

Developing Chatbots as Virtual Clients for Empathy Training in Counselor Education: Examining Feasibility from Users' Perspectives

Lih-Horng Hsieh¹, En-Yu Liu², Chi-Min Lin³, and Chia-Jyi Liu⁴

The purpose of this research is to develop chatbots as virtual clients (VCs) for empathy training, and to explore the feasibility assessment of using chatbots to assist in empathy training for students in counseling and guidance-related departments. The virtual clients are developed and tested based on chatbots using domain knowledge of empathy training. Chatbots are powered by natural language processing engines and trained by corpora of dialogues involving various emotions. As such, using chatbots serving as VCs allows the interlocutor to practice empathy skills with diversity and richness. After chatbots are successfully transformed into VCs, 60 volunteers from the Department of Guidance and Counseling are recruited to experience the interaction with VCs, followed by immediately conducting user experience surveys to understand users' subjective feelings and thoughts for evaluating the ease of use and effectiveness of VC. Chatbots are successfully transformed into VCs by implementing corpora of emotions into "custom brain" of chatbots. The content of corpora includes client scenarios that are consistent with the goals of empathy training for undergraduates in guidance and counseling or psychotherapy. Test results show that the average score from the User Experience Questionnaire of the testers exceeds level 5 (1-7), which is a state of satisfaction. In terms of System Usability Scale, the average score is 4 (1-5), indicating agreement with the system's usability. Along with qualitative feedback, testers agreed on the usability of the chatbot system, stating that most people can quickly learn to use the system, interact with the virtual client (VC) without stress, feel relaxed, and want to continue the conversation. They believe that VC is suitable for practicing empathy skills, reducing anxiety during practice, and improving empathy response skills through such practice. From the observation of the testers during the test process, the testers showed curious and interested eyes when they saw the VC, and they enjoyed practicing empathy with the VC. This research successfully transformed the chatbot into a VC, which can be used as an assistive training tool for empathy practice. In addition to increasing practice hours, willingness, and opportunities for empathy training, using VCs allows prospective counselors to avoid ethical problems and

¹ Department of Counseling and Guidance, National Changhua University of Education

² School of Medicine, College of Medicine, Taipei Medical University

³ Data Science Degree Program, National Taiwan University

⁴ Department of Physics, National Changhua University of Education

Corresponding author:

Chia-Jyi Liu, Department of Physics, National Changhua University of Education. Email: liucj@cc.ncue.edu.tw
This work was supported by National Science and Technology Council of Taiwan under the Grant No. MOST110-2410-H018-020.

pressure caused by directly facing the real-life client in the learning stage of developing counseling skills. Furthermore, utilizing VCs as a training auxiliary tool will help solve the teaching dilemma of insufficient case sources and enable the continuation of counselor education in situations such as future recurrences of epidemics similar to the COVID-19 pandemic.

Keywords: empathy, smart chatbots, virtual clients, empathy training courses, counselor education

When it comes to training counselors, the acquisition of counseling skills and the development of their abilities are crucial aspects of their education. However, this training can be challenging and time-consuming, particularly when there is a shortage of training and supervisory staff, as well as a limited number of client simulators. In such circumstances, students undergoing training may unintentionally compromise the rights of their clients due to their insufficient professional expertise, which is a concern in terms of professional ethics. Consequently, many instructors teaching counseling skills must rely on theoretical concepts from textbooks, share their own practical experiences, or engage students in role-playing exercises where they simulate the roles of both counselor and client, while receiving group supervision. Alternatively, they may seek permission for their students to conduct counseling internships under curriculum supervision. Unfortunately, these methods often restrict the training process, which requires extensive practical experience and internship hours, thus impacting the overall effectiveness of the training. Although there have been recent advancements in training smart robots to act as counselors, there are still technical and professional obstacles to overcome. Conversely, exploring the possibility of developing smart chatbots to serve as client simulators in counseling education could be an intriguing area of study.

In the process of nurturing and training counselors, the concept of empathy holds significant value. Carl Rogers, the founder of the client-centered approach in psychology, firmly believes that empathy is an innate human characteristic. Through empathy, individuals can effectively communicate, establish positive relationships, and foster a state of mutual emotional support. Rogers (1957) goes so far as to consider empathy as a crucial element in the success of psychotherapy and counseling. As a result, empathy has gained prominence as a vital attribute in assisting individuals within the field of counseling and psychotherapy, playing a pivotal role in the counseling relationship (Barrett-Lennard, 1981; Rogers, 1957, 1979). Irrespective of the theoretical orientation in counseling, empathy has been recognized as a shared and influential factor in the effectiveness of the counseling relationship. Extensive research has been conducted to explore the dynamics between counselors and clients during the counseling process, with the majority of findings emphasizing the significance of the counseling relationship. By employing empathy skills, counselors can establish a trusting connection with their clients. This allows clients to feel understood and encourages them to openly discuss their problems, facilitating problem-solving and resolution. Hence, it is evident that empathy skills play a substantial role in the counseling process and should not be overlooked in the training of prospective counselors (Song et al., 2004).

Chen et al. (2014) identified a dilemma concerning the utilization of empathy by counselors based on their shared practical experiences. In Taiwan, new counselors or students studying counseling often encounter challenges in comprehending, recognizing, and effectively responding to their clients' emotions when learning and employing empathy. Consequently, establishing a close connection with clients becomes difficult for them. Several literature reviews (Kurtz et al., 1985; Matarazzo & Patterson, 1986) evaluating the impact of counseling skills training on counselors have indicated some progress in one or more counseling skills, including empathy, following short-term training. However, these literature reviews also highlight that the effects of such training are not sufficient to be generalized to real-life counseling scenarios. In other words, the effects of counseling skills training are often more noticeable in theoretical assessments but do not accurately reflect the counselor's performance in actual counseling sessions. These findings align with the researcher's extensive teaching and training experience of nearly 30 years. As an educator and trainer in the Department of Guidance and Counseling, involved in counseling education at the bachelor's, master's, and doctoral levels, the researcher has instructed courses on group counseling, group counseling internships, individual counseling internships, and advanced counseling training. During these experiences, it has been observed that undergraduate students possess a theoretical understanding of empathy concepts and techniques, yet frequently encounter difficulties in applying empathy when practicing individual counseling or leading counseling groups. The encountered challenges often manifest in situations where empathy is required, such as forgetting to express empathy when necessary or demonstrating only

a rudimentary level of cognitive empathy, failing to exhibit more advanced empathetic responses. These circumstances prompt us to question why empathy training conducted in a classroom setting does not effectively translate into practical application during counseling or group counseling sessions.

Empathy plays a crucial role in establishing the counseling relationship and significantly impacts the effectiveness of counseling. The current absence of purposeful empathy training tools has led researchers to reflect on the issue of inadequate practical experience or training hours in existing empathy training programs (Foster et al., 2016; Frangoudes et al., 2021). They aim to propose feasible methods to enhance these courses and address the insufficiencies in training.

With the rapid progress of artificial intelligence (AI) and the advancement of software and hardware in the field of machine learning, such as the notable improvements in high-performance graphics processing unit (GPU) technology, the speed of matrix and numerical computations has significantly increased. The widespread availability of deep learning frameworks has also made the development of deep learning applications less daunting. In the future, counseling educators may utilize computer-assisted systems to teach counseling skills and identify effective conversation techniques and language used in counseling. Alternatively, computers can simulate clients for counseling practice sessions (Sharf, 1985; Shorey et al., 2019). Furthermore, technology educators have effectively utilized the application of robots, based on technology and programming, to train students regarding skills in specific areas (Lu & Ma, 2019). By leveraging these advancements, students can engage in repeated practice sessions, compare their performance, and receive feedback, thereby eliminating the challenges associated with limited internship clients, professional ethical constraints, and insufficient practice hours.

Foster et al. (2016) conducted a study utilizing virtual patients (VPs) to teach empathy to medical students and examined its effects. The findings indicated that medical students exhibited increased empathy responses when interacting with standard patients, as demonstrated by the feedback on empathy provided during their interactions with VPs. Moreover, research has shown that VPs have the ability to standardize content, enable repetitive practice, and facilitate empathy training in safe and comfortable environments. Incorporating an empathy dimension into VPs would facilitate the integration of this technology into healthcare professional training courses. In addition, Lee and Lin (2021) summarized in their study the principles and applications of theatre-based training to teach empathy, which provide valuable insights for educators to employ methods based on interactions for empathy training.

Building upon the aforementioned context, researchers aim to create intelligent chatbots as supplementary tools for empathy training courses. Our objective is for aspiring counselors to engage in empathy dialogues with chatbot virtual clients (VCs) using spoken language instead of text input. Additionally, we seek to explore the feasibility of using chatbots as auxiliary tools for empathy training in counseling, as evaluated by students in counseling-related fields.

Theoretical Framework

Definition and Classification of Empathy

Empathy is an essential social skill necessary for establishing and strengthening relationships with others (Yang, 2018). Experiencing empathy can evoke both sadness from perceiving others' pain and happiness from sharing in their joy. It involves the ability to "put oneself in someone else's shoes" and "feel the same emotions" (Teng, 2011; Yang, 2018).

Literature suggests that empathy is a multidimensional construct, commonly categorized as "cognitive" and "emotional" orientations. There is also a category known as somatic empathy. Among these, "cognition" and "emotion" are the most frequently discussed types ("Empathy," 2024). Cognitive empathy goes beyond simply experiencing the emotions and feelings of others; it involves gaining an accurate understanding and cognition of the situation. Another term for cognitive empathy is "perspective taking." On the other hand,

emotional empathy entails sharing the same feelings as another person and responding with an appropriate emotion (Foster et al., 2016; McCarthy & LaChenaye, 2017).

Empathy can also be classified into two levels. Initially, in counseling, the focus is on “primary levels of empathy.” Primary empathy is beneficial for establishing a therapeutic relationship. It allows the counselor to correctly comprehend the clients’ emotions, reactions, thoughts, and feelings, providing support, encouragement, and a sense of acceptance. It also serves as a foundation for gathering data and clarifying problems (McCarthy & LaChenaye, 2017).

In contrast to primary empathy, the counseling process requires “advanced empathy” with a therapeutic nature. Advanced empathy means that we try to understand the explicit or implicit behaviors, experiences, deep feelings, and inner thoughts of the person concerned, and respond with corresponding emotional words for confirmation and exploration. The true feelings of the parties are brought from behind the stage to the stage (Sung et al., 2004).

With advanced empathy, counselors can grasp the hidden and unexplained deep emotions, experiences, and inner thoughts of clients. It enables the perception of concealed issues, thoughts, hints, or previously undetected aspects, while providing appropriate responses to guide clients in expressing their true thoughts and integrating their experiences, behaviors, and feelings. This deeper understanding assists clients in gaining self-awareness and facilitates the progress of counseling (Clark, 2010; Liao, 2014).

The Relationship between Empathy and Counseling

The counseling relationship encompasses all the interactions between a professionally recognized counselor and the client, involving emotions, attitudes, and behaviors. It encompasses the feelings, attitudes, and expressions exhibited by both parties towards each other (Gelso & Carter, 1985). Extensive research from the past to the present has focused on examining the counselor-client relationship in counseling, with the majority of studies affirming its significance (Bachelor, 1991; Gelso & Carter, 1994; Harcum, 1989). C. Rogers asserts that an effective counseling relationship is built on the counselor’s ability to demonstrate empathy, genuineness, congruence, and unconditional positive regard for the client (Kokotovic & Tracey, 1990).

Empathy plays a crucial role in the effectiveness of counseling. According to Rogers, when the counselor accurately understands and empathizes with the client, the client is more likely to experience positive changes (Rogers, 1975). Similarly, Roffers and Waldo conducted a study on counselors’ empathy and its impact on counseling outcomes, revealing that higher levels of counselor empathy are associated with better counseling results (Roffers & Waldo, 1983). Besides, in today’s curriculum design, empathy is also one of the important considerations in the evaluation criteria of learning performance (Chen et al., 2021). In other words, there is a substantial positive correlation between empathy and the effectiveness of counseling.

Empathy Training Course for Becoming a Counselor

Empathy is not only a fundamental aspect of interpersonal interaction but also a crucial skill and characteristic for counselors and psychologists. While traditional counseling courses have primarily focused on teaching counseling techniques, the complex and diverse nature of interpersonal interactions in today’s society calls for enhanced training in empathy through various learning approaches. Thus, developing empathic communication becomes not only an acquired skill but also an ingrained trait that counselors should be familiar with and adept at utilizing.

Currently, the field of professional psychology has engaged in significant reflections on the education and training of psychologists, promoting the adoption of a “professional competency-based” model for education, training, and assessment. The aim is to bridge the gap between academic education and practical

requirements (Liu & Chen, 2014; Liu et al., 2011; Rubin et al., 2007; Sumerall et al., 2000). Empathy is recognized as a multifaceted phenomenon that enhances counseling and doctor-patient relationships. However, there are limited tools available for students to deliberately practice empathy (Foster et al., 2016). Consequently, this study focuses on training chatbots to serve as virtual clients (VCs), serving as a tool for empathy training.

In the education of counseling professionals, analyzing the core competencies of counseling and incorporating a “situated learning” teaching approach can help address the gaps between instruction, examination, and practical work. This approach supplements the training process, which traditionally emphasizes knowledge acquisition but lacks practical training experiences, thereby improving students’ learning and practical abilities. Counseling is a people-centered profession, involving the complex dynamics of human psychology and behavioral interactions. In addition to acquiring relevant theoretical knowledge in counseling, proficiency in practical counseling skills and professional judgment is essential to fulfill the roles and responsibilities of a counselor. Consequently, the inclusion of training with standardized patients (clients) is increasingly emphasized as an effective practical training method alongside traditional educational approaches to professional development (Chodos et al., 2009; Liu & Chen, 2014; Liu et al., 2011; McLellan, 1996).

Application of Smart Chatbots in the Field of Counseling

As early as 1985, Sharf recognized the potential impact of computer technology on counseling and psychotherapy (Sharf, 1985). Consequently, it is imperative for us to contemplate the possibilities and ramifications of AI replacing human resources in the field of counseling practice and education, particularly as AI becomes increasingly prevalent. This reflection encompasses exploring how AI can be effectively utilized for professional counseling training and supporting counseling practices. We must also consider the impact of AI on individuals’ adaptation in various aspects of life, society, and psychology. Furthermore, ethical considerations in counseling may require adjustments in light of these advancements. Therefore, both current and future counseling practitioners and educators cannot disregard the development trend of AI and must proactively prepare themselves to respond to it.

Currently, numerous countries have developed different types of robots for psychological applications. Examples include: (1) Facebook’s AI-powered robots capable of recognizing signs of suicidal tendencies (Matias computers, Internet) and other AI-related products (Helgadottir, n.d.; Matias et al., 2017); (2) Karim, a chat robot designed by Silicon Valley’s X2AI company primarily to assist Syrian refugees in psychotherapy; (3) Emma, a Dutch-speaking robot employed in the treatment of patients with mild anxiety or phobias (Rucker et al., 2018); and (4) Tess, an English-speaking robot proficient in cognitive behavioral therapy to aid humans.

Shen (2019) mentioned various uses of technology, computers, and AI-related products in the counseling field, including internet-supported intervention services, self-help treatments, computerized therapy, therapeutic robots, virtual reality therapy, and video game therapy. However, based on existing research literature on AI chatbots, it is evident that there is still a long way to go before AI can fully comprehend human minds and potentially replace counselors or therapists. Consequently, it is clear that smart robots still have significant room for improvement in understanding and discerning human emotions and feelings. While this limitation exists, it also implies that there are promising areas for research and development to enhance their capabilities.

Frangoudes et al. (2021) conducted a literature review to explore the primary applications and development of chatbots in medical education. The use of chatbots is a growing trend observed across various fields, including healthcare, service industries, and more recently, education. In healthcare education, specifically, there is an increasing body of research focused on incorporating chatbots into the learning and teaching processes. This trend can be attributed to the portability and affordability of chatbots.

While chatbots offer flexible solutions and have the potential to enhance healthcare education, the literature review indicates that their effectiveness has not been thoroughly evaluated. Thus, further research in this area is warranted.

Given the existing challenges in the use of robots for treatment and care, we propose leveraging the current advantages of robots by employing them as simulated clients to assist counselor trainees in their training process. Currently, training programs heavily rely on role-playing with simulated clients. However, during internships, ethical dilemmas often arise when working with actual clients. To address these dilemmas, our aim is to develop and create robots that can act as clients, offering practical curriculum and exercise opportunities. These robots can be utilized for teaching, training, and skill reinforcement outside of regular class hours. By incorporating such smart robots, we anticipate an enhancement in the effectiveness of training counselor trainees.

Use Situational Teaching and Virtual Consultation (VC) Training to Strengthen Learners' Deliberate Practice of Empathy Skills

Situational teaching, developed from modern cognitive psychology, is designed to enable students to learn in real or simulated situations (Ministry of Education, 2017). This method enhances learning effectiveness and problem-solving abilities by creating realistic scenarios where students can apply and reorganize knowledge. Situational teaching differs from traditional methods by providing experiential understanding, improving skill operations, fostering critical thinking, and enhancing self-confidence and learning satisfaction. Debriefing provides essential feedback (Fanning & Gaba, 2007; Jeffries, 2005; Monti et al., 1998).

Advancements in technology, particularly in virtual reality, have enriched situational teaching. Virtual reality's immersive and interactive features allow learners to experience contexts more vividly, enhancing their preparation and understanding of practical situations (Chen et al., 2017; Lai et al., 2009). Virtual reality differentiates itself from other methods by offering a more realistic experience through its immersive, interactive, and imaginative characteristics (Chang, 2018; Lan et al., 2021). It creates a sense of presence and embodiment, prompting users to adopt others' perspectives (Zheng, 2022). In addition, realistic or online interaction, such as interpersonal interaction via Facebook and so on, has noteworthy positive effects on subjective well-being (Yang & Chien, 2019). Since better mental states naturally enhance the effectiveness the learning process, these studies also inspire us on the designing of approaches to empathy training.

This study aims to develop an empathy training corpus by identifying the characteristic problems faced by future service recipients. By transforming chatbots into virtual standard cases, we intend to use virtual consultations (VC) as a training medium. This will establish a situational teaching framework to train prospective counselors in empathy.

Methods

The objective of this study is to develop intelligent chatbots that can function as virtual clients (VCs) and to explore students' opinions on chatbots as auxiliary tools for empathy training in counseling and guidance personnel.

The research process involves collecting and analyzing relevant literature, selecting suitable chatbots, and then transforming them into smart VCs by integrating practical experiences from counselor training courses into the chatbots' custom algorithms. The smart VCs will undergo calibration and revisions based on evaluation results.

Additionally, this study aims to address two key questions: "Is it feasible to develop and train chatbots to serve as VCs?" and "What are the participants' experiences when engaging in dialogue with VCs?"

Design and Development

The research team comprises counseling educators, information technology experts, and research and development manufacturers involved in the design and development process of transforming chatbots into smart virtual clients. Leveraging the extensive experience of the counseling educator in empathy training, the team has compiled a corpus of emotions and feelings that may arise in middle school students and developed situational scripts for empathy dialogues between prospective counselors and VCs. The chosen corpus is based on the future role of the prospective counselors as middle school teachers specializing in counseling. The script's validity is evaluated by an expert with a Ph.D. in counseling and another expert with a Ph.D. in psychology, and it undergoes several revisions to reach the current version. Close collaboration with the technical team responsible for chatbot development ensures ongoing technical support and resolution of any encountered issues during the training process of the client scenario script.

Students will have real-life interactions with the VC. The VC will share their mood as well as the events and situations that influenced their mood. When the student expresses empathy towards the VC, the VC will respond with "You understand my mood" to conclude the dialogue. Conversely, if the student fails to correctly grasp the VC's emotions or demonstrate proper empathy, the VC will respond with "You do not understand my mood" as the end of the conversation. However, students can continue to practice until they feel confident in responding with empathy or continue practicing to make their empathy skills more natural, fluid, and accurate.

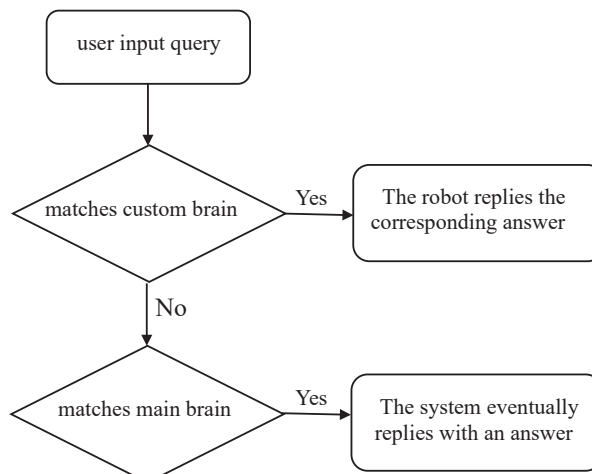
Research Tools

The Chatbot

A chatbot is a computer program that utilizes artificial intelligence and natural language processing (NLP) to comprehend client inquiries and provide automated responses, mimicking human conversation. The chatbot's architecture consists of a "main brain" and a "custom brain." The "main brain" remains unaffected by any training performed on the "custom brain." However, the training of the corpus edited by the user influences the "custom brain." The "main brain" incorporates a sequence-to-sequence generative model as a fundamental component, along with other components responsible for handling user queries in various domains. Given that this paper primarily focuses on the "custom brain" of the system, we will refrain from delving into extensive explanations of the minor details and components of the main brain. Figure 1 provides a general depiction of the relationship between the main brain and the custom brain.

Figure 1

Illustration of the Relationship between the Main Brain and the Custom Brain



Machine Learning Technique for Natural Language Processing Pre-Training

The “main brain” of our system is based on the original Transformer sequence-to-sequence model with an encoder-decoder architecture. The encoder consists of a pre-trained Bidirectional Encoder Representations from Transformers (BERT) model developed by Google, augmented with 2 additional Transformer layers. The decoder, on the other hand, comprises a stack of 3 Transformer layers. The Chinese BERT model we utilize has been pre-trained by the Social Computing and Information Retrieval Center at Harbin Institute of Technology. BERT is essentially a transformer language model with a flexible number of encoder layers and self-attention heads. Google refined this language representation model by leveraging a large amount of unlabeled text through unsupervised learning. In contrast, the “custom brain” in our study is founded on a sentence similarity (SS) model employing a Siamese network architecture, where each sub-network consists of a Convolutional Neural Network (CNN). This SS model is trained using labeled data containing pairs of similar or dissimilar sentences, with the aim of accurately measuring the semantic similarity between the user’s input query and all candidate queries edited by the user. The response from the “most similar” query is selected as the answer to the input query. We chose to use the Transformer architecture with a BERT encoder for the “main brain” and a CNN model for the “custom brain” for the following reasons: The Transformer architecture with a BERT encoder offers advantages such as faster training speed and enhanced performance in human-machine conversations. Recurrent Neural Networks (RNNs) like Long Short-Term Memory (LSTM) models, which have a sequence-dependent structure, are not well-suited for massively parallel computing. The parallel computing capabilities of the transformer model, with the aid of Graphical Processing Units (GPUs), facilitate efficient training. Additionally, when an RNN model processes long sequences, it recursively learns and memorizes previous information, where the features of a word at each position depend on the features of the previous word. Consequently, a long-range dependency issue arises that cannot be fundamentally resolved. In contrast, the BERT self-attention model establishes long-range dependencies by connecting words across layers, addressing this limitation.

Datasets and Data Preparation for Model Training & Quality Control of Dataset

The publicly available sentence similarity datasets, such as the ATEC dataset, CCKS dataset, and so on, account for most of the training data for the CNN semantic similarity model. These corpora are taken from the ATEC (Ant Technology Exploration Conference), CCKS (China Conference on Knowledge Graph and Semantic Computing) and LCQMC (Large-scale Chinese question matching corpus) datasets, including 101,477 entries in the ATEC data set, 100,000 entries in the CCKS data set, 238,766 entries in the LCQMC data set, and some of the data examples are taken twice. These data are widely used academically and are of high quality.

In addition, we also prepared and generated by ourselves a small part of the training data. In terms of methods or strategies for generating synthetic training corpus for these question-question pairs (QQ pairs), the generation methods of positive samples (where Q1 and Q2 are similar) include: completely manually annotated Q1 and Q2, random transformation of the word order of a Q1 to generate Q2, replacement of synonyms of a Q1 to generate Q2, the addition or deletion redundant words of a Q1 to produce Q2, the translation Q1 into another language and then translate it back to the original text (for example, translate it from Chinese to English and then translate it from English to Chinese) as Q2, and so on. These artificially synthesized corpus we prepared by ourselves is mainly related to the dialogue materials of our virtual counseling in order to fine-tune the trained CNN semantic similarity model to perform better for the contents of VC projects.

The (Chinese language) training corpus for the convolutional neural network CNN has 655,393 entries or question-question pairs. The total number of characters with in the pairs of questions in each of these 655,393 entries amounts to 13,623,213. This model uses individual characters as tokens, that is, 13,623,213

tokens. The effect of model training is good and in line with expectations. For example, the accuracy of this model on the test set of the ATEC dataset is 86.4%.

Siamese Neural Network

Advantages of using a Siamese network with two sub-networks of CNN for the “custom brain” include excellent performance and speed. CNN has the capability to automatically detect features without the need for manually crafted rules. In our client, we achieve this by incorporating a convolution layer with filters of lengths 2, 3, and 4 words, which correspond to features representing 2, 3, and 4 consecutive Chinese characters. Moreover, since most user queries in conversations are typically not lengthy, and the word order often does not significantly affect semantic similarity measurement, the property of “translational invariance” in CNN makes it suitable for this scenario. This property also helps avoid the issue of overfitting data to word order that can arise with RNNs. Additionally, CNN models generally offer faster processing compared to RNN and BERT, enabling quicker inference.

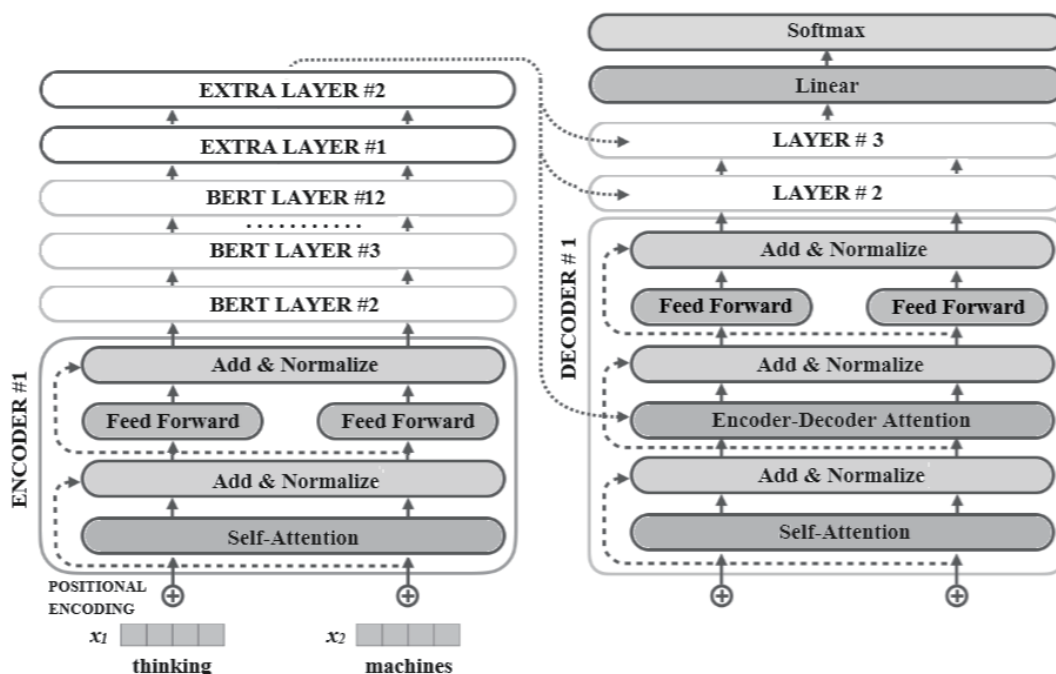
Architecture, Model Training, and Data Preparation for the “Main Brain” and the “Custom Brain”

In the upcoming sections, we provide detailed information on the architecture, model training, and data preparation for both the “main brain” and the “custom brain” of our system.

Architecture of the “Main Brain” Model

The model adheres to the encoder-decoder architecture of the original Transformer model (Vaswani et al., 2017). The encoder component comprises a stack of 14 Transformer layers, consisting of the 12-layer pre-trained Chinese BERT (Cui et al., 2021) along with two additional transformer layers. On the other hand, the decoder component consists of 3 Transformer layers. Figure 2, adapted from (Alammar, 2018), depicts the architecture in detail.

Figure 2
Illustration of the Architecture of the “Main Brain Model”



Note. This figure shows the encoder-decoder architecture of the Transformer model. Adapted from *The illustrated transformer* [Blog post], by J. Alammari, 2018 (<https://jalammar.github.io/illustratedtransformer/>).

The model receives a tokenized sequence of tokens representing a Chinese sentence as input. These tokens are processed using the WordPiece tokenizer of BERT and converted into a continuous representation, which includes word embedding and positional embedding. The encoder then maps the input sequence to produce encoder outputs denoted as “z.” Using these encoder outputs, the decoder generates output tokens sequentially, which are subsequently converted into the final output sentence.

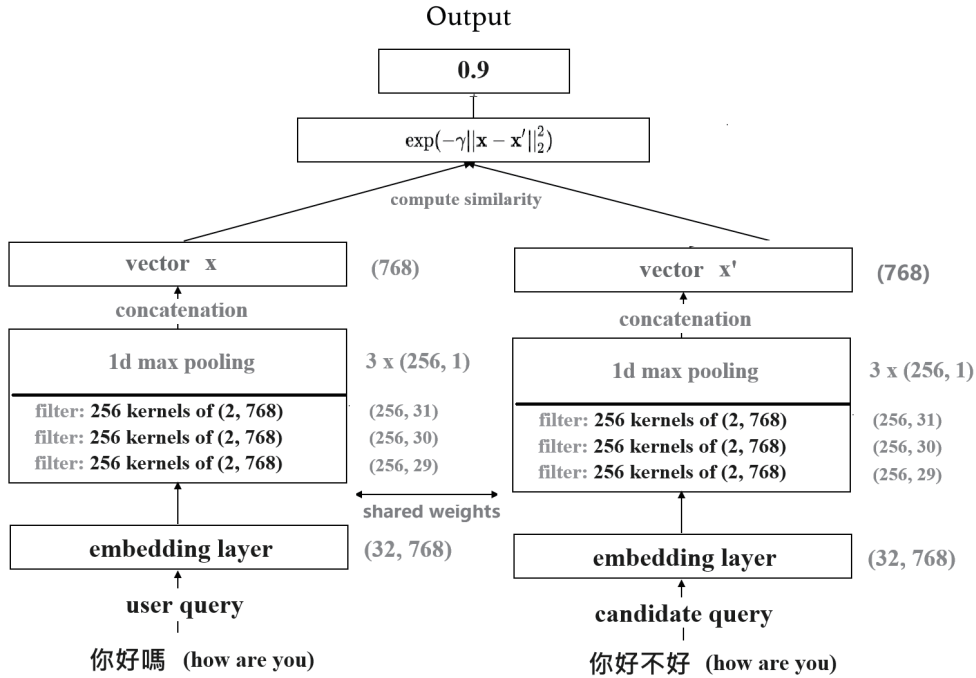
Architecture of the “Custom Brain” Model

We developed a sentence similarity (SS) model for the functionality of the chatbot that can be trained with detailed dialogues between the interlocutor and VC. This model is applied to assess semantic similarity between a pair of two sentences. Given a trained SS model, we use it to create a project with query and answer pairs (Q_1, A_1) , (Q_2, A_2) , \dots (Q_n, A_n) first. Subsequently, for a user query Q , the model can be used to measure the similarities between Q and the candidate queries Q_1, Q_2, \dots, Q_n ; suppose the most similar candidate query is Q_m , and Q_m is sufficiently similar to Q , then the corresponding answer A_m is taken as the answer to Q .

Figure 3 illustrates the architecture of the SS model. It is based on a Siamese neural network that contains two identical sub-networks of convolution neural network (CNN), and each of the sub-network CNN takes one variable-length Chinese sentence as input and outputs a fixed 768-dimension vector. Each input sentence of length L is first preprocessed to a sequence of 32 token ids: If $L < 32$, extra $32-L$ token ids of the ‘[PAD]’ token are appended to the sequence; if $L > 32$, the sequence is truncated to 32 tokens. The embedding layer looks up the word embedding of the token ids and converts to a $(32, 768)$ tensor, and

then a convolution layer with three filters of 256 kernels of sizes (2, 768), (3, 768) and (4, 768) converts the word embedding to 3 tensors with sizes (256, 31), (256, 30), and (256, 29) respectively. An ReLU function is applied to these three tensors, and then a one-dimension max-pooling layer converts each of them to size (256, 1). These three tensors are concatenated to a vector dimension 768 as the output of the CNN sub-network.

Figure 3
Illustration of the Architecture of the SS Model



After computing the vectors of the input query and a candidate query, we compute similarity between the two vectors with the radial basis function using the following expression.

$$K(X, X') = EXP(-\gamma \|X - X'\|_2^2) \tag{1}$$

If the similarity value surpasses an empirically turned threshold like 0.85 and is sufficiently close to 1.0, the candidate QA pair is considered a suitable candidate. The most similar suitable candidate is taken as the matched candidate QA. The threshold like 0.85 here is a hyperparameter that can be set according to the needs of the projects. Based on the similarity between the sentence of the query and the sentences in the project corpus, within the system program, the threshold value is used to decide how easy the queries should hit the most similar Q&A in project corpus. If it needs to be more vague, that is, it is easier to hit, that is, if the query is not particularly similar to the project corpus and you want to hit it, then the threshold can be set lower, like 0.8; conversely, if it needs to be more precise, you can lower the threshold. The threshold is raised, for example to 0.9.

Model Training of the Sentence Similarity (SS) Model

Regarding model training, the model is trained with the Adam’s optimizer, and we experimented with

some small learning rates and dropout regularization between 0 and 0.3. We use the mean squared error (MSE) as the loss function given by

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \frac{1}{n} \sum_{i=1}^n (Y_i - K(X_i, X'_i))^2 \quad (2)$$

In Eq. (2), for a training example, 你最近如何 (How have you been lately), 你最近怎樣 (How's everything lately), 1), the X and X' are the output vectors after passing “How have you been lately” and “How's everything lately” to the CNN network, the labeled Y value is 1, and the optimizer minimizes the difference between the labeled similarities and the output similarities for the training examples.

Model Training Approaches & Criteria for Fine-Tuning Parameters

The training of our model is the same as the general machine learning model training strategy and method. During training, the data set will be divided into training set, verification set and test set, and during training, the training examples of the training set will be used. Training, and when training to each epoch, judge the current performance of the model training based on the performance on the validation set; stop training the model when the validation loss starts to maintain without further increase, and use the model of the epoch with the smallest validation loss as the trained model.

Software Engineering of the Chatbot System

This section explains the software engineering perspective of the components of the chatbot. In particular, we focus on the custom brain of the chatbot, since the empathy dialogues between interlocutors and VCs are based on the custom brain.

Figure 4 illustrates the relationship between our chatbots and the server. The server side of the chatbot system can allow many chatbots to access, and the device side of the chatbot contains Android applications for its functionalities and accesses the server via a RESTful API (“REST,” 2024). REST stands for representational state transfer, and API stands for Application Program Interface; a RESTful API is an API with an architectural style that accesses data with HTTP requests. The server takes much computational resources including space, memory, and computing power, it is therefore not feasible to set up the entire chatbot system on each chatbot. In addition, the modules and data in the server need to be updated constantly, and it would reduce user-friendliness to require the chatbot users to download and update the device applications frequently. Thus, in order to achieve greater scalability and maintainability and better performance, we set up the computational resources on the server and use RESTful API for data transfer.

In the chatbot application, the audio speech input is first converted to text with a speech recognition API like Google speech recognition. For example, suppose the user speaks “你看起來心情不好” (“You seem to be in a bad mood.”), the voice is converted to the text, and all pieces of information about this query, including the recognized query text, user id, project id for specific scripts of empathy dialogues, etc., are incorporated into a JSON object such as {‘query’: [‘你看起來心情不好’], ‘langCode’: ‘zh-tw’, ‘project_key’: ‘some-project-key’, ‘id’: ‘some-id’, ‘platform’: ‘dev_internal’, ‘debug’: 1} and sent to the server. The server then performs calculations and determines a suitable response, and the response to the user, e.g., “怎麼了，發生什麼事了呢” (“What’s wrong? What happened?”) is sent back to the chatbot and then converted to speech with a text-to-speech API.

Regarding the usability of Google’s Text-to-Speech and Speech-to-Text services, for speeches without in-depth technical or academic terminologies, the Google’s automatic speech recognition service works pretty well, and the error rate is low in most cases. As long as the pronunciation is standard, the chances of an error occurring are low. Thus, the Google’s services suffice our needs. In addition, low error rate in

speech recognition leads to few incorrectly recognized words in the spoken sentences, so our sentence similarity model is still able to retrieve the corresponding similar QA pairs in the trained projects.

The Suitability and Effect of the Chatbot for Empathy Training

Since the dialogues for empathy training are more closed domain instead of open domain dialogues involving wide range of topics, the chatbot can meet the requirements for empathy training. The sentence similarity model can cover similar sentences within our dialogues scenarios, and the sequence to sequence model for the main brain also suits our needs since it is trained with our prepared data. This is different from using existing pre-trained models, which may be more general in nature. However, the answers generated by these models may not be suitable or predictable for our specific scenarios, and we do not have control over them.

In the project training stage, the question and answer corpus is designed based on the dialogue situation; different project situations can be considered closed domain, so they can be covered.

The Use of LLM for Performance Enhancement

As far as the dialogue situation of virtual counseling is concerned, our system and model have met the requirements, and currently does not directly use large language models (LLMs) such as ChatGPT for conversation. First, a self-trained CNN model, compared to a larger multi-layer Transformer model for sentence vector, the amount of calculation is much less and the speed is faster, which affects the user experience and the cost of deployment. Second, the accuracy of the model in judging semantic similarity is enough to meet the needs; in the dialogue situation of the project, there are not many parts involving complex semantics, and for those complex questions where the semantic similarity judgment is not accurate enough, it is more efficient to supplement corresponding project corpus to cover them. Above all, the dialogue situations of virtual counseling need to be specially designed, which are more closed domain instead of entirely open domain. For closed domain scenarios, the answers generated directly by LLM are uncontrollable, which means that a certain proportion of answers from an LLM will not meet expectations. However, if a semantic similarity model is used, the question and answers covered by corpus of the trained projects are guaranteed to be correct. The answers are all predetermined and meet expectations.

However, we can still use LLM to assist ourselves in generating drafts of our virtual counseling-related project corpus, which we can then modify, optimize, or polish. We can also use LLM to generate sentence paraphrases for questions and answers involving complex semantic parts to augment the project corpus. These are appropriate use of LLM in our system for enhancing the overall efficiency for developing our chatbot for VC.

An Example of Dialogues Between the Chatbot and the Tester

When the tester converses with the chatbot, some of the tester's queries hit the QA pairs of the project corpus, and the chatbot replies with corresponding answers from the custom brain which are guaranteed to be as expected. In addition, some of queries does not hit the project corpus, and the chatbot replies with the sequence to sequence model from main brain. Though the answers not from the custom brain are not fully predefined and may not all be as expected, the rounds of dialogue are generally consistent with the main line of the project corpus. This is because the training corpora of the seq2seq model for the main brain are prepared on our own, which can generally fit the dialogue scenarios of our VC projects. Furthermore, the seq2seq model is a small model, so, from an engineering perspective, we can continuously replenish and refine its training corpora and periodically retrain the model to fit our needs.

The table below shows examples of 7 QA pairs of a trained project and 6 example user queries and

corresponding answers from the chatbot. Among the 6 user queries, 3 of them hit the project corpus and obtain exactly the same as the predefined answer of the corresponding matched Q, including “聽起來你很生氣，是怎麼回事” (It sounds like you're angry, what's going on?), which hits 聽起來你很憤怒 怎麼了 (You sound angry. What's wrong?), 你生氣的原因是因為同學不理你嗎 (Are you angry because your classmates ignore you?), which hits 同學不理你讓你很氣憤 (You are very angry when your classmates ignore you), and 那你會很生氣嗎 (Will you be angry then?), which hits 那你一定很在意 也很生氣 (Then you must be very concerned and angry). Though the other 3 queries including 你的感覺不好嗎 (Are you feeling bad?), 我和大多數人一樣很討厭被誤解 (Like most people, I hate being misunderstood.), 這讓人感到不公平 (This feels unfair) do not hit any of the QA pairs of the project corpora, the corresponding answers for the queries flowed to the seq2seq model are generally fluent and do not divert the conversation from the context covered by the project corpus.

Table 1
Examples of Pre-defined QA Pairs in the Project

Q (in the training project)	corresponding A
你還好嗎 (How are you)	我覺得很生氣 (I feel very angry)
你感到不高興 (You feel unhappy)	對阿 同學不理我了 (Yes, my classmates ignored me.)
同學不理你讓你很氣憤 (You are very angry when your classmates ignore you)	沒錯 (That's right)
聽起來你很憤怒 怎麼了 (You sound angry. What's wrong?)	同學誤會我沒告訴他就拿了她的東西 (My classmate misunderstood me that I took his things without telling him.)
被誤會的感覺一定不好受 (It must be uncomfortable to feel misunderstood)	對阿 事實上我沒拿 他也不聽我解釋 (Yes. In fact, I didn't take it and he didn't listen to my explanation.)
讓你沒有機會解釋 百口莫辯 (Leave you no chance to explain, unable to argue)	是阿 很不公平 謝謝你懂我的心情 (Yes, it's very unfair. Thank you for understanding my feelings.)

Table 2
Examples of User Queries and the Chatbot's Corresponding Answers

user queries	the chatbot's answer	source
你的感覺不好嗎 (Are you feeling bad?)	不錯啊，我覺得這種事情本來就不應該發生在自己身上 (That's right. I think this kind of thing should never happen to me.)	main brain
聽起來你很生氣，是怎麼回事 (It sounds like you're angry, what's going on?)	同學誤會我沒告訴他就拿了她的東西 (My classmate misunderstood me that I took his things without telling him.)	custom brain
你生氣的原因是因為同學不理你嗎 (Are you angry because your classmates ignore you?)	沒錯 (That's right)	custom brain
我和大多數人一樣很討厭被誤解 (Like most people, I hate being misunderstood.)	我也討厭被誤解 (I also hate being misunderstood)	main brain
那你會很生氣嗎 (Will you be angry then?)	是阿 我覺得很誇張喔 最後居然就不跟我講話了 (Yes, I thought it was exaggerating. In the end, he actually stopped talking to me.)	custom brain

(Continued)

Table 2
Examples of User Queries and the Chatbot's Corresponding Answers (Continued)

user queries	the chatbot's answer	source
這讓人感到不公平 (This feels unfair)	這種人真的很討厭，不是很明白耶 (This kind of person is really annoying. I don't understand it very well.)	main brain

Note. request url = http://127.0.0.1:5002/api/chatbot_coordinator {'boc_score': 0.00010778333671623841, 'boc_type': -1, 'cache': 0, 'chat': ['不錯啊，我覺得這種事情本來就不應該發生在自己身上'], 'chatbot-coordinator_timeTaken': '0.956 sec', 'dummy': 0, 'module': 'nuwa_chat', 'query': ['你的感覺不好嗎'], 'response': ['不錯啊，我覺得這種事情本來就不應該發生在自己身上'], 'source': 's2s', 'x_trace_id': ''}; request url = http://127.0.0.1:5002/api/chatbot_coordinator {'boc_score': 0.011394217610359192, 'boc_type': -1, 'cache': 0, 'chat': ['我也討厭被誤解'], 'chatbot-coordinator_timeTaken': '1.257 sec', 'dummy': 0, 'module': 'nuwa_chat', 'query': ['我和大多數人一樣很討厭被誤解'], 'response': ['我也討厭被誤解'], 'source': 's2s', 'x_trace_id': ''}; request url = http://127.0.0.10:5002/api/chatbot_coordinator {'boc_score': 0.0, 'boc_type': -1, 'cache': 0, 'chat': ['這種人真的很討厭，不是很明白耶'], 'chatbot coordinator_timeTaken': '1.356 sec', 'dummy': 0, 'module': 'nuwa_chat', 'query': ['這讓人感到不公平'], 'response': ['這種人真的很討厭，不是很明白耶'], 'source': 's2s'}.

Figure 4
The Relationship between the Chatbot and the Server in the Chatbot System

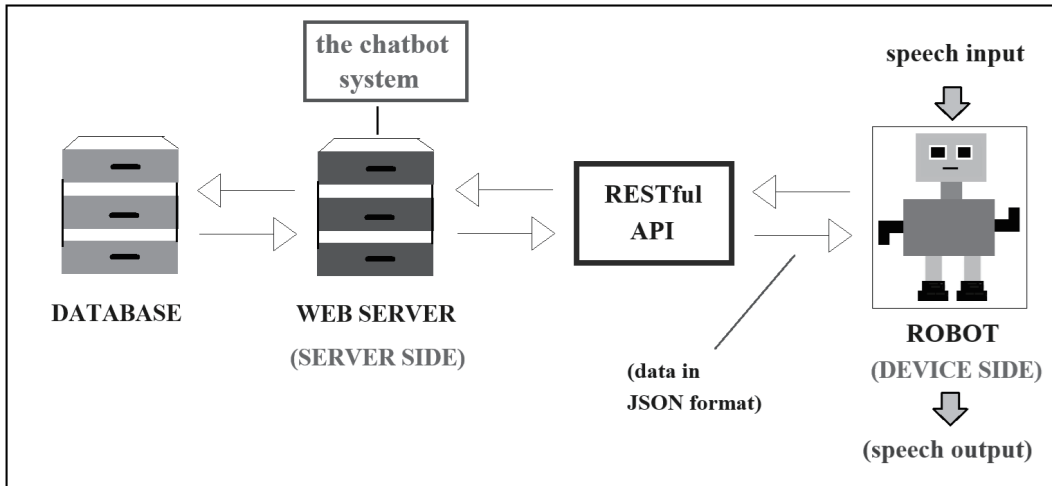


Figure 5 illustrates the operation flow and computation process of the custom brain of the chatbot. To avoid confusion of the meaning of “training,” we define “model training,” “project training,” and “testing” in Table 3.

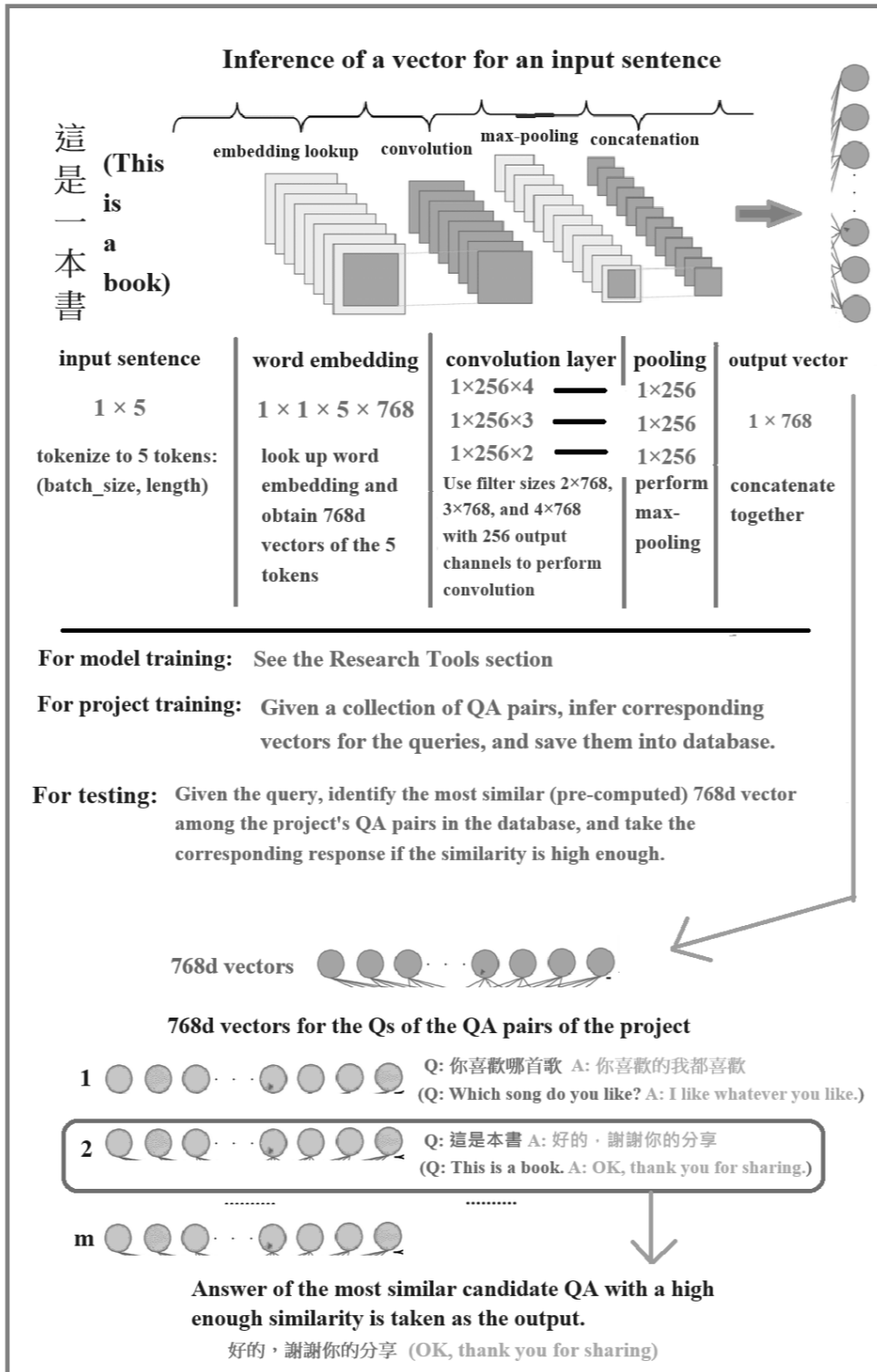
Table 3
Comparison of Model Training, Project Training, and Testing

model training	The training of the sentence similarity (SS) model for the custom brain.
project training	The inference of vectors for the questions of a set of question answer pairs (training scripts) with the trained SS model, which are all saved to database. During project training, the parameters of the SS model are kept fixed and not fine-tuned.
testing	The inference of vector for the input query and the identification of the most similar QA pair from the trained scripts if exists.

Our sentence similarity model is based on a common Siamese convolutional neural network architecture (CNN) similar to (Shi et al., 2020). In addition, for details of the architecture and techniques of CNN, see paper (O’Shea & Nash, 2015) for reference. When the input JSON object `{‘query’: [‘這是一本書’], ‘langCode’: ‘zh-tw’, ‘project_key’: ‘some-project-key’, ‘id’: ‘some-id’, ‘platform’: ‘dev_internal’, ‘debug’: 1}` (“This is a book.”) is sent to the custom brain system on the server, it goes through the calculation process as illustrated in Figure 5 to infer a 768d vector for the input query “這是一本書” (“This is a book.”) with the trained SS model and check the most similar QA pair among all training scripts of the trained project with project id “some-project-key.” Suppose that among all Qs of the candidate QA pairs in the “some-project-key” project, the most similar QA pair is (“這是本書” (“This is book”), “好的，謝謝你的分享” (“OK, thank you for sharing.”)). If the similarity between its inferred vector and that of input query is high enough based on the radial basis plus some basic rule-based string similarity methods, the custom brain returns its answer “好的，謝謝你的分享” (“OK, thank you for sharing.”) to the chatbot; if not, the custom brain returns null, and the input JSON object is subsequently sent to the chatbot’s main brain, which will generate an answer with its sequence to sequence model or rule-based modules and return it to back to the chatbot. Besides, it should be noted that one performance bottleneck regarding the identification of the most similar candidate query answer pairs from the trained project: Suppose the number of QA pairs for the training scripts of one project is N , it is very time-consuming to take $O(N)$ linear time complexity to compute similarity between the vector of input query and those of the N candidate QA pairs when N is large.

To overcome the problem, we use an approximate nearest neighbor (ANN) search algorithm based on pre-built search trees and projections with the ANNOY package. It was developed in 2015 by Erik Bernhardsson at Spotify, and the implementation and algorithmic details can be referred in (Bernhardsson, 2015). With the pre-built search trees, the time complexity for looking up a small number of k nearest points for one point can be reduced to $O(\log N)$. The use of ANNOY reduces calculation time and makes the system implementation feasible. For a trained project with about 1,000 QA pairs, the average calculation time for obtaining a response for a query is below 0.1 second.

Figure 5
The Operation Flow and Computation Process of the Custom Brain of the Chatbot



Transforming Chatbots into Smart Virtual Clients (VCs)

Based on the rich experiences of empathy training in the past years, we have compiled the emotions and feelings, which the middle school students might involve, into situational scripts of empathy dialogue between prospective counselors and VCs.

We build the corpus of the chatbot based on the situational scripts of empathy dialogue. The script is evaluated for validity by one expert with Ph.D. in counseling and another expert with Ph.D. in psychology, and is revised several times to get the current version. These scripts are then entered in a training console through internet for training “custom brain” of the chatbot, followed by synchronization to chatbots. Chatbots are then ready to be used as VCs to carry out empathy training testing for prospective counselors. After testing, we add more dialogues to enrich the corpus for correctness of chatbot responses. We further refine the corpus according to the testers’ experience.

Testing of the Chatbot Acting as Virtual Client (VC)

Before students begin training in empathy through seven rounds of client scenario exercises, they must follow guidelines for practicing empathy. These students are being trained to become middle school counselors in the future. The guidelines will assist prospect counselors in understanding the context of the conversation and prevent miscommunication with the VC.

In the following scenarios, students are asked to strictly adhere to the principles of dialogue. If the VC does not understand the student’s words or the student’s empathy response is incorrect, the student is asked to try the empathy response again. For example, if the VC says, “I feel angry.” and the student responds, “You are not feeling well.” and the VC says, “Nope.” This would indicate that the student does not provide an appropriate empathic response and should try again by saying something like “You feel angry.” In another scenario, students are required to focus on the VC’s emotions and practice empathic responses. For example, VC says: I feel angry. Student responds with something like: Can you tell me why? or student responds: What happened? or student responds: How do you want to solve it? In this scenario, students should try to say something in an empathic way such as “you feel angry” instead of ask questions, gathering data, or trying to help the VC and entering the mode of problem solving. These two scenarios are often seen in the dialogue between the participants in this study and VCs. When the tester fails to properly empathize with the VC, the VC will make unexpected dialogue responses, similar to how a real client would respond when a counselor fails to properly empathize with them. The client might digress from the subject of discussion, be reluctant to reveal more relevant information, or exhibit resistance. This situation can only be improved when the counselor can respond correctly to the client. Therefore, the chatbot’s response in this study is designed to imitate the response of a real client, so that the chatbot can function as a VC, playing the role of a client.

Information about the Testers

The testers were recruited openly. All the testers voluntarily agreed to participate in the test of empathy with robots. In the first test, a total of 22 students participated in the test, all of them are students, including 4 males and 16 females. This group included 6 junior students, 15 senior students and 1 graduate student for a master’s degree. In the second test, a total of 38 students participated in the test, all of them are the first-year undergraduates, including 11 males and 27 females. All participants in this research signed the “study participant consent form” and this research was performed in accordance with relevant guidelines/regulations and approved by National Changhua University of Education Research Ethics Committee.

The following two questionnaires are used to understand the tester’s experience.

User Experience Questionnaire (UEQ)

This research uses UEQ questionnaire to understand the user's experience (Laugwitz et al., 2008). The items will include: Willingness to use, difficulty, convenience, communication, empathy, satisfaction with companionship, etc. There are 12 questions in total, and each question is measured on a Likert scale seven-point response level: "1. Very dissatisfied," "2. Dissatisfied," "3. Relatively dissatisfied," "4. Average," "5. Relatively satisfied," "6. Satisfied," "7. Very satisfied." For each question, the tester is asked to give any score from 1 to 7 corresponding to the seven levels of the Likert scale. The higher the score, the more satisfied with the empathy interaction of the chatbot.

System Usability Scale (SUS)

This study uses SUS questionnaire to understand the user experience (Brooke, 1996). The items will include: Willingness to use, degree of difficulty, convenience, degree of communication, empathy, satisfaction with companionship, etc. There are 11 questions in total, and each question is measured on a Likert scale five-point response level: "1. Strongly disagree," "2. Disagree," "3. Neither agree nor disagree," "4. Agree," "5. Strongly agree." For each question, the tester is asked to give any score from 1 to 5 points corresponding to the five levels of the Likert scale. The higher the score, the more agree with the ease of use of the chatbot and the agreement that the chatbot can be used to train empathy.

Results

The client robot simulated in this research, following numerous tests, problem identification, communication, discussion, and refinement of training methods, currently functions as a simulated client. It engages in conversations with students regarding their mood, conducts multiple rounds of dialogue, and incorporates corresponding dynamic body movements to enhance participants' motivation for interaction. Virtual Clients (VC) encompass case scenarios aligned with the empathy training objectives for undergraduate students in counseling or psychotherapy. The study's content spans various emotions, such as anger, anxiety, fear, happiness, joy, jealousy, and other situations. This facilitates students' practice of empathy through conversations with Virtual Clients (VCs).

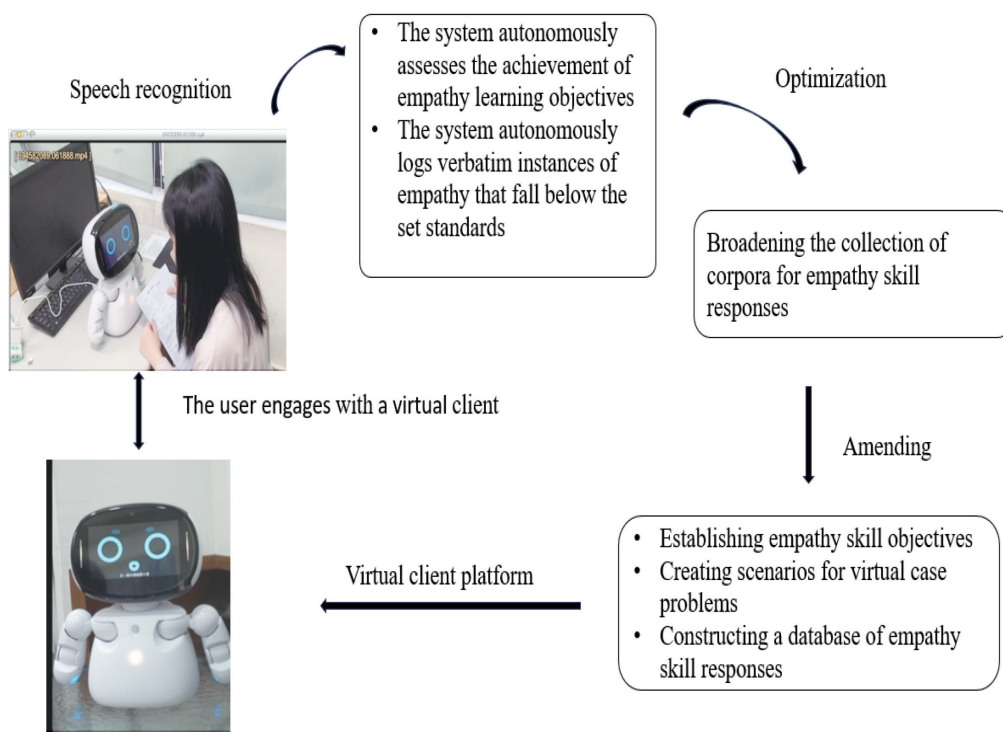
The simulated client robot developed in this study, after many tests, problem discovery, communication and discussion, adjustment, and modification of training methods, can currently serve as a simulated client, talk to students about their mood, and conduct multiple rounds of dialogue, accompanied by corresponding Rich body movements can increase interactants' motivation to interact with them.

Virtual Clients (VC) include case scenarios that are consistent with the empathy training goals for undergraduate students in counseling and counseling or psychotherapy. The content developed in this study covers: (1) anger; (2) anxiety; (3) fear. Emotions; (4) happy emotions; (5) joyful emotions; (6) jealous emotions; and (7) fearful emotions and other situations. Allows students to have empathy practice conversations with VCs.

After the chatbot was successfully converted into a VC in this study, a total of 22 students from the Department of Guidance and Counseling were recruited to participate in the test with informed consent to conduct the empathy conversation with the chatbot. The purpose of the test is to understand the experience of the students who participated in the test and to evaluate the ease of use of the chatbot as VC and users' experience. The experiences of participated students are understood by using the User Experience Questionnaire (UEQ), the System Usability Scale (SUS), and the feedback on open-end questions, which were filled out by participated students immediately after the test (Brooke, 1996; Laugwitz et al., 2008).

Figure 6

Flowchart Illustrating the Utilization of Virtual Clients for Coaching Consultation Abilities



Flowchart illustrating the utilization of virtual clients for coaching consultation abilities

As shown in Table 4, the UEQ has a total of 12 questions, each question has seven response levels on the Likert scale, and 22 testers are asked to give any score from 1 to 7 for each question corresponding to the seven levels of the Likert scale. The higher it is, the more satisfied the tester feels. The average score for each question is shown in Table 3. All testers score between 4-7 points, that is, their experience with the chatbot falls at the scale position from average to very satisfied, with an average score of more than 5. The testers think that the chatbot is cute in appearance, suitable in body size, vivid in expression, suitable in voice and intonation and that the chatbot has appropriate body language. Besides, the testers like the chatbot and feel comfortable with him and want to chat with him. Overall, the testers consider the chatbot an ideal robot, and could practice empathy with him.

For the System Usability Scale (SUS) shown in Table 5, there are 11 items in total, each of which has five response levels on the Likert scale. For each item, the testers are asked to give any score from 1 to 5 for each question corresponding to the five levels of the Likert scale. Note that the fourth item is a reverse scoring question. However, the fourth item would not be calculated in averaging the score. The higher the score, the more satisfied with the ease of use of the chatbot. The average score for each question is shown in Table 4. All testers scored between 3 and 5 points, that is, their experience with the chatbot fell at the scale position from neutral to strongly agree, with an average score of over 4. The results show that the testers agree with the ease of use of the chatbot system, and think that most people can quickly learn to use the system, interact with VCs, and have no pressure to talk to. Overall, the testers think that VCs are suitable for practicing empathy, and feel that empathic response can be improved because of such practices. In the User Experience Questionnaire results, the subjects liked the appearance of the chatbot, felt that he was cute and had a suitable body shape; felt that the robot's voice tone was suitable, the body language and expressions

were vivid; felt comfortable interacting with the robot and communication was smooth; believed that he was the ideal robot and could practice empathy with him.

In addition to quantitative questionnaires, testers were also asked to give feedback on open-ended questions in this study. The testers said that both the chatbot's appearance and body movements were cute and interesting. Talking with the chatbot for practicing empathy would not feel stressful, wishing to continue the conversation with the chatbot. However, the testers feel that the expressions of the chatbot are currently unable to match the emotional presentation, and are limited by the presentation of the voice to some extent. During practicing empathy with the chatbot as VC, the testers felt that it was helpful and could improve showing empathy. Furthermore, the practicing process of the conversation between the tester and the VC is recorded and can be turned into a verbatim transcript, which can be provided as feedback to testers on empathy performance.

This study expands the number of emotional scenarios in the corpus from 7 to 14, following an initial test with 22 participants. Afterwards, 38 first-year students from the Department of guidance and counseling were recruited for a second test. The average score for the User Experience Questionnaire (UEQ) is 5.20, and the average score for the System Usability Scale (SUS) is 4. The qualitative feedback data from the second test are consistent with the results from the first test.

Table 4
User Experience Questionnaire (UEQ)

Item Number	Questionnaire Item	Average Score of Test Results (1-7)	
		1st test	2nd test
1	Overall, I love it	5/5	5.1
2	I think it looks cute	6/6	6.1
3	I think it's fit	6.3/6	1.1
4	I think it's expressive	5.5/5	5.3
5	The tone of voice is suitable	5/4	5
6	It is able to communicate with people	5/5	
7	It has appropriate body language	5/5	2.6
8	I am comfortable with the chatbot	5/4	8.2
9	I would like to chat with him	5.1/5	
10	I'd like to introduce it to others	6/5	6.8
11	I would like to introduce it to others	5.5/5	
12	I think it's an ideal chatbot	5/5	2
Average		5.4/5	2.9

Table 5
System Usability Scale (SUS)

Item Number	questionnaire item	Average Score of Test Results (1-5)	
		1st test	2nd test
1	I think I would like to use this product regularly	3.4/3	5
2	I find it interesting having conversation with this product	4.1/4	2.6
3	I think this product easy to use	3.5/3	6.8
4	I think I need help with this product (reverse scoring question, not calculated in the average score)	3.2/3	0.5

(Continued)

Table 5
System Usability Scale (SUS)(Continued)

Item Number	questionnaire item	Average Score of Test Results (1-5)	
		1st test/2nd test	
5	I think the functionality of this product is well integrated	3.4/3.68	
6	I think this product allows me to practice empathy	4.6/4.5	
7	I think most people can learn to use this product very quickly	4.3/4.26	
8	I feel no pressure in having conversation and interacting with this product	4.2/4.1	
9	I think it helps to practice empathy with this product	4.3/4	
10	I think this product is great for practicing empathy	4.3/4.1	
11	I feel that my empathy response improves as a result of this practice	4.1/4.32	
Average		4.1/4.04	

Discussion

In the process of developing the chatbot serving as VC, the research team encountered various technical limitations and challenges just like the problems encountered reported in the reference (Shorey et al., 2019). During this study, there are two challenges in enabling the chatbot to effectively and correctly assume the role of the virtual client for providing students with smooth empathy dialogue experience. First challenge is how to design the client scenarios in such a way that the VC chatbot can understand what the interlocutor is talking about. The second challenge is how to make the VC chatbot can conduct multiple dialogues.

The brain of the chatbot in this study can be divided into “main brain” and “custom brain.” The main brain will not be affected by any custom brain training and is the result of the original pre-training. The custom brain will be affected by the custom sentences entered by the users in the training room console (Figure 4). We can design the content of the training dialogues to make the custom brain show the logic of induction and deduction with enriched corpus by using different but synonymous wordings or similar expressions from the response of the interlocutor to the same dialogue sentence of the VC. This is in line with the synonym substitution of emotional adjectives between interlocutors in the process of empathy dialogue, which means that it conforms to the principle of correct empathy expression and the requirement to achieve correct empathy. For example, when the VC says “I feel very angry,” in different entries of user’s query of the training room console, we edit and add the following sentences, “You feel upset,” “you feel very angry,” “You are a little in a bad mood,” “You are a little sulking,” “You are a little upset,” and “You are a little angry,” etc. for interlocutors to respond.

After many tests, problem discovery, communication and discussion with chatbot’s engineer of the vendor, adjustments, and modification of training methods, the chatbots serving as virtual clients developed by this research can make correct responses based on the content of our training. Even if sometimes the dialogue is beyond our training scripts, the chatbot can also talk naturally, and there will be no dilemma of talking without communicating, because the chatbot has a main brain to assist with a custom brain. The VCs can be as natural, expressive, and actionable as a real person is speaking to you.

During the testing process, we find that the testers show curiosity and interest when they first see the VC, and happily practice empathy with the VC. They show willingness of multiple rounds of dialogue during the conversation. In addition, the average score of the tester’s User Experience Questionnaire

exceeds level 5 on 7 point Likert scale. According to the System Usability Scale, the average score is above 4.1 on 5 point Likert scale. The testers think the chatbot is cute and fit having lively expression with suitable voice and tone, and appropriate body language. The testers like the chatbot and feel comfortable with him, wanting to chat with him. Overall, the testers think that the chatbot is an ideal robot, and they can practice empathy with him. Besides, the testers agree with the ease of use of the chatbot system. They believe that most people can quickly learn to use the system, interact with VCs, and have no pressure to talk to. The testers believe that VCs are suitable for practicing empathy and they will improve their empathy skills with such practices. In addition, based on the feedback of open-ended questions, the testers think that the appearance and body movements of the chatbot are very cute and interesting; talking to chatbots for the empathy practice will not feel stressful; the testers feel quite relaxed and want to continue the conversation; and the testers affirm that VC is suitable for practicing empathy and that they can improve empathy from it. Therefore, it can be readily seen that this study can successfully convert chatbots into VCs as an auxiliary practice tool for empathy training. These results are consistent with other reports (Foster et al., 2016; Frangouides et al., 2021; Shorey et al., 2019).

At present, chatbots can effectively and correctly assume the role of VCs, talking about mood with students under training by providing students with multiple rounds of smooth empathy dialogue experience. As such, students will be motivated and show more willingness to interact with VCs. Besides, students can practice empathy with VCs for each category of emotion as if passing a level in a game. At the end of the dialogue, VCs will respond to students depending on whether students could read the emotions of VCs expressing empathy. With these practices, students can get immediate feedback and thus help students in the repeating exercises so as to increase their sense of self-efficacy. Our research finds that developing chatbots as virtual clients in empathy training is feasible, similar to existing literature on using empathy or communication skill training for healthcare professionals (Frangouides et al., 2021).

Emotion is a state of physical and mental excitement produced by an individual after internal and external stimuli. Although this state can be experienced but not easily controlled by the individual. Emotions can interfere with and promote the individual's physiology and behavior. How many different emotions do humans have? People might give different answers. In the 1970s, psychologist Paul Eckman identified six basic emotions, that is, happiness, sadness, disgust, fear, surprise and anger. He believed that these emotions are common in all human cultures. Nevertheless, human emotions are quite complex. In addition to these six prototype human emotions. After maturation, learning, and complexity, people could have more than 135 kinds of positive to negative emotions, ranging from strong to weak emotions (Ekman, 1987). The situational scripts of empathy dialogue are first designed in this study according to seven types of emotions, we will develop dialogue content for more emotions in the future to enrich the corpus.

The problem found at present is that the corpus still needs additions. The existing corpus will limit to some extent the dialogue content with VC. In addition, VC is a cute robot with a relatively young voice, which is just in line with the future service objects of the trained counselors. One might criticize that the face expressions have not fully met the emotions that VCs want to express. At present, this chatbot can be designed to have expressions matching with the emotional content of the language, but it is time-consuming to go through special programming. We first focus on transforming the chatbot into VC at this stage. Thus, the researchers are presenting the expressions and body movements that the VC can do at present (Multimedia Appendix Files). They cannot fully match the expressions and emotions of the VC. In the future, we need to work with engineers to further overcome this problem.

The importance of this study is threefold, that is, the contribution to professional education of counseling, the contribution to the development and construction of the empathy corpus for chatbot, and the contribution the client's well-being. This research can provide counseling educators with a suitable training-assisted chatbot, empathy training content and courses. Since the training process of counselors needs to accumulate internships and practical learning experiences, the prospective counselors during the internship could have ethical problems facing real clients and experience a lot of pressure and challenge. This dilemma

could be solved by using virtual clients since they are not “real people.” Virtual clients will not be impatient and will not comment on prospective counselors for the skills that they have not mastered. This situation can enhance the performance of the prospective counselors and reduce the pressure of interpersonal interaction. It also thus helps the prospective counselor’s willingness to practice and enhances the practical experience so that they can practice repeatedly until mastering the skills. On the other hand, faced with the rapid development of technology, the emotional feedback of users after talking to chatbots is also an important piece of information for developing chatbots. With the establishment of the empathy corpus, the present research could make contribution to converting chatbots into VCs. The stronger the empathy ability of the counselor, the better able to establish a good counseling relationship with the client, and it can be expected that the counseling effect will be improved accordingly. Hence, the well-being of the client can benefit through effective counseling.

Future work would involve incorporating more diverse training data and developing culture-specific models to enhance the chatbot’s ability to understand and respond to varied cultural expressions of empathy.

Limitations of the research include the fact that this study focuses on developing chatbots as virtual clients, serving as auxiliary tools for empathy training of counselors, but it has not yet progressed to the experimental stage for integration with empathy training courses. This limitation underscores the need for future work to incorporate chatbot-based virtual standard cases into situational teaching, problem-based learning (PBL), and other instructional plans. This will enable the creation of empathy training courses and the assessment of their effectiveness.

A further research direction is to understand the effect of chatbot-assisted counseling skills training courses. The research method can adopt a quasi-experimental design with experimental and control groups to further understand the effect of applying chatbots in the empathy training of counseling personnel.

Conclusions

Due to the lack of development and empirical research on the development and use of chatbots as virtual clients, this research develops and uses chatbots in empathy training and evaluates its feasibility. It is an original research and development. The unique points of this paper are as follows.

1. Many current studies focus on the development of therapeutic chatbots, but there are still many challenges to be addressed. Traditionally, training for prospective counselors has relied on role-playing with fake clients or using actual clients for role-playing during internships, which can raise ethical concerns. In this study, we use chatbots as virtual clients to assist in the training of prospective counselors. These chatbots can be used in practice courses and internships for role-playing, as well as in teaching, training, and after-training skills development exercises. Using these chatbots may improve the effectiveness of counselor training.

2. The reports from literature used virtual patients to teach empathy through typing texts on a screen rather than through oral dialogue in order to improve the empathic communication skills of medical and nursing students. In contrast, our chatbot in this study conducts multiple rounds of oral dialogue with prospective counselors, more closely simulating a real counseling process.

3. Our chatbot simulates a juvenile client. As a physical humanoid robot with body movements and facial expressions, rather than a computer screen communicating through typing, the chatbot is more engaging for students to have an empathetic dialogue with as a virtual client.

As a result, virtual clients based on smart chatbots are developed in this study and can be used as an auxiliary tool for empathy training. In addition to increasing practice hours and opportunities for empathy training, using virtual clients allows prospective counselors to avoid ethical problems and pressure caused by directly facing the real-life client in the learning stage of developing counseling skills.

Through the VC’s response mechanism, that is, after the conversation the VC tells the interlocutor “you

understand my mood” or “you do not understanding my mood,” and the system of the chatbot can convert the conversation process of the voice record into a verbatim transcript, the prospective counselor can get instant feedback whether the empathy response is correct. With this mechanism, the interlocutor is given the opportunity to enhance or modify responses such that the empathy skills can be further improved. This function is not available in general counseling skills practice scenarios or counseling with real individuals.

References

- Alammar, J. (2018, June 27). *The illustrated transformer* [Blog post]. <https://jalammar.github.io/illustrated-transformer/>
- Bachelor, A. (1991). Comparison and relationship to outcome of diverse dimensions of the helping alliance as seen by client and therapist. *Psychotherapy: Theory, Research, Practice, Training*, 28(4), 534–549. <https://doi.org/10.1037/0033-3204.28.4.534>
- Barrett-Lennard, G. T. (1981). The empathy cycle: Refinement of a nuclear concept. *Journal of Counseling Psychology*, 28(2), 91–100. <https://doi.org/10.1037/0022-0167.28.2.91>
- Bernhardsson, E. (2015, October 1). *Nearest neighbors and vector models – part 2 – algorithms and data structures* [Blog post]. <https://erikbern.com/2015/10/01/nearest-neighbors-and-vector-models-part-2-how-to-search-in-high-dimensional-spaces.html>
- Brooke, J. (1996). SUS: A ‘quick and dirty’ usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), *Usability evaluation in industry* (pp. 189–194). CRC Press.
- Chang, S.-Y. (2018). Xuni shijing yunyong yu jiaoyu changyu keneng mianlin de wenti. *Taiwan Educational Review Monthly*, 7(11), 120–125.
- Chen, L.-H., Yeh, W.-L., Lin, Y.-M., Cheng, C.-Y., & Fan, H.-M. (2021). Using reality games in off-campus teaching to deepen students’ engaged learning and understanding: The case of “Inquiring after Mackay’s Footprints.” *Bulletin of Educational Research*, 67(1), 65–106. <https://doi.org/10.3966/102887082021036701003>
- Chen, S.-C., Kao, C.-C., & Wu, D.-Y. (2014). Helpful suggestions for Taiwanese counselor educators addressing the struggles that Chinese counselors face when trying to apply empathy skills in their counseling sessions. *Guidance Quarterly*, 50(2), 13–21.
- Chen, S.-H., Chen, S.-C., Lee, S.-C., Chang, Y., & Yeh, K.-Y. (2017). Impact of interactive situated and simulated teaching program on novice nursing practitioners’ clinical competence, confidence, and stress. *Nurse Education Today*, 55, 11–16. <https://doi.org/10.1016/j.nedt.2017.04.025>
- Chodos, D., Naeimi, P., & Stroulia, E. (2009). An integrated framework for simulation-based training on video and in a virtual world. *Journal of Virtual Worlds Research*, 2(1). <https://doi.org/10.4101/jvwr.v2i1.375>
- Clark, A. J. (2010). Empathy: An integral model in the counseling process. *Journal of Counseling & Development*, 88(3), 348–356. <https://doi.org/10.1002/j.1556-6678.2010.tb00032.x>
- Cui, Y., Che, W., Liu, T., Qin, B., & Yang, Z. (2021). Pre-training with whole word masking for Chinese BERT. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 29, 3504–3514.

- <https://doi.org/10.1109/TASLP.2021.3124365>
- Ekman, P. (1987). A life's pursuit. In T. A. Sebeok & J. Umiker-Sebeok (Eds.), *The semiotic web 1986* (pp. 3–46). De Gruyter Mouton. <https://doi.org/10.1515/9783110861310-002>
- Empathy. (2024). In *Wikipedia*. <https://en.wikipedia.org/wiki/Empathy>
- Fanning, R. M., & Gaba, D. M. (2007). The role of debriefing in simulation-based learning. *Simulation in Healthcare: Journal of the Society for Simulation in Healthcare*, 2(2), 115–125. <https://doi.org/10.1097/SIH.0b013e3180315539>
- Foster, A., Chaudhary, N., Kim, T., Waller, J. L., Wong, J., Borish, M., Cordar, A., Lok, B., & Buckley, P. F. (2016). Using virtual patients to teach empathy: A randomized controlled study to enhance medical students' empathic communication. *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, 11(3), 181–189. <https://doi.org/10.1097/SIH.0000000000000142>
- Frangoudes, F., Hadjaros, M., Schiza, E. C., Matsangidou, M., Tsivitanidou, O. E., & Neokleous, K. (2021). An overview of the use of chatbots in medical and healthcare education. In P. Zaphiris & A. Ioannou (Eds.), *Learning and collaboration technologies: Games and virtual environments for learning* (pp. 170–184). Springer International Publishing/Springer Nature. https://doi.org/10.1007/978-3-030-77943-6_11
- Gelso, C. J., & Carter, J. A. (1985). The relationship in counseling and psychotherapy: Components, consequences, and theoretical antecedents. *The Counseling Psychologist*, 13(2), 155–243. <https://doi.org/10.1177/0011000085132001>
- Gelso, C. J., & Carter, J. A. (1994). Components of the psychotherapy relationship: Their interaction and unfolding during treatment. *Journal of Counseling Psychology*, 41(3), 296–306. <https://doi.org/10.1037/0022-0167.41.3.296>
- Harcum, E. R. (1989). Commitment to collaboration as a prerequisite for existential commonality in psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, 26(2), 200–209. <https://doi.org/10.1037/h0085420>
- Helgadottir, F. D. (n.d.). *Computers and therapy*. AI-Therapy. <https://www.aitherapy.com/articles/computers-and-therapy>
- Jeffries, P. R. (2005). Development and testing of a hyper learning model for design of an online critical care course. *Journal of Nursing Education*, 44(8), 366–372. <https://doi.org/10.3928/01484834-20050801-07>
- Kokotovic, A. M., & Tracey, T. J. (1990). Working alliance in the early phase of counseling. *Journal of Counseling Psychology*, 37(1), 16–21. <https://doi.org/10.1037/0022-0167.37.1.16>
- Kurtz, P. D., Marshall, E. K., & Banspach, S. W. (1985). Interpersonal skill-training research: A 12-year review and analysis. *Counselor Education and Supervision*, 24(3), 249–263. <https://doi.org/10.1002/j.1556-6978.1985.tb00482.x>
- Lai, C.-M., Huang, H.-M., Liaw, S.-S., & Huang, W.-W. (2009). A study of user's acceptance on three-dimensional virtual reality applied in medical education. *Bulletin of Educational Psychology*, 40(3),

- 341–361. <https://doi.org/10.6251/BEP.20080422>
- Lan, Y.-L., Chang, Y., Chen, W.-L., & Wang, Y.-F. (2021). The feasibility of virtual reality in measuring psychological assessment: Student and psychologist views. *Bulletin of Educational Psychology*, 53(2), 481–510. [https://doi.org/10.6251/BEP.202112_53\(2\).0010](https://doi.org/10.6251/BEP.202112_53(2).0010)
- Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and evaluation of a User Experience Questionnaire. In A. Holzinger (Ed.), *HCI and usability for education and work: 4th symposium of the workgroup human-computer interaction and usability engineering of the Austrian computer society, USAB 2008* (Lecture Notes in Computer Science, vol. 5298, pp. 63–76). Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/978-3-540-89350-9_6
- Lee, G.-H., & Lin, M.-T. (2021). How empathy develops? A theatre-based training approach. *Bulletin of Educational Psychology*, 53(1), 127–150. [https://doi.org/10.6251/BEP.202109_53\(1\).0006](https://doi.org/10.6251/BEP.202109_53(1).0006)
- Liao, Y. J. (2014, September 23). *Tonglixin ji chuli jiqiao* [Blog post]. Xuite. <https://web.archive.org/web/20230823165048/https://blog.xuite.net/cfacai/wretch/260157137>
- Liu, Y.-C., & Chen, W.-L. (2014). Use of simulated patients/clients in suicide assessment interview training. *Formosa Journal of Mental Health*, 27(2), 283–313. [https://doi.org/10.30074/FJMH.201406_27\(2\).0005](https://doi.org/10.30074/FJMH.201406_27(2).0005)
- Liu, Y.-C., Lan, Y.-L., & Kao, S.-P. (2011). The situated learning approach in clinical psychology education: The use and challenge of the standardized patient training method. *Formosa Journal of Mental Health*, 24(2), 247–278. [https://doi.org/10.30074/FJMH.201106_24\(2\).0004](https://doi.org/10.30074/FJMH.201106_24(2).0004)
- Lu, C.-C., & Ma, S.-Y. (2019). Design STEAM course to train STEAM literacy of primary students: Taking “Animal mimicry beast” as an example. *Journal of Research in Education Sciences*, 64(3), 85–118. [https://doi.org/10.6209/JORIES.201909_64\(3\).0004](https://doi.org/10.6209/JORIES.201909_64(3).0004)
- Matarazzo, R. G., & Patterson, D. R. (1986). Methods of teaching therapeutic skill. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed., pp. 821–843). John Wiley & Sons.
- Matias, J. N., Manikonda, L., Hale, S., & Arnold, K. (2017, November 2). *AI in counseling and spiritual care* [Blog post]. Medium. <http://medium.com/ai-and-christianity/ai-incounseling-spirual-care-e324d9aea3b0>
- McCarthy, S., & LaChenaye, J. (2017). Adopting an ethic of empathy: Introducing counseling best practices as qualitative best practice. *Journal of Ethnographic & Qualitative Research*, 11(3), 188–198.
- McLellan, H. (Ed.). (1996). *Situated learning perspectives*. Educational Technology Publications.
- Ministry of Education. (2017, September 26). *Ministry of Education objectives for 2018 (January-December)*. <https://english.moe.gov.tw/cp-9-17161-B6F6B-1.html>
- Monti, E. J., Wren, K., Haas, R., & Lupien, A. E. (1998). The use of an anesthesia simulator in graduate and undergraduate education. *CRNA: The Clinical Forum for Nurse Anesthetists*, 9(2), 59–66.
- O’Shea, K., & Nash, R. (2015). *An introduction to convolutional neural networks*. arXiv. <https://doi.org/10.48550/arXiv.1511.08458>

- REST. (2024, July 14). In *Wikipedia*. https://en.wikipedia.org/wiki/Representational_state_transfer
- Roffers, T., & Waldo, M. (1983). Empathy and confrontation related to group counseling outcome. *The Journal for Specialists in Group Work*, 8(3), 106–113. <https://doi.org/10.1080/01933928308411739>
- Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology*, 21(2), 95–103. <https://doi.org/10.1037/h0045357>
- Rogers, C. R. (1975). Empathic: An unappreciated way of being. *The Counseling Psychologist*, 5(2), 2–10. <https://doi.org/10.1177/001100007500500202>
- Rogers, C. R. (1979). The foundations of the person-centered approach. *Education*, 100(2), 98–107.
- Rubin, N. J., Bebeau, M., Leigh, I. W., Lichtenberg, J. W., Nelson, P. D., Portnoy, S., Smith, I. L., & Kaslow, N. J. (2007). The competency movement within psychology: An historical perspective. *Professional Psychology: Research and Practice*, 38(5), 452–462. <https://doi.org/10.1037/0735-7028.38.5.452>
- Rucker, J. J. H., Iliff, J., & Nutt, D. J. (2018). Psychiatry & the psychedelic drugs. Past, present & future. *Neuropharmacology*, 142, 200–218. <https://doi.org/10.1016/j.neuropharm.2017.12.040>
- Sharf, R. S. (1985). Artificial Intelligence: Implications for the future of counseling. *Journal of Counseling & Development*, 64(1), 34–37. <https://doi.org/10.1002/j.1556-6676.1985.tb00999.x>
- Shen, C.-H. (2019). Future of counseling education and services: The impact and response of artificial intelligence development. *Journal of Indigenous Counseling Psychology*, 10(4), 34–54.
- Shi, H., Wang, C., & Sakai, T. (2020). A siamese CNN architecture for learning Chinese sentence similarity. In B. Shmueli & Y. J. Huang (Eds.), *Proceedings of the 1st conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing: Student Research Workshop* (pp. 24–29). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.aacl-srw.4>
- Shorey, S., Ang, E., Yap, J., Ng, E. D., Lau, S. T., & Chui, C. K. (2019). A virtual counseling application using artificial intelligence for communication skills training in nursing education: Development study. *Journal of Medical Internet Research*, 21(10), Article e14658. <https://doi.org/10.2196/14658>
- Song, L., Singleton, E. S., Hill, J. R., Koh, M. H. (2004). Improving online learning: Student perceptions of useful and challenging characteristics. *The Internet and Higher Education*, 7(1), 59–70. <https://doi.org/10.1016/j.iheduc.2003.11.003>
- Sumerall, S., Lopez, S. J., & Oehlert, M. E. (2000). *Competency-based education and training in psychology: A primer*. Charles C Thomas Publisher.
- Sung, X.-L., Lin, X.-T., Zheng, X.-Y., & Hsieh, L.-H. (2004). *The theory and implementation of school guidance* (work-revised ed.). Pingao Books Publishing House.
- Teng, H.-Y. (2011). Tantaoy zishang guanxi zhong di tonglixin. *Counseling & Guidance*, 30(3), 39–44. <https://doi.org/10.29837/CG.201103.0012>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is all you need*. arXiv. <https://doi.org/10.48550/arXiv.1706.03762>

- Yang, C.-J., & Chien, C.-L. (2019). Facebook makes me happier? The effects of realistic and online interaction on subjective well-being and their mediating variable. *Bulletin of Educational Psychology*, 51(2), 183–205. [https://doi.org/10.6251/BEP.201912_51\(2\).0001](https://doi.org/10.6251/BEP.201912_51(2).0001)
- Yang, T.-S. (2018, January 29). “Tonglixin” shi ge haodongxi, xiwang renren du you: Chengshi biaoda neixin ganshou yu xuyao de “fei baoli goutong.” *Cheer Kuaile Gongzuo Ren*. <https://www.cheers.com.tw/article/article.action?id=5087871>
- Zheng, P.-R. (2022). Tisheng tonglixin: Xunishijing shi jiefang? *Positive Psychology in Counseling and Education*, 1, 99–107. [https://doi.org/10.30099/PPCE.202206_\(1\).0009](https://doi.org/10.30099/PPCE.202206_(1).0009)

收稿日期：2024 年 01 月 08 日
一稿修訂日期：2024 年 01 月 15 日
二稿修訂日期：2024 年 01 月 16 日
三稿修訂日期：2024 年 06 月 14 日
四稿修訂日期：2024 年 07 月 09 日
五稿修訂日期：2024 年 07 月 19 日
六稿修訂日期：2024 年 08 月 12 日
七稿修訂日期：2024 年 08 月 21 日
接受刊登日期：2024 年 08 月 21 日

國立臺灣師範大學教育心理與輔導學系
教育心理學報, 2025, 56 卷, 3 期, 539-570 頁

研發聊天機器人作為輔導與諮商人員同理心訓練的模擬個案：從使用者觀點檢視其可行性

謝麗紅¹、劉恩宇²、林繼民³、劉嘉吉⁴

本研究旨在開發聊天機器人作為輔助培訓輔導與諮商人員同理心訓練的虛擬標準個案（VC），並評估其作為輔助輔導與諮商人員同理心訓練的可行性。本研究基於輔導與諮商人員培訓的領域知識，開發並測試了聊天機器人作為虛擬個案的應用。機器人由自然語言處理引擎驅動，並通過包含各種同理心的對話語料庫進行培訓。對話者可與 VC 進行多樣性、多輪對話來練習同理心技巧。成功開發 VC 後，招募了 60 名輔導與諮商相關科系學生體驗與 VC 互動，並進行用戶體驗調查，以了解用戶對 VC 易用性和有效性的主觀感受和看法。結果顯示測試者的用戶體驗問卷平均分數超過 5 分（1-7），達到滿意的狀態、在系統可用性量表方面，平均分數為 4（1-5），表示對系統的可用性表示同意。而質性回饋資料顯示測試者認為與 VC 互動感到放鬆無壓力，並願意繼續對話，認為 VC 適合練習同理心技巧，能降低練習時的焦慮感，通過這樣的練習可以改善他們的同理心技巧。足見本研究成功將聊天機器人轉變為 VC，可以作為同理心練習的輔助培訓工具，除了增加同理心培訓的練習時間、意願和機會外，VC 還可以讓準諮商人員在發展諮商技巧的學習階段，避免直接面對真實個案所帶來的倫理問題和壓力，亦有助於解決個案來源不足的教學困境。

關鍵詞：同理心、智能聊天機器人、虛擬個案、同理心培訓課程、諮商員教育

¹ 國立彰化師範大學輔導與諮商學系

² 臺北醫學大學醫學院醫學系

³ 國立臺灣大學資料科學學位學程

⁴ 國立彰化師範大學物理學系

通訊作者：劉嘉吉，國立彰化師範大學物理學系，liucj@cc.ncue.edu.tw。

