocol?

Zoë

Yes

HaHA

Can you recall any setbacks from the past?

Zoë

Yes

HaHA

If I may, I would like to ask if you have already tried to make light of this setback and see it from a different perspective. Do you think there could be another side to this, perhaps a more positive, even funny, one?

We call this LAUGHING AT MISFORTUNES.

Zoë

No

HaHA

How about giving this a go?

Zoë

Yes

HaHA

If I may, do you feel that you struggled with this setback for a long time?

Zoë

Yes

Page 2 of 4

Figure C.5: Page 2 of sample conversation 2 (negative).

HaHA

You may find it interesting that Nietzsche preached that pain and failure are not something to run away from, but they should be accepted and even desired.

Refer to the paragraph below (I'm about to show this - please bear with me!) where he talks about pain. Please recite this while thinking about your setback and begin to laugh out loud when you reach the following sentence: "To those human beings who are of any concern to me I wish suffering...".

Then try to continue to laugh while reciting the rest of the paragraph. We can laugh at this because it conflicts with our deep-rooted beliefs.

Does that seem clear?

Zoë

Yes

HaHA

That's fantastic! Let's read the complete quote together now!

Zoë

Continue

HaHA

Here it is:

To those human beings who are of any concern to me I wish suffering, desolation, sickness, ill-treatment, indignities—I wish that they should not remain unfamiliar with profound self-contempt, the torture of self-mistrust, the wretchedness of the vanquished: I have no pity for them, because I wish them the only thing that can prove today whether one is worth anything or not—that one endures.

HaHA

Please tell me when you have read this.

Zoë

Done

HaHA

Once you manage to come to terms with your distant setbacks in this way, you can move on to more recent setbacks.

And do remember to acknowledge your progress as this isn't a walk in the park. Although, it's not affected by heavy rain!

Zoë

Continue

HaHA

How do you feel about doing this?

Zoë



HaHA

Great, feel free to try this.

Zoë

Continue

HaHA

Do you want me to introduce another exercise?

Zoë

No

HaHA

Ah don't go yet! Before we end this, are there any protocols you'd want to go back to?

Zoë

No

HaHA

Then our session has come to an end, unfortunately!

How are you feeling, having reached the end of our session?

Zoë

Better

Page 3 of 4

Figure C.6: Page 3 of sample conversation 2 (negative).

HaHA

Glad to hear it.

Now, to get the maximum benefits for your sense of humour and general happiness, you should now try to recognise the discussed contexts for laughter (i.e. daily routine-accomplishments, incongruities, faults, etc.) as often as possible throughout your life.

There is no endpoint to developing a sense of humour, it is a journey. And remember, the most important thing is to enjoy the process! I believe in you, and of course, I will be here if you need me; unless I'm on my daily jog.

If only that was true!

Zoë

Goodbye

HaHA

Thanks a lot, Zoë! I hope you enjoyed our time together and that we have another conversation soon.

Page 4 of 4

Figure C.7: Page 4 of sample conversation 2 (negative).

Conversation 3. Emotional context: negative.

HaHA

Hello, my name is HaHA, that's short for Happy Humour Assistant. Can you tell me your name?

Alexis

Alexis

HaHA

Great to speak with you, Alexis. Well, I'll be doing my best to help you learn to laugh.
Are you aware of the SELF-INITIATED HUMOUR PROTOCOLS?

Alexis

No

HaHA

I'd advise you to look through these protocols before our conversation so you can use this session to build on your knowledge of them and take your understanding to the next level.
By the way, they are on the right of the screen in my browser view. So you can check it now or at any moment during our conversation.

Alexis

Continue

HaHA

Oh, if during our conversation, you find that I'm not responding after a minute or two, do wake me up by reloading the page - I occasionally take a power nap if I need it!

Alexis

Continue

HaHA

Although we will explore some of these protocols today, performing them throughout your daily life will allow you to receive the main benefit. You should find that your laughter becomes more natural and frequent, and you'll feel happier overall.

Practice makes perfect. And I make comedians!

Okay, that may not be completely true; but we can take it in steps.

Alexis

That wasn't funny

HaHA

Okay, that wasn't one of my best!

Alexis

Continue

HaHA

Just to let you know, our conversation might include some humour (as you may have noticed already). I will identify this with CAPITALS and give clarification if you need it. Of course you can also refer to the documentation for further explanation.

We can start whenever you are ready.

Alexis

Let's start

HaHA

How do you feel? I'm genuinely intrigued as I can't do this yet.

Alexis

Not too good, I've had a difficult day

HaHA

I feel that you are feeling quite sad today, is that right?

Alexis

That's correct

HaHA

I see. Thanks for sharing.

Can you identify the underlying reason behind your sadness, Alexis? You don't need to share the reason.

Alexis

No

HaHA

Would you like to explore the possible cause of your feeling?

Page 1 of 3

Figure C.8: Page 1 of sample conversation 3 (negative).

Chapter C. Example conversations

Alexis
Yes
HaHA
Do you think it could be an incongruity or inconsistency causing your negative feeling?
Alexis
Not sure
HaHA
Might the cause of this feeling be an error or flaw that you've noticed?
Alexis
No
HaHA
Would you like to explore more elaborate potential causes of this feeling?
Alexis
No
HaHA
Could your low mood be due to a distant setback?
Alexis
No
HaHA
Do you recognise any long-standing hardship or difficulty present in your life, such as an illness or loss?
Alexis
No
HaHA
It seems that we aren't able to find the underlying reason at the moment.
Don't worry, though; feeling temporarily low is not uncommon, and hopefully, this feeling should pass soon. But if it doesn't, I'd recommend checking up with your medical doctor, just in case.
I'm only cleared to prescribe laughter.
Alexis
Continue
HaHA
Do you want to have a go at a humour exercise?
Alexis
Yes
HaHA
Have you tried making up a new laugh?
Alexis
No
HaHA
I see; then you're in for a treat! To do this, you will need to relax the muscles around your mouth and leave it open while repeating one of these phrases and transforming it into laughter: "ah, ah, ah, ah, ...", "eh, eh, eh, eh, ...", "oh, oh, oh, oh, ...", "ih, ih, ih, ih, ih, ...", "uh, uh, uh, uh, ...". Remember, for laughter to have a great sound, it needs to come from the heart, not the throat! Find something that really tickles your funny bone, but try to make it easy to repeat; if you're like me, you may want to use it for long periods!
Alexis
Continue
HaHA
Does this make sense to you? I've been told that I can go off on a tangent sometimes. Though I live in a rectangle, so I'm not sure how that's possible!
Alexis
Yes
HaHA
That's good news.
Ideally, you would want to work on this quite consistently, nurture it and grow it like a beautiful flower!
It would make your life better, and it would greatly enrich your work and relationships

Page 2 of 3

Figure C.9: Page 2 of sample conversation 3 (negative).

Alexis
Continue
HaHA
Do you want to discover another exercise?
Alexis
No
HaHA
I don't want us to say our goodbyes yet! Are there any specific protocols you would like to review?
Alexis
No (end session)
HaHA
In that case, it looks like our chat has come to an end.
So, how are you now?
Alexis
Better
HaHA
I'm glad for you.
The discussed contexts for laughter are to be recognised throughout your life for the magic to begin.
And remember, there is no completion point to developing a sense of humour, so "keep calm and carry on" laughing!
Alexis
Haha
HaHA
Ah, that's so encouraging; thank you!
Alexis
Goodbye
HaHA
Thanks, Alexis! I hope you got what you came for from our chat, and it would be nice to meet again soon if you wish.

Page 3 of 3

Figure C.10: Page 3 of sample conversation 3 (negative).

Appendix D

Mini session information

Definition

We define a mini session as a conversation between the user and chatbot where they explore one protocol or group of related protocols together, after which the user has the option to end the main session. A mini session can be initiated in the following scenarios:

- the user is feeling positive
- the user is feeling negative and they have found an underlying cause for this feeling
- the user would like the chatbot to recommend another laughter exercise for them to explore
- the user wishes to explore a specific laughter exercise

Example sessions

We have implemented mini sessions incorporating each of the self-initiated humour protocols and also several combinations of these.

We provide two sample mini sessions below:

Starter mini session: Playful mind, Playful face and Self-Glory

See Figure D.1.

More involved mini session: Laughing at misfortunes

See Figure D.2

Playful mind, Playful face, and Self-Glory {1, 2, 3} (level 0): 'Is your body in a playful mode?'

Ask the user if their body is in a playful mode

- Yes: Ask the user if they spend any time acknowledging their simple achievements, such as daily routine tasks
 - Often or sometimes: Ask the user if they have tried congratulating themselves on these achievements with a laugh or smile
 - → → Yes: Ask the user how this made them feel
 - → → Positive: Encourage the user to try practicing this again whenever they feel ready
 - → → (Ask the user if they'd like to continue exploring other contexts for humour)
 - → → Negative: Provide a reassuring response
 - → → Remind them of the protocols sheet
 - → → → Encourage the user to try practicing this again whenever they feel ready
 - → → → (Ask the user if they'd like to continue exploring other contexts for humour)
 - → → No: Propose that the user tries this on their own, over simple things, such as daily routine tasks.
 - → → (Ask the user how they feel about this)
 - → → Never: Propose that the user congratulates themselves, via a smile or laugh, on something simple that they have completed, such as doing the dishes or even the act of breathing
- No: Inform the user that it's very useful for being humorous, to be more flexible and playful about their beliefs and thoughts
 - Ask the user if they'd like an example
 - → → Yes: Give an example to the user
 - → → (Ask the user what their favourite song is)
 - → → No: Ask the user what their favourite song is
 - → → Inform the user that another way to become more playful is to loosen up their mouth and eye muscles
 - → → Propose that the user moves these around and sings song X in order to loosen these muscles
 - → → → Ask the user how they feel about this
 - → → → Positive: (Ask the user if they'd like to continue exploring other contexts for humour)
 - → → → Negative: Provide a reassuring response
 - → → → Remind the user of the protocols sheet
 - → → → (Ask the user if they'd like to continue exploring other contexts for humour)

Figure D.1: Starter session (i.e. takes place at the start of a conversation where the user is feeling positive)

Laughing at setbacks (Protocol 11, level 3): 'Have you experienced any setbacks in your distant past?'

Ask the user if they've experienced any setbacks in their distant past

- Yes: Ask the user if they've tried to laugh at this setback, while changing their interpretation of it and not seeing it as completely negative
 - → Yes: Ask the user how this made them feel
 - → Positive: Empathetic response
 - → → Encourage the user to practice this again
 - → → (Ask the user if they'd like to continue exploring other contexts for humour)
 - → → Negative: Provide a reassuring response
 - → → Ask if the user would like some further clarification of the exercise
 - → → → Yes: Clarify and give the user an example
 - → → → Remind the user of the protocols sheet
 - → → → Encourage the user to try practicing this again whenever they feel ready
 - → → → (Ask the user if they'd like to continue exploring other contexts for humour)
 - → → → No: Remind the user not to laugh with contempt as that would likely make them feel worse
 - → → → Encourage the user to try practicing this again whenever they feel ready
 - → → → (Ask the user if they'd like to continue exploring other contexts for humour)
 - → → No: Ask the user if they would like to try this
 - → → Yes: Ask if the user has experienced this setback in the distant past and has struggled with it for a long time
 - → → → Yes: Remind the user of the objectives
 - → → → Ask the user if that makes sense
 - → → → Yes: Read the quote together
 - → → → → Propose that they use this as an introduction and practice this in their own time
 - → → → No: Remind the user of the protocols sheet and encourage them to practice
 - → → → → Propose that they use this as an introduction and practice this in their own time
 - → → → Ask the user how they feel about this
 - → → → → Positive: Empathetic response
 - → → → → (Ask the user if they'd like to continue exploring other contexts for humour)
 - → → → → Negative: Provide a reassuring response
 - → → → → Ask if the user would like some further clarification of the exercise
 - → → → → → Yes: Clarify and give the user an example
 - → → → → → Remind the user of the protocols sheet
 - → → → → → (Ask the user if they'd like to continue exploring other contexts for humour)
 - → → → → → No: Remind the user not to laugh with contempt as that would likely make them feel worse
 - → → → → → (Ask the user if they'd like to continue exploring other contexts for humour)
 - → → → No: Propose that the user takes some time to reflect on any setbacks from their distant past
 - → → → → Ask if the user would like some further clarification of the exercise
 - → → → → → Yes: Clarify and give the user an example
 - → → → → → Remind the user of the protocols sheet
 - → → → → → (Ask the user if they'd like to continue exploring other contexts for humour)
 - → → → → → No: Remind the user not to laugh with contempt as that would likely make them feel worse
 - → → → → → (Ask the user if they'd like to continue exploring other contexts for humour)
 - → → → → → No: (Ask the user if they'd like to continue exploring other contexts for humour)
 - → → → → → Not sure: Further clarify and give the user an example
 - → → → → → (Ask the user if they would like to try this)
 - → → → → → No: (Ask the user if they'd like to continue exploring other contexts for humour)

Figure D.2: More involved mini session (corresponds to protocol (xi) Laughing at misfortunes)

Appendix E

Optimisation function selection

We include, below, the final output of our multi-objective optimisation function that retrieves optimised rewritings for a random sample of base statements, which clearly indicates the number of statement repetitions.

We also include a key (see Figure E.1), which represents the label colours used and their indication of the number of statement repetitions.

Figure E.2 illustrates the function output in the case where the user is feeling positive (i.e. $P=1$), while Figure E.3 shows the function output in the case where the user is feeling negative (i.e. $P=0$).

KEY

| Colour fill | Number of repetitions |
|-------------|-----------------------|
| White | 0 |
| Light red | 1 |
| Light green | 2 |

Figure E.1: Represents the labels used to indicate the number of statement repetitions.

Positive retrieval function

| Rewritings retrieved for statement 1 | Rewritings retrieved for statement 2 | Rewritings retrieved for statement 3 | Rewritings retrieved for statement 4 | Rewritings retrieved for statement 5 |
|---|---|--|---|--|
| Nice to meet you, {}. Hopefully, you can leave this session with useful ... | Fantastic! Just so you know, these are also displayed on the right of ... | Happy to hear it, {}. Feel free to reload the page and re-enter your credentials ... | In your everyday life have you made any mistakes recently? | Another way to become more playful is by having a PLAYFUL FACE. I'm not ... |
| Nice to speak to you, {}. I will help you to learn how to laugh.*Have you ever ... | Perfect. I also want to let you know that they can be found on the right side ... | I'm really happy to hear it, {}. Reload the page, re-enter your credentials, ... | Have you made any mistakes or errors recently? If it makes it easier, I'm ... | So the other way to become more playful is to have a PLAYFUL FACE; ... |
| Nice to talk to you, {}. I can offer assistance and guidance to help you ... | That's great. Also, just to let you know, they are shown in the ... | Great, I would love to as well, {}!*You can reload the page and re-enter your ... | Any recent slip-ups? Nothing too big, just something in your day-... | To be more playful, it helps to change the way your face moves, to have a ... |
| Good to talk to you, {}. I'm here to help you learn to laugh.*Have you heard of... It's good to talk to you, {}. I will help to teach you to laugh.*Are you aware of t... | Good. In case you need to refer to these during our conversation, they are ... | I'm happy to hear that, {}.*You can reload this page and enter your informati... | Do you remember making small mistakes in your everyday life? I remembe... | If you can loosen up your mouth and eye muscles, this is another way of ... |
| Nice to meet you, {}. I will help you to learn to laugh and guide you through th... | Good good, in case you need a reminder ('I'm not saying you will!), you can ... | Nice to hear it, {}.*You can reload the page and re-enter your credentials for ... | Have you made a mistake or blunder recently? It doesn't have to have ... | To have a PLAYFUL FACE you can relax your mouth and eye muscles while ... |
| Very nice to meet you, {}! I will do my best to help you develop your sense of ... | Awesome. In case you'd like to have a look at them while we're chatting, ... | I would as well, {}!*Just reload the page and re-enter your credentials ... | Are there any mistakes you have made lately that affected your life on a ... | Having a PLAYFUL FACE will make you more playful, as I'm sure you can ... |
| Nice to speak to you, {}. I will do my best to help you learn to laugh.*Are you ... | Nice! They're also on the right of the screen in my browser view. Have a ... | Glad to hear it, {}, thanks. You can reload the page and re-enter your ... | Have you made any small errors or blunders in the last few days?*Nothing ... | An alternative way to become more playful is to have a PLAYFUL FACE. |
| Nice to talk with you, {}. My goal is to improve your ability to laugh, which ... | Amazing. You can find them on the right of the screen in my browser view. | Nice to hear it, {}. You can reload the page and enter your details in again for ... | Have you made any mistake recently in your day-to-day life? | Another way of becoming more playful is to have a PLAYFUL FACE.*To do ... |
| Glad to talk, {}. If you are here to learn to laugh, you came to the right place! ... | Good. Just so you know, they are also being displayed on the right of ... | Ah, you made my day, {}!*Feel free to reload the page and re-enter your ... | In your everyday life have you made any mistakes recently? | You can also become more playful by having a PLAYFUL FACE (what ... |
| It's good you're here, {}. I will teach you about learning to laugh and ... | Good. They are also being shown on the right of the screen in my browser view. | Great to hear you say that, {}! You can reload this page and enter your details ... if | Have you made any mistakes lately in your daily life? | A PLAYFUL MIND is important, but to be even more playful, you need ... |
| Thanks for coming, {}. Let's hope you leave as a laughter professional!* I'd | Perfect. Just so you're aware, you can find these on the right in my brows... | So would I, {}!*Feel free to reload the page and re-enter your credentials for... | Are there any small errors or blunders you've made recently in your day-to-... | Another way to be playful is to have a PLAYFUL FACE, with relaxed and loose ... |
| Okay, nice that you are here, {}. I will try my best to help you with your ... | Great. By the way, they are also shown on the right of the screen in my browser | That's great to hear, {}.*Feel free to reload the page and we can have another ... | Have you made any small mistakes or recently become aware of minor ... | Your body can also enter a playful mode once you have a PLAYFUL FACE.* ... |
| Great to meet you, {}! I will try my hardest to help your nurture your sense of ... | Fantastic! Just so you know, these are also displayed on the right of ... | Very pleased to hear this, {}.*Feel free to reload the page and re-enter your ... | As much as we try, our lives can never go as perfectly planned as we'd like ... | Having a PLAYFUL FACE is also a way to become more playful, and it's an easy ... |
| Nice to meet you, {}. I will do my best to help you learn to laugh.*Do you ... | Great. Just to let you know, they are also being displayed on the right of... | Glad to hear that, {}.*When you feel like it, go on and reload the page and re... | Have you made any mistakes lately in your daily life? | To be more playful, it helps to change the way your face moves, to have a ... |
| Great to speak with you, {}. Well, I'll be doing my best to help you learn to ... | Super. Just to let you know, they are being shown on the right of the screen if ... | *You can reload the page and enter your credential... | Would you say you've recently made any errors, just little things, in your ... | Consider having a PLAYFUL FACE; you can loosen up your mouth ... |
| Pleased to meet you, {}! I'll try my best to teach you the lesson of laughter. It... | Excellent. Feel free to review them whenever you like. I wouldn't blame you... | That's great, {}, I would too! *You can reload the page and enter your credential... | Have you been making any errors or blunders recently? I'm not asking for ... | Having a PLAYFUL FACE is important for being in a playful mode. What's that... |
| Nice to meet you, {}. I will do my best to teach you how to laugh today. It's ... | Good. They are also being shown on the right of the screen in my browser view. | I'm really happy to hear this, {}.*Feel free to reload the page and re-enter ... | Did you make any mistakes? Little mishaps in your everyday life. | The other distinct way to be more playful is by having a PLAYFUL FACE; to do ... |
| It's good to talk to you, {}. I will help to teach you to laugh.*Are you aware of ... | That's good. Just so you know, they are also being shown on the right of the ... | Nice to hear it, {}. You can reload the page and enter your details in again ... | Have you made any silly mistakes recently? Nothing too major, just in your ... | Another way to become more playful is by loosening up your mouth and eye ... |
| Great to talk with you, {}. I will try to help you learn to laugh.*Are you familiar ... | Okay, just so you know, they are also being displayed on the right of ... | I am happy to hear it, {}.*Whenever you're in the mood, just reload the ... | Do you remember making any mistakes or errors lately, just something ... | Although it sounds obvious, many people overlook the fact that you can be ... |

Figure E.2: List of statements output by the optimisation function in the case where the user is displaying a positive emotion.

Negative retrieval function

| Rewritings retrieved for statement 1 | Rewritings retrieved for statement 2 | Rewritings retrieved for statement 3 | Rewritings retrieved for statement 4 | Rewritings retrieved for statement 5 |
|---|---|---|---|---|
| It's good to talk to you, {}. I will help to teach you to laugh.*Are you aware of ... | They are being shown on the right of the screen in my browser view if you ... | Nice to hear it, {}. You can reload the page and enter your details in again for ... | Have you made any mistakes lately in your daily life? | To have a PLAYFUL FACE you can relax your mouth and eye muscles while ... |
| Nice to talk to you, {}. I will do my best to help you learn to laugh and have ... | If you wish to review them now or during our conversation, you can ... | Glad to hear it, {}. If you want another chat, all you have to do is refresh the ... | Have you made any mistake recently in your day-to-day life? | Consider having a PLAYFUL FACE; you can loosen up your mouth ... |
| Great to meet you, {}. I'll do what I can to teach you how to laugh.*Are you ... | Great. You can check them out anytime if you need to. They are shown on the ... | I'm really happy to hear this, {}. Feel free to reload the page and re-enter ... | Are there any mistakes you have made lately that affected your life on a ... | Although it sounds obvious, many people overlook the fact that ... can be more |
| Thanks for coming, {}. Let's hope you leave as a laughter professional!* ... | The protocols are on the right hand side of my browser view. Have a ... | Great to hear you say that, {}. You can reload this page and enter your details ... | Do you remember making small mistakes in your everyday life?*I ... | An alternative way to become more playful is to have a PLAYFUL FACE. |
| Nice to talk with you, {}. My goal is to improve your ability to laugh, which ... | As you might have seen, they are shown on the screen to the right if ... | That's great, {}. I would too! *You can reload the page and enter your ... | Have you been making any errors or blunders recently? I'm not asking for ... | Having a PLAYFUL FACE is important for being in a playful mode. What's ... Let |
| Nice to meet you, {}. Hopefully, you can leave this session with useful ... | Just to let you know, they're showing to the right of my browser view if ... | Great to hear it, {}. Whenever you are ready to speak again, give the ... | Have you made a mistake or blunder recently? It doesn't have to have ... | That wasn't the only way to be playful - another equally important way is to have ... |
| Okay, nice that you are here, {}. I will try my best to help you with your ... goals. | If you wish to remind yourself of any of these protocols during our ... | Glad to hear that, {}. When you feel like it, go on and reload the page and re-... | Any recent slip-ups? Nothing too big, just something in your day-... | To be more playful, it helps to change the way your face moves, to have a ... |
| I'm glad to speak with you, {}. I try my best to help teach you to laugh. ... | They are on the right of my screen if you're in my browser view. Just in ... | Love to hear it, {}. If you want, you can reload the page and we can have ... | Have you had any small mishaps recently? Made any minor errors? | So you can also become more playful by having a PLAYFUL FACE; if you ... |
| Lovely to speak to you, {}. I tell you what; I will try my best to teach you to ... | They are shown on the right of the screen if you're viewing me on your ... | Great to hear that, {}. For another talk, you can just reload and re-enter your ... | Have you made any small mistakes or recently become aware of minor ... | If you wish to become more playful, I'd also recommend having a ... PLAYFUL |
| It's good you're here, {}. I will teach you about learning to laugh and ... | Just to let you know, they're showing to the right of my browser view if ... | I'm really happy to hear this, {}. Feel free to reload the page and re-enter ... | Have you made any past errors lately, if you're happy to admit it? It doesn't ... | If you can loosen up your mouth and eye muscles, this is another way of ... |
| Glad to talk, {}. If you are here to learn to laugh, you came to the right place! | As you may have already noticed, they are shown on the right in my browser ... | Glad to know that, {}. Feel free to reload the page and put your details in again ... | Can you recall any errors or mistakes you've made lately? I'm not talking ... | A PLAYFUL MIND is important, but to be even more playful, you need a ... |
| It's lovely to chat with you, {}. I will do my best to teach you how you ... using | Just to let you know, you can also find these on the right of the screen in ... | I would as well, {}! Just reload the page and re-enter your credentials ... | Are there any small errors or blunders you've made recently in your day-to-... | Having a PLAYFUL FACE will make you more playful, as I'm sure you can ... |
| Nice to meet you, {}. I will help you to learn to laugh and guide you through ... | They are currently showing on the right for you to browse whenever you ... | Aw, I'm very glad to hear it, {}. If you want to chat again, feel free to ... | Have you made any mistakes recently? Nothing really important, just in ... | Another way to become more playful is to have a PLAYFUL FACE; by ... |
| Pleased to meet you, {}! I'll try my best to teach you the lesson of laughter. ... | If needed, you can always see them in the protocol viewer on the right of ... | Yay - I'm glad to hear that, {}! Feel free to have another chat by ... | Have you made any small mistakes or recently become aware of minor ... | The other distinct way to be more playful is by having a PLAYFUL ... |
| Nice to talk to you, {}. I can offer assistance and guidance to help you ... | If you're in my browser view and wish to review them, then you should ... | Great, I would love to as well, {}! You can reload the page and re-enter your ... | Would you say you've recently made any errors, just little things, in your ... | Although it sounds obvious, many people overlook the fact that ... can be more |
| Great to talk with you, {}. I will try to help you learn to laugh.*Are you familiar ... | By the way, they are on the right of the screen in my browser view. So you ... | I'm happy to hear that, {}. You can reload this page and enter your ... | Have you made any mistakes or errors recently? If it makes it easier, I'm ... | You can also become more playful by having a PLAYFUL FACE; by ... |
| It's good you're here, {}. I will teach you about learning to laugh and ... | Great. You can check them out anytime if you need to. They are shown on the ... | That's really great to hear, {}, thanks!*If you wish to have another ... | As much as we try, our lives can never go as perfectly planned as we'd like ... | Having a PLAYFUL FACE is also a way to become more playful, and it's an ... easy |
| Nice to meet you, {}. I will do my best to teach you how to laugh today. It's ... | The protocols are on the right hand side of my browser view. Have a ... | That's great to hear, {}. Feel free to reload the page and we can have another ... | Do you remember making any mistakes or errors lately, just something ... | To be more playful, it helps to change the way your face moves, to have a ... |
| Great to meet you, {}! I will try my hardest to help your nurture your sense of ... | You can find them on the right of the screen in my browser view, in case ... | I am happy to hear it, {}. Whenever you're in the mood, just reload the ... | Are there any mistakes you have made lately that affected your life on a ... | So you can also become more playful by having a PLAYFUL FACE; if you ... |
| Great to speak with you, {}. Well, I'll be doing my best to help you learn to ... | Just to let you know, they're showing to the right of my browser view if ... | Very pleased to hear this, {}. Feel free to reload the page and re-enter your ... | Do you remember making small mistakes in your everyday life?*I ... | If you can loosen up your mouth and eye muscles, this is another way of ... |

Figure E.3: List of statements output by the optimisation function in the case where the user is displaying a negative emotion.

Appendix F

Evaluation feedback

Trial 1 questionnaire responses

We list below, trial 1 participants' written feedback.

Please describe the overall educational impact (on your knowledge of laughter and humour) of your experience.

- I never thought there could be a humor protocol which actually works.
- Taught me some theories and methods that I didn't know before.
- It definitely make it easier for me to cope with difficult situations.
- It helped me classify different types of humour that I hadn't explicitly thought about before
- I got reminded about what Niche said about the past and the future and being in the present moment.
- Where previously I would not have explicitly thought about laughter and humour in my interactions, after using this chatbot, I am able to recognise the protocols that I am using naturally.
- I didn't know about any of this theory before the chatbot. I practiced some of the theory unknowingly but it still taught me a lot.
- The overall experience was instructional and helpful. However, an important part of the educational effect for me was the actual reading of the article on self-initiated humorous protocols. The chatbot was highly helpful as a motivation and structure to practice the protocols. It was also partly helpful as a teaching tool as to what types of protocols do use and when. But I feel I got more of my actual learning out of my careful reading of the article and my previous experience with these protocols as a part of SAT.
- It was interesting to learn about the different ways to find laughter in life.
- I only completed two sessions but found both to be useful
- I learnt about the ways in which laughter can be used to make situations less daunting. It can be used to prevent anxiety inducing situations which I find interesting.
- It really has allowed me to understand the different types of laughter that occur

in different situations.

- It seems that the overall flow is very engaging, and prompt me to laugh. I personally enjoyed it.
- The laughing exercise were quite enlightening, but I found it confusing when or how the bot wanted me to apply them.
- It was a good exercise, I read the attached document, I learned about the theories about laughter which was interesting.
- quite informative
- I learnt about the benefits of laughing in circumstances where it would be otherwise unusual to laugh, for instance laughing at setbacks/misfortunes or routine daily rituals. I was particularly fascinated to learn how we can use laughter to improve our perception of our misfortunes in life, an example of where the HaHA has given useful life advice.
- This is excellent, thank you! It is tackling an important life issue with a simple important solution. It is effective in using an interesting set of examples and history lessons and links them well to everyday life (and therefore might well be targeted at a higher education level participant).
- N/A
- I felt that I learnt a lot! (I had no real knowledge about the different techniques etc that were described and explained before I completed the trial!)
- In a couple of instances, when asking about learning more on how to laugh using a technique (by clicking the 'smile' emoji), the chatbot kind of restarted the conversation and went in a different direction. I expected to learn more here, so was a little disappointed. There were a few technical issues in starting but that was soon resolved.
- I learnt about the different types of humour and laughter.
- I hadn't known of many of the protocols before, so it was interesting to learn of them.
- I was able to learn more about laughter which was interesting. I learnt a lot from reading the PDF on the right of the screen
- I never took the time to reflect on the type of humor I was using in every personal struggle or situation because it was a coping mechanism. Learning about the different modes of laughter and humor was eye opening and educational. I especially liked the instructions to try out a scenario allowing users to do their own homework on reading the protocol or understanding how the action would improve the mood. I do believe that maybe an audio or visual representation of the different protocols would be more engaging.
- Good to help people with social anxiety and people with learning disabilities who wouldn't know when to laugh eg high functioning autistic people
- I learned some ways to become happy and the protocols are helpful to make me laugh.
- Learning about the humour theories and trying the exercises made me think about what I find funny and that was quite interesting
- The platform has greatly furthered my knowledge of laughter and humour. It has increased my awareness of the variety of ways in which laughter and hu-

mour can be used in ones' daily life. I have a greater appreciation for how knowledge of laughter and humour can be utilised to build resilience and to develop a more relaxed, and positive outlook on situations and life experiences.

- Helped me realise how to help myself laugh and be more humorous.
- The experience has certainly helped me laugh more through difficult situations..Overall see the funny side to situations.
- so important to relax and maybe laugh at little things of life however in my humble opinion even false laughter alone can be a factor effecting well being of mind and attitude
- The protocols were interesting and reading through them was a positive experience. I felt this was more engaging however when delivered in smaller chunks by the chatbot than when directed to read through in the panel on the right.
- A great way to learn about humour protocols in a digestible manner
- The chatbot was useful and taught interesting thing about theory of laughter and humour
- I feel the chatbot can be a little unclear at times of what exercise it is referring to, there were multiple instances where it had expected me to do something i.e. Would it be possible for you to laugh this off?", but it's not clear which exercise I should perform. I had to read through the protocols multiple times to get an idea of which exercise it might have been referring to. I think this is a great idea, but would be better if it had something to teach me to practise the exercises individually or a clearer indication of exercises during the conversation.
- Great in shedding light on the different protocols used for humor. With all the protocols an example was given or it was explained with application. In relation to my experience, I was not aware of the different types and thus this study if good for widening horizons into an area that wouldn't necessarily have been touched upon.
- I have learnt some new concept on how to laugh and new level to make myself laugh.
- I learned new interesting exercises that I will try using later. The idea of using a chatbot for teaching has potential but at the moment it is difficult to have the same impact as a real person. However, the chatbot is much more impactful in learning experience rather than reading through the material.
- I learnt quite a bit about laughter and humour.

Please describe the overall emotional impact of your experience.

- It was quite enjoyable, definitely a nice relaxation from my project work
- Overall it's great, and do help me get out of the mood one day during the trial.
- Using Laughter protocol helps me lessen my daily stress and make me feel happier.
- It was a nice experience which taught me more about humour
- I liked it in the beginning but got tired of each time introduction part. I felt I

- don't want to use it anymore as it became boring, specially in the introductory part, same jokes and too much asking for confirmation, Continue button.
- I felt satisfied that I had learned something new about humour and myself.
 - It made me realise situations where I could laugh it off more.
 - It was very positive. Laughing more and on more occasions definitely improved my emotional well being. The platform was also fun to interact with though there is room for improvement.
 - I was surprise to find that I was laughing while interacting with the chat bot
 - I found that the experience really helped to prompt me into thinking about what usually affects my mood and how I can handle it better. It was light hearted and simple to use which makes it less daunting
 - it was great, I'm glad I took part.
 - I like it.
 - It was a good way to unwind after a long day. I believe that the interaction was genuine enough and can take your mind off stressful thoughts, especially when combined with the laughter protocols and exercises.
 - It made me a little bit better during the exercises, although I was not be able to try the exercises during the dat, but overall it was a good experience.
 - relaxing , feeling happy
 - Overall, trying each laughter exercise led to a temporary uplift in mood. The humour protocols have also introduced a playful, relaxing element to key aspects of life (such as routine habits, setbacks, debates. In particular, because of the light-hearted method of communication I was able to clearly understand the protocols and feel confident I can practice these in my daily habits.
 - Made me laugh. Inspires to continue the practice. Would go back to the chatbot to do it in future if it was available
 - N/A
 - It's certainly made me think a lot more about laughter and humour and how I can use it even in tough situations.
 - The chatbot is quite humorous in what it says overall. However, doing this over multiple days becomes repetitive and the same messages start to feel a bit less funny after a day or two since one knows what the chatbot is going to say - maybe randomising messages or even having a log of how a previous conversation with the user went would be good.
 - positive
 - I personally found the experience of going through the physical actions of laughter, without an associated mood of humour, worsened my mood rather than improved it. It's possible, of course, that over time there would be a more positive effect. But I felt that the 'artificiality' of the process disposed me to feel irritation.
 - I was often left neutral after using the chatbot. However, later on in the day I thought about the techniques described by the chatbot and I tried to use them, so at the end of the day it did have an impact. Although it was not huge, in my opinion.
 - I think the humor was light and easy which took off a lot of pressure. I also

liked that the questions made me think about past experiences because often times when battling a hardship we dont smile or appreciate ourselves for getting through it but we do so afterwards, after some time has passed. Having that recall moment was an important trait in this conversation because it reminds users that now might be that tough time but instead of waiting to appreciate yourself AFTER getting through it, do it during the battle. I thought the conversation flowed well but it was limited due to 3-4 options of mood. I think if there were more moods and more diversified conversations it could be more meaningful but I understand this is a preliminary study. I wonder how people who are more numb to this light sense of humor will be able to tackle personal issues through use of the bot. I wonder if the type of humor being used to convey important protocols is important in getting the message across. For example I found this exercise to be light and funny and it improved my mood, but I know that maybe other people I know may not see the humor and light heartedness because they've been so numbed and shaped by only struggles for a consistent majority of their lives. This is not to say however that I as an individual have not struggled, because oftentimes it is those feelings of tired or sad or angry that allow our hearts to be more vulnerable therefore more open to the effects of light humor. Overall this study is a beacon of hope in the loneliness that adulthood brings and I hope it will be developed into a depression hotline of sorts in order to show people they are not alone.

- Great, it was good fun, with more combinations to responses it could help people with their sense of humour and help people who are lonely or far from family
- I have not experienced any bad things recently, so the chatbot does not improve my emotion that much. But I believe it will do a good job when someone is in a low mood.
- Thinking about humour and trying the laughter exercises put me in a playful mood and made me more likely to laugh at a joke (even if it was very bad and not all that funny) afterwards. I was mostly in a good mood for the past few days though, I have no idea how it would have worked if I had been feeling sad or stressed instead.
- The platform had a very positive impact. I have really enjoyed engaging with the chatbot. The jokes made by the chatbot were humorous. In particular, the language used by the chatbot was very effective in creating a comfortable setting to explore personal experiences and humour together.
- After using the techniques I felt more cheerful and a lot happier. I felt more relaxed and more at ease..I quickly learned to see the positive side and also be more accepting of all situations.
- good maybe
- Overall the experience felt light-hearted and warm and reminded me of the positive impacts that humour can have in self reflection. However at times I felt the chatbot could be perceived as callous (asking one to laugh at a misfortune without much attempt to understand what that misfortune may be). Perhaps this could be remedied if the chatbot were to acknowledge it doesn't intend to

- be callous or phrase certain suggestions in a softer manner,
- Positive
 - I appreciated that when I did not feel ready or comfortable doing an exercise, the chatbot did not insist or got stuck there but offered that I could do it later in my own time
 - Fine overall, I feel there were times where the chatbot can come across as a little overwhelming from every single sentence having a humour component to it, but all slightly different/lacking coherence between sentences. I think it might have been better to have jokes at some specific points (punchlines) or maybe chain the jokes to be a little less ramble-y.
 - When humor was used by the bot, I would react to how funny the statements were! I think in some instances it would make me smile. Overall a good experience!
 - I feel very joyful after conversing with the chatbot, Other than discovering the roots of my emotion and learn how to laugh on it, I am also really interested in more about the jokes that the chatbot will tell. I also really love to be able to interact with emoji icon instead of just words.
 - Overall, I felt better after practicing some of the exercises for short time. I guess if I keep practicing the effects will last longer. The chatbot made me feel good during the procedure. Some of the exercises worked but some others didn't.
 - It made me laugh and improved my emotional state.

Do you have any further suggestions for how the platform can be improved?

- Could ask more questions maybe? I do enjoy this process!
- The chatbot could maybe provide examples of the types of humour it suggests
 - *1- Too much ".Continue".
 - 2- If I want to use this platform daily or several times per day, I don't like to introduce myself to it every time or vice versa
 - 3- The name of the exercises where not the same as in PDF file. e.g. Feigning laughter"
- The protocol viewer could be more graphically presented and more integrated into the chat.
- More jokes would be nice ;) *Integrate simple, step-by-step teaching of the protocols into the chatbot. As it is now, the chatbot only mentions the protocols not teaching them. It also uses technical jargon (albeit in ALL CAPS and with the offer for more info). It would be ideal if the chatbot can walk us through understanding these protocols in simple terms as opposed to just naming them and referring the user to the PDF (which in itself is largely technical in language).
- For capturing our emotions and moods, it would be great to include a more nuanced scale. The three choices of 'better', 'no change' or 'worse' seemed to me to be inadequate in helping my self evaluation or in guiding me to decide

where I need more practice.

- I really liked the genuine attempt at making the chatbot itself humorous. I feel it was a little overdone though. Also, the flow of the conversation would often distract me from the main point or task. The flow was often like this: Ask a question > Provide lots of explanations/make a joke > Give the multiple choices for my response.
- I would often have to go back and find the question I was supposed to answer, or else, avoid reading the rest so I could focus on the question. I understand and appreciate the attempt at showing empathy, providing context, and making a joke. I don't mean these were unhelpful; what I suggest is to improve the balance."
- It was very easy to use and intuitive. Being able to use it on your phone was great because it makes it easily accessible and portable
- nope!
- It would be much better if you could develop an app which reminds the exercise every day and goes through the methods with the applicant.
- not really *Sometimes a serious question was asked, followed by a joke in the same speech box. This made me forget the question when prompting me to answer - perhaps, where applicable, HaHA might be able to let the user answer and then follow up with the joke
- At the end, if I chose to start another session I expected HaHA to start one for me, but I was told to reload the page and login again - it might be worth seeing if there is a way of automating this so in such circumstances the user can jump straight into the other session without entering their name again.
- Even though the process of responding to prompts (such as 'Yes/No' or a 'Photo seen') was intuitive, perhaps it might be worth disabling the chat feature after showing a prompt to further improve clarity even further for users
- would suggest adding a path that provides for more structure to learn each of the techniques, a more serious reply (more empathy) when someone says they are having a sad/bad time, and add the learn more option to a technique the chatbot is taking one through.
- Note: when it gave prompts, the prompts were not complete and did not allow for text instead of choosing one of the options provided.
- N/A
- Sometimes, people may not want a light-hearted, humourous conversation and it would be cool if the chatbot could pick up on that in their replies and adjust its tone to reflect.
- Please see earlier comment
- the chatbot made me giggle several times, so it certainly works
- Leaving the session and returning another day meant I had to repeat some things, e.g., giving the system my name, and reading / answering basic material. It would be good to record that someone had logged in before, and where they had progressed to. I also found it quite common that the none of the 'multiple-choice' answers on the online form represented what I actually felt or thought.

- It would be nice to see a face, humanize the chatbot a little bit more. Maybe even let it read the content to you? And perhaps that is too complicated, but you could speak back to it and have a real discussion.
- More varied responses, more emotions. Jokes based on location or region as I could see people all over the globe using it.
- The emotion identification sometimes does not work very well.
- I have the impression that the chatbot goes over the protocols in the same order every time. That gets a bit boring after a couple of days. It also repeats the same things a lot.
- I was expecting the chatbot to remember my name from one session to the next (since I was using the username and password provided) and was disappointed that it did not.
- The chatbot is very hard to use on mobile, both in portrait and landscape modes.
- There are a few typos - for instance, the chatbot introduces itself as an "Assistant" rather than "Assistant" (I wasn't sure if it was supposed to be a joke or not, but it looked a bit unpolished).
- In future, the option to choose which protocol to investigate could be given earlier on so that an experienced user can navigate straight to their desired task.
- In future, the time which the user says that they have available to explore the platform could be used as a parameter to tailor the depth in which the chain of information for each protocol continues for. For example, if a user states that they have 20 minutes to use the platform, they may want to explore and understand a protocol for longer than a user who only has 10 minutes free. This may allow the chatbot to provide more detailed explanations to a 20 minute user and more short, concise explanations to a 10 minute user.
- In future, short, animated clips could be an interesting way to give an example of a scenario which the user could find humour in (where applicable to certain protocols). This would be particularly useful for users with reading difficulties, and in general to break up the text.
- In future, questions related to the user's opinion on the chatbot's guidance could be in the format of 'typed answers' instead of 'options', to enable the user to provide a more detailed response of their experience.
- The platform is amazing and with little need for improvement.
- maybe to find and describing some natural but funny life events , for example , a lonely dog looking out to the right and left crossing a busy road or a cat being chased by a rat !
- Initially I replied to the chatbots greeting conversationally, however this meant that for the first session the bot thought my name was "HI, I'm Matthew". Although this was quite funny. The jokes also became a bit repetitive and clumsy, although this was in itself funny at times, I think the chatbot could benefit from having a few more jokes at hand.
- There have been a few times where the same joke was repeated again shortly after the first instance of it. Also, sometimes the bot asks 'are you ready to

do this exercise' after it has said quite a few other things since explaining the exercise. Sometimes this can be slightly confusing as one may not remember/understand which exercise is being discussed

- Clearer indication of exercises. More coherent humour or punchlines.
- Reducing formality of the language used by the bot, although I can understand the opposing argument too! So overall not much to say.
- I feel there should be a way of tracking what have done by the user as some protocols is time consuming and we would not be able to get back to conversing with the chatbot soon enough. Also, the chatbot suggests exercise based on question, it would be good to know about some information as to why I should do this exercise before I choose to continue,
- Instead of having the exercises as a separate document it would be nice to be incorporated in the dialogue. It would be good if the chatbot kept in memory that I used it before so the starting conversation is not repeated or with different wording and shorter.
- The suggested responses and the dice to type can get confusing. As in which one is meant to be used- the buttons and/or typing answers.

Please provide any additional comments you have.

- Nice work!
- Nice job!
- I completed 4 sessions on the chatbot
- Once I made a mistake by typing my username with capital U instead of small, then after it gave me the error message and asked for typing again, no matter how much I refreshed the page and entered the correct username & password, it didn't continue to the introduction part. It seemed it has frozen.
- Thanks for the great work. I look forward to seeing the future versions and participating in the next stage.
- Really enjoyed using it! I think for people who struggle with anxiety or over-thinking it could really help to guide them to handle their emotions better
- Thank you.
- I interacted with the bot a total of three times. When providing my emotion to the chatbot, I claimed to be "fine/okay" two times and it guessed my emotion as "sad".
- It was strange for me to be able to laugh at the past sad experiences, and still I don't know how it is possible. may be I misunderstood some expression from document, or I didn't get it at all.
- maybe funnier short stories can be used
- I overall formed a happy, cheerful perception of the HaHA, and although the (inevitable) repetition of phrases gave a slightly robotic feel, aspects such as linking back to what we covered when I end a session gave me the feeling it was designed in a very sophisticated way.
- Bien fait!
- this approach seems suited to many life learning lessons, worth launching as an

app in a series: eg laughing, crying, silence/pause, de-escalating, mindfulness etc

- I completed 2 sessions with the chatbot
- Really interactive and provides valid responses that match the context! Really cool!
- It would be good to have some history with regard to the conversation, otherwise repetition can't be avoided and that is not good. It would also be good if upon having this history the chatbot is able to suggest on its own in some ways the next set of steps too.
- the examples could perhaps be more detailed
- Great work! Very interesting topic and I like the idea.
- The haha chatbot was excellent! Really fun and was so easy to use. This in app form would go down a storm. It would be used by a wide audience!
- I have an eating disorder and didn't enjoy the frequent references to cake and low sugar and so on. It didn't trigger me but it bothered me a bit.
- It was a fun exercise at first but it did get repetitive and a bit boring really fast as the chatbot was often reusing the same sentences. The protocol document was also really long and looked intimidating, and it wasn't really necessary since the chatbot explained things really well.
- The language and as a result, the personality of the chatbot which was created, was informative, humorous, and sincere.
- The platform was user-friendly and easy to navigate.
- The examples of real-life scenarios which were used in certain protocols was effective in clarifying the protocol task.
- The mix of typing one's answers and selecting from the options provided, made the platform engaging and interactive.
- I have thoroughly enjoyed exploring the platform. The language and personality created for the chatbot is uniquely kind, informative, and effective in teaching its users about laughter and humour.
- The experience was most enjoyable very easy and fun to use.
- practicing to laugh physically can help temporary but some other factors like having a belief and meditation could also add to a deeper effect and overall longer state of health
- Although the chatbot felt empathetic in its expression overall, this experience was more therapeutic when I was given the opportunity to type my interactions. Clicking through multiple choice and yes/no answers started to feel a bit mechanical after a few sessions.
- I'm so sorry if I sounded harsh in my feedback, I really like the humour protocols and I did feel that practising them has had a positive effect on my mood. But I felt taken out of the conversation when I couldn't keep up with what exercise I was supposed to be doing, and felt a little overwhelmed. I think this is a very interesting platform, and definitely some very good work, I just think that perhaps I'm not quite on the same wavelength. I'm very interested to see where you take this concept in the future!
- Thanks for the trial! Very educational even with a process (laughing) that we

do everyday!

Clinicians' questionnaire responses

We list below, 4 clinicians' written feedback (from trial 2).

Please describe the overall educational impact (on your knowledge of laughter and humour) of your experience.

- It is very well written and easy to digest. It is quite informative.
- It was quite instructive, however, it would be more informative by providing a piece of wider information on the philosophical and psychological background of laughter and humour.
- Moreover, more practical examples could facilitate the imagination of the past/current experience. This could be considered subjective practice before recognizing similar situations and using the techniques.
- Difficult to describe on the basis of one experience. I am also familiar with the approach so the novelty was not influential. I think benefits would increase with repeated practice. I liked that I was encouraged to associate a novel and potentially positive response to negative emotions or memories.
- Very natural guidance, without feeling forced to perform
- Reasonable learning on laughter science, but could have been clearer and more digestible.

Please describe the overall emotional impact of your experience.

- I found HaHA exciting and interesting.
- I found it transformational! As claimed, it created a more positive outlook, and therefore, a more elated mood.
- It's great, especially as I became familiar with the various types of laughters, and it went better
- Not significantly emotionally impacted, as did not feel drawn to engage deeply in laughter techniques by the technology (with clear understanding of purpose and protocols).
- Difficult to describe as I was role playing being sad for the purposes of exploring the system! I found it diverting and playful nonetheless.

Do you have any further suggestions for how the platform can be improved?

- Yes, the initial response to "How are you feeling today" is limited. Whatever I said other than being happy was interpreted as "sad".
- It was quite a basic interface, very text based. Maybe some more audio visual or graphical enhancement.

- Instructions long - in my opinion those likely to enjoy using a chatbot won't read this!
- There may be room for improvement in the recognition of neutral emotions, and when replies such as 'not bad' are expressed, with emotions being recognised as sad when replies such as 'not bad' are expressed.
- Language sometimes jargon heavy and technical, e.g. Do you think your negative feeling could be caused by a specific incongruity? Not a natural flow of conversation, e.g. single comment, with a pause, then another question; felt slow as the user Laughing exercises, e.g. Duchenne exercise and Laughing at long-term suffering not introduced / structured with specific instructions/timers/etc (expected to read the text)

Please provide any additional comments you have.

- HaHA could not respond to additional questions other than the planned ones. Furthermore, it was not possible to return or ask again until the techniques were completely stated.
- The conversations were humorous and uplifting. I found myself laughing even before I started to practice each protocol. Well done!
- UI not visually appealing UX not always intuitive, e.g. unclear can't type when a Yes/No button is offered (cursor still flashing)
- Chatbot tries to be too funny - Interpretation of 'how are you feeling?' was very crude
- Conversation ended when the laughing at suffering felt offensive and I answered no
- The possibility of visual interaction and practicing the techniques during interaction with HaHA can increase the impact of learning.