Title	Design of personalized interaction dialogue system for mobile otome games
Sub Title	
Author	王, エイエン(Wang, Wanying) 杉浦, 一徳(Sugiura, Kazunori)
Publisher	慶應義塾大学大学院メディアデザイン研究科
Publication year	2024
Jtitle	
JaLC DOI	
Abstract	
Notes	修士学位論文. 2024年度メディアデザイン学 第1129号
Genre	Thesis or Dissertation
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO40001001-00002024-1129

Master's Thesis
Academic Year 2024

Design of Personalized Interaction Dialogue System for Mobile Otome Games

Keio University
Graduate School of Media Design

WANYING WANG

A Master's Thesis
submitted to Keio University Graduate School of Media Design
in partial fulfillment of the requirements for the degree of
Master of Media Design

WANYING WANG

Master's Thesis Advisory Committee:

Professor Kazunori Sugiura	(Main Research Supervisor)
Professor Nanako Ishido	(Sub Research Supervisor)

Master's Thesis Review Committee:

Professor Kazunori Sugiura	(Chair)
Professor Nanako Ishido	(Co-Reviewer)
Professor Matthew Waldman	(Co-Reviewer)

Abstract of Master's Thesis of Academic Year 2024

Design of Personalized Interaction Dialogue System for Mobile Otome Games

Category: Design

Summary

With the rapid development of mobile devices and internet technology, mobile games, particularly otome games targeting female players, have witnessed significant growth in recent years. These games meet the emotional needs of female players through exquisite art styles, rich storylines, and immersive emotional interactions. However, challenges such as declining user retention rates and content homogenization persist in the otome game market, highlighting the need for innovative interaction designs to enhance user experience and long-term engagement.

This study focuses on designing a personalized interaction system for mobile otome games by developing a "Large Language Model (LLM)-based personalized dialogue system" to optimize player interaction experiences. A questionnaire survey was conducted, collecting data from 156 female players to analyze their motivations, preferences, and personalization needs. The results indicate that "romantic fantasy", "character attachment", "storyline experience", and "art style" are the core factors influencing player experience. Furthermore, personalization features, such as character appearance customization, personalized dialogue generation, and interactive services, significantly enhance player immersion and satisfaction.

This study proposes a framework for a personalized dialogue system based on LLMs and multi-agent techniques, enabling dynamic adjustments to dialogue content and interaction logic. This system facilitates more intelligent and personalized game experiences, strengthening emotional resonance between players and virtual characters while improving game interactivity and user retention rates.

The theoretical contribution of this study lies in expanding the research framework for otome game interaction mechanisms, while its practical significance provides developers with innovative approaches to optimizing otome game design. This study offers valuable insights for future developments in personalized interaction systems for mobile games.

Keywords:

otome games, personalized dialogue system, Large Language Models (LLMs), user engagement, game interaction design, immersive experience, personalization customization

Keio University Graduate School of Media Design

WANYING WANG

Contents

Acknowledgements **viii**

1 Introduction **1**

1.1. Research Background 1

1.2. Research Questions 2

1.3. Research Purposes 2

1.4. Research Framework 3

2 Related Works **4**

2.1. Otome Games 4

2.1.1 Definition and Characteristics 4

2.1.2 Classification and Types 5

2.1.3 Evolution of Otome Games 10

2.1.4 Research Perspectives for Otome Games 11

2.1.5 Academic and Practical Significance 12

2.2. Game Immersion 12

2.2.1 Definition and Theoretical Framework of Immersion 13

2.2.2 Game Interaction and Virtual Characters 16

2.3. Personalized Dialogue Generation 19

2.3.1 Dialogue Generation 19

2.3.2 Personalized Dialogue Generation based on Deep Learning 21

2.3.3 Personalized Dialogue Generation based on Agent 22

3 Survey Design **23**

3.1. Questionnaire Distribution and Collection 23

3.2. Survey Statistics and Analysis 24

3.2.1 Descriptive Analysis 24

3.2.2 Reliability and Validity Analysis 31
3.3. Personalized Dialogue System 47
3.3.1 Concept 47
3.3.2 Personalized Dialogue System based on LLMs Agent . . . 49

4 Experimentation on Interactive system 53
4.1. Experiment 53
4.1.1 Objective 53
4.1.2 Methodology 53
4.1.3 Procedure 53
4.1.4 Content 54
4.2. Experiment Results 66
4.2.1 Anova Analysis 66
4.2.2 Questionnaire Analysis 68

5 Conclusion 71
5.1. Results 71
5.2. Limitation 71
5.3. Future Works 72

References 73

Appendices 87
A. Programming for Multi-Agent Personalized Dialogue Generation for Otome Game 87

List of Figures

1.1	Research Framework	3
2.1	Angelique	6
2.2	Kiniro no Corda	6
2.3	IkemenRevolution	8
2.4	Love and Producer	8
2.5	Love and Deepspace	9
3.1	Pie chart of percent of Age question	25
3.2	Pie chart of percent of Highest Level of Education question	26
3.3	Pie chart of percent of Occupation question	27
3.4	Pie chart of percent of Playing Frequency question	28
3.5	Pie chart of percent of Playing Time question	29
3.6	Pie chart of percent of Monthly Income	30
3.7	Pie chart of percent of Monthly Expenditure	30
3.8	Mean Comparison of Play Motivation	35
3.9	Pearson Correlation Visualization	40
3.10	Concept of Personalized Dialogue System	48
4.1	Characters in Game A	55
4.2	Dialogue system in Game A(1)	56
4.3	Dialogue system in Game A(2)	57
4.4	Characters in Game A	59
4.5	Dialogue System in Game B(1)	60
4.6	Dialogue System in Game B(2)	61
4.7	Character: Apple	62
4.8	Character: Pear	63
4.9	Character: Grape	64

4.10 Mean of Interaction time 67
4.11 Mean of Interaction time 68

List of Tables

2.1 Comparison and Development Trends 9
2.2 state of your mind 18

3.1 Frequency of Demographic Characteristics 24
3.2 Frequency of Playing Characteristics 26
3.3 Reliability Analysis 31
3.4 Validity Analysis 34
3.5 Play Motivation Descriptive Statistic 34
3.6 Item Descriptive Statistic 36
3.7 Pearson Correlation 38
3.8 Pearson Correlation (continued) 39
3.9 Components of Personalization Dimension 41
3.10 Whole dimension regression 42
3.11 Preference for personalization regression 43
3.12 Willingness to pay for personalization regression 45
3.13 Influence of personalization regression 46
3.14 List of emotions and their PAD values [1] 51

4.1 One-Way ANOVA Analysis for Interaction 66
4.2 One-Way ANOVA Analysis for Questionnaire 68

Acknowledgements

I am indebt to Professor Kazunori Sugiura for guiding not only about research but with many aspects of my life.

Chapter 1
Introduction

1.1. Research Background

With the rapid development of the internet and the widespread adoption of mobile devices such as smartphones, mobile games—characterized by their simplicity, portability, low technical barriers, and ability to effectively fill fragmented time—have become particularly popular, especially among younger audiences. As smartphone performance continues to improve, the quality and quantity of mobile games have significantly increased, solidifying their position as a mainstream form of entertainment. According to a GameLook data report, the global mobile game market is projected to reach 98.7 billion by 2024, with Asia remaining the largest regional market. Four countries—China, the United States, Japan, and South Korea—account for 77.9% of the global mobile game market, with China alone contributing 37.4%. In 2018, China's gaming industry generated more than 25% of the global gaming revenue, ranking first worldwide in terms of market share and player base. Mobile games were the primary driving force behind this growth, accounting for 61% of the revenue share.

Female users have emerged as a significant consumer segment within the mobile gaming market. According to a Newzoo study, the number of female gamers worldwide surpassed 1 billion in 2019, representing 46% of the total player base. By 2020, this proportion increased to 47%, with female players demonstrating higher engagement in mobile games, where they constituted 48% of the player base. Within this context, *otome* games, a key subgenre of games targeting female audiences, have gained increased visibility on mobile platforms over the past five years. Although *otome* mobile games entered the market relatively late, their growth has been remarkable compared to other genres. As a prominent representative of the "her games" category, otome mobile games exhibit tremendous market potential, leveraging the needs of female players.

Despite the current economic scale of the mobile gaming industry exceeding 100 billion and the rapid success of otome mobile games, challenges remain, particularly in terms of declining user retention rates. This study aims to analyze the interactive experiences of otome mobile games from the players' perspective and propose game interaction designs that can better enhance player experiences and long-term engagement.

1.2. Research Questions

- What factors affect players' gameplay experience of the interaction system of mobile Otome games?
- What are the characteristics of the interaction functions in existing mobile Otome games?
- Can the design of optimized interaction system improve players' game experience?

1.3. Research Purposes

- Theoretical Significance: Within the body of literature on game interaction features, research on interaction mechanics in otome games or related theoretical studies remains underdeveloped. Given the substantial growth potential of the female gaming market, most existing studies focus on general games, lacking a research perspective tailored to female audiences. This study, grounded in the perspective of female players, aims to expand the scope of otome game research and incorporate female-oriented considerations into the discourse, thereby enriching the theoretical understanding of this field.
- Practical Significance: As mobile otome games are relatively nascent, the market is still in an exploratory phase. The rapid popularity of such games has revealed issues of high content and gameplay homogenization among

contemporaneous products. This study seeks to move beyond the traditional text-based adventure game format, exploring diversified interaction models for otome games. It examines the feasibility of transforming otome games towards more interactive and multifaceted gameplay models while achieving sustainable economic operations. Additionally, the study offers practical insights and reference points for the creation and development of future mobile otome games.

1.4. Research Framework

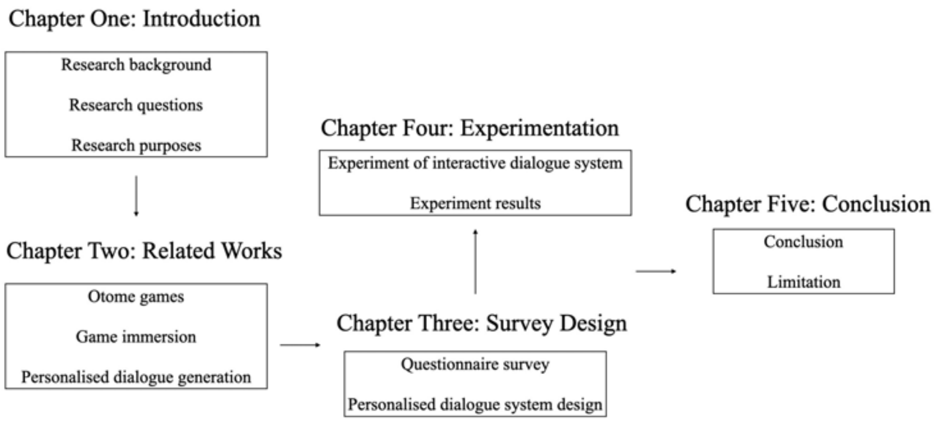

Figure 1.1 Research Framework

Chapter 2
Related Works

2.1. Otome Games

2.1.1 Definition and Characteristics

Otome games, originating in Japan, are a type of romance simulation game primarily targeting women. These games emphasize narrative experiences in which players take on the role of a female character and interact with male characters to advance emotional storylines or plot development. [2].

Definition and Expansion

Target Audience Otome games primarily cater to women aged 18–35, a demographic often characterized by high purchasing power and strong demand for cultural and entertainment content.

Content Structure The games revolve around "emotional interactions," with players advancing the storyline through multiple choices, leading to varied outcomes such as "happy endings" or "tragic endings."

Key Features

High Immersion The games provide players with an immersive experience from a female perspective, allowing them to participate in virtual romantic relationships.

Character Design Male characters are designed with diverse personalities, such as gentle, aloof, or humorous, to meet the emotional preferences of different players.

Multimedia Presentation Games feature high-quality visuals, voice acting by renowned actors, and engaging background music to enhance emotional investment. [3]

2.1.2 Classification and Types

Otome games can be categorized based on their business model and design objectives into "buy-to-play" and "free-to-play" with in-app purchases. [4]

Buy-to-Play Otome Games

Buy-to-play games require players to make a one-time payment to access the full game content. This model was prevalent in early otome games, particularly on PC and console platforms.

Characteristics

Fixed Pricing Players pay a one-time fee for the game, with no additional costs required.

Complete Content Access All game content, including storylines, characters, and endings, is unlocked upon purchase.

High Quality These games are often positioned as premium products, featuring deep narratives, well-developed characters, and exceptional artistic design.

Advantages Players can focus entirely on the storyline and character interactions without interruptions from in-app purchases. Ideal for players who value immersive experiences and wish to enjoy a complete story.

Representative Games Japanese franchises such as "Angelique", "Harukanaru Toki no Naka de", and "Kiniro no Corda". Independent games like "Hana Awase" and "Anohana".

(Source: Otome game: Angelique)

Figure 2.1 Angelique

(Source: Otome game: Kiniro no Corda)

Figure 2.2 Kiniro no Corda

Free-to-Play Otome Games with In-App Purchases

Free-to-play games attract players by offering basic game content for free and monetize through in-app purchases. This model has become dominant for otome games on mobile platforms in recent years.

Characteristics

Free Download Players can download and experience part of the game for free.

In-App Purchases Additional content, such as items, chapters, or character cards, can be unlocked through payments.

Continuous Updates Games often maintain user engagement with regular events and new storylines.

Advantages

Low Entry Barrier Players can try the game before deciding to spend money.

Flexible Spending Options Players can choose whether to spend on items, additional storylines, or character draws based on their preferences.

Disadvantages Game experiences can be disrupted by monetization mechanisms, such as key storylines requiring payment to unlock. Criticism of "emotional commodification" arises, where the emotional experiences offered by the game may feel overly commercialized.

Representative Games Japanese market:*Ikemen Revolution*
Chinese Market: Popular Chinese otome games like *Love and Producer*,*Light and Night* and *Love and Deepspace.*

(Source: Otome game: Ikemen Revolution)

Figure 2.3 IkemenRevolution

(Source: Otome game: Love and Producer)

Figure 2.4 Love and Producer

(Source: Otome game: Love and Deepspace)

Figure 2.5 Love and Deepspace

Comparison and Development Trends

Table 2.1: Comparison and Development Trends

Dimension	Paid Games	Free-to-Play Games
Entry Barrier	High (one-time purchase)	Low (free download)
User Engagement	Driven by story and artistry	Maintained via events /updates
Revenue Model	One-time payment	Continuous in-app purchases
Suitable Platform	PC, Console	Mobile
User Feedback	Emphasis on immersive experiences	Experiences influenced by monetization

In recent years, with the widespread adoption of mobile games and the expansion of the player base, the free-to-play model has gradually become the mainstream for otome games. However, buy-to-play games still hold a certain position in niche markets, particularly among independent titles that emphasize narrative experiences and high-quality production. [4] [5]

2.1.3 Evolution of Otome Games

Origin and Early Development (1994–2000)

The genre of otome games originated in 1994, when Japan's KOEI Corporation officially defined the category with the release of Angelique. At its core, otome games offer a "female-centered" virtual romantic experience by allowing players to assume the perspective of a female protagonist. During this period, otome games were primarily developed for PC and console platforms, often presented in the form of text-based adventure games. [3] [5]

Genre Expansion and Global Dissemination (2000–2015)

Diversification of Gameplay As the player base grew, otome games began incorporating additional gameplay elements such as life simulation and role-playing mechanics, exemplified by titles like Kiniro no Corda.

Global Spread Through translations and online communities, otome games gradually gained traction in markets such as China and South Korea, transitioning from niche entertainment to a cultural phenomenon with increasing attention.

Development in China (2015–Present)

Introduction and Breakthrough The 2017 release of Love and Producer marked a turning point, bringing otome games from niche culture to mainstream popularity in China. The game attracted a large female audience with its exquisite visuals and deeply engaging narratives. [2]

Industry Scale Data indicates that the market scale for female-oriented games expanded from billions of RMB in 2017 to nearly one trillion RMB by 2023, making it a significant segment within the gaming market.

Development Characteristics

Localization Game content increasingly incorporates Chinese cultural elements to better appeal to local audiences.

Domimance of the Free-to-Play Model Developers rely on frequent updates and diverse events within the in-app purchase framework to enhance player engagement.

2.1.4 Research Perspectives for Otome Games

Player Motivation and Behavior

The motivations of otome game players are multidimensional and complex. Research identifies the following primary factors:

Emotional Needs Interaction with virtual characters fulfills emotional needs for companionship and affection. Female players tend to choose characters that align with their aesthetic preferences and personality traits. [5] [6]

Escapism Otome games provide a virtual world where players can temporarily escape real-world pressures. Game content is designed with expansive narrative backgrounds to enhance immersive experiences. [2]

Social Interaction Players share their gaming experiences via social media and community discussions, building virtual social networks. [5]

Consumption Intentions and Behavior

Research based on the Technology Acceptance Model (TAM) highlights the following factors influencing consumption behavior:

Perceived Factors Perceived ease of use and usefulness significantly influence players' willingness to make purchases. Key drivers for sustained spending include gameplay convenience, engaging storylines, and high-quality character design. [4]

Emotional Dependence Players unlock storylines and purchase items to deepen emotional interactions with characters, fostering emotional attachment to virtual characters.

Gender Culture and Social Impact

Research on otome games offers both cultural and social insights:

Gender Empowerment Otome games challenge traditional "male-dominated" design paradigms, fostering the expression of female agency and consciousness.

Controversy and Criticism Some studies argue that the idealized portrayal of male characters in otome games may reinforce gender stereotypes, potentially affecting players' perceptions of real-world relationships. [2]

2.1.5 Academic and Practical Significance

Academic Significance

Otome games provide valuable samples for studying female culture, revealing unique consumption behaviors in the digital entertainment sector. They validate the applicability of theories such as the Uses and Gratifications Theory and the Technology Acceptance Model in the context of female-oriented games. [4]

Practical Significance

The findings guide developers in optimizing design to enhance emotional resonance and immersive experiences in games. They provide a basis for market strategies, contributing to the sustainable development of female-oriented games. [7]

2.2. Game Immersion

With the growing popularity of video games, immersion has become a key metric for evaluating game quality. It describes the psychological state in which players are deeply engaged in a virtual world. Immersion not only significantly influences player satisfaction and loyalty but also drives the continuous optimization of game design and user experience. This review systematically examines relevant studies to explore the theoretical foundations, influencing factors, and design applications of immersion, aiming to guide future research in this field.

2.2.1 Definition and Theoretical Framework of Immersion

Definition of Immersion

Immersion refers to the experience of players being fully absorbed and deeply interacting with a virtual world during gameplay. Brown and Cairns (2004) classified immersion into three levels through qualitative research: "engagement," "engrossment," and "total immersion." Among these, total immersion is characterized by a state where players are entirely disconnected from reality and fully immersed in the virtual environment. [8]

Application of Flow Theory

Csikszentmihalyi's flow theory provides a critical foundation for immersion research. Flow experiences consist of eight core elements: clear goals, immediate feedback, a balance between task difficulty and player skills, focus, a sense of control, time distortion, loss of self-consciousness, and intrinsic rewards. This theory is often used to explain the mechanisms triggering immersion. [9]

SCI Model

Ermi and Mäyrä (2005) proposed the SCI model, categorizing immersion into three types: [10]

Sensory Immersion Achieved through high-quality sound effects and visuals.

Challenge-Based Immersion Derived from overcoming complex tasks and challenges in games.

Imaginative Immersion Relies on emotional resonance fostered by narrative and character design.

Key Factors Influencing Immersion

Game Design Features

Visual and Audio Design High-resolution graphics, dynamic sound effects, and realistic virtual environments significantly enhance sensory immersion. For instance, virtual reality (VR) technology pushes sensory immersion to new heights.

Narrative and Character Development Narrative and character design can evoke emotional resonance in players, enhancing imaginative immersion. For example, personalized character customization not only deepens immersion but also fosters player loyalty.

Interaction Mechanisms Intuitive and seamless interaction design reduces players' cognitive load. Features such as combining touchscreens with motion controls allow players to interact more naturally with mobile games.

Player Individual Differences

Player Traits Cognitive styles, emotional needs, and gaming preferences influence immersion. For example, players who value challenges are more likely to experience challenge-based immersion in strategy games, whereas narrative-driven players focus more on emotional connections from storytelling.

Psychological State Emotional states and cognitive workload are crucial variables affecting immersion. Studies show that an optimal level of challenge can trigger flow experiences, while excessive or insufficient difficulty may lead to anxiety or boredom.

Environmental and Technological Support

Hardware Technology VR and augmented reality (AR) technologies significantly enhance immersion by offering lifelike experiences. Advances in hardware such as high-refresh-rate displays and haptic feedback devices further support immersive experiences.

Environmental Factors External disturbances like noise and lighting can affect immersion. Optimizing the gaming environment can help players focus better on the game.

Practical Applications of Immersion

Mobile Games Hu Qianqian (2019) suggested that mobile game design should emphasize narrative, visual, and auditory immersion. Dynamic sound effects, optimized difficulty levels, and well-designed reward systems can effectively enhance player experiences.

First-Person Shooter (FPS) Games Experimental studies by Nacke et al. (2009) revealed that the combat design in FPS games could enhance immersion by elevating players' emotional arousal. Physiological measures such as galvanic skin responses and electromyographic activity were used to quantify emotional reactions.

Game Loyalty Immersion fosters player loyalty by satisfying psychological needs such as exploration, escapism, and achievement. Studies indicate that personalized game design (e.g., customizable character appearance and abilities) can enhance sustained player engagement through immersion.

Future Research Directions

Cross-Cultural Studies Most existing immersion research focuses on Western cultural contexts. The perception and mechanisms of immersion in different cultures warrant further exploration.

Technological Integration The incorporation of artificial intelligence (AI) presents new possibilities for personalized immersive experiences. For instance, AI-generated dynamic narratives and real-time difficulty adjustments tailored to player skill levels could further enhance immersion.

Long-Term Effects Current research predominantly examines short-term immersion experiences. The evolution of immersion over prolonged gameplay and its impact on player behavior remain underexplored.

2.2.2 Game Interaction and Virtual Characters

Game interaction, as a significant branch of human-computer interaction, has achieved remarkable progress in recent years, particularly in role-playing, emotional interaction, and personalized experiences, driven by advancements in artificial intelligence. This paper systematically reviews the primary research directions in game interaction technologies based on current literature, including virtual environment perception and interaction, personalized dialogue generation, character-based simulated dialogues, emotion- and personality-driven response generation, and frameworks for preference and personality modeling.

Virtual Environment Perception and Interaction

The realism and diversity of virtual environments are fundamental to enhancing game interaction experiences. The V-IRL (Virtual Intelligence in Real Life) platform bridges virtual and real-world environments, providing an interaction platform centered on visual and geospatial data. By integrating maps and street-view data, this platform enables virtual agents to perform complex tasks within virtual representations of real cities, such as route optimization, location recommendation, and environmental analysis. Benchmarks across various geographic and visual models demonstrate V-IRL's superior flexibility, task diversity, and realism.

Specifically, V-IRL supports multi-task capabilities based on language and vision models. For instance, the simulated "path planning" task showcases how agents leverage shortest-path algorithms to efficiently explore urban areas, while the "location recommendation" task combines language models with geospatial data to offer customized suggestions based on user preferences. These applications provide critical references for dynamic scene generation and complex interaction logic design in games.

Personalized Dialogue Generation

Personalization is a crucial factor in enhancing player immersion. The Personal-Dialog dataset and model leverage large-scale dialogue data to construct conversational systems with multidimensional personality attributes, such as gender, age, and interests. A personalization feature fusion module integrates structured personality information with contextual inputs to generate dialogues suited to different scenarios.

This approach demonstrates significant potential for application in games. For example, by generating dialogue content reflective of player traits, non-player characters (NPCs) can exhibit enhanced intelligence and adaptability. Such methods are particularly important in open-world games, role-playing games (RPGs), and educational games, providing players with customized and immersive experiences.

Character-Based Simulated Dialogue

Role-playing scenarios in games require virtual characters to exhibit specific linguistic styles and behavioral patterns. The ChatHaruhi project introduces a dialogue generation method based on character memory banks. By extracting original scripts and integrating improved prompting techniques, this approach enables language models to better emulate specific characters' tones and expressions.

The study employs two strategies: extracting iconic dialogues from scripts to form a memory bank and dynamically generating dialogues based on user input and character backstories. This method not only improves semantic consistency in conversations but also enhances emotional connections between players and virtual characters. For instance, in narrative-driven games, players can engage in more authentic interactions with characters, deepening their understanding of both the characters and the game world.

Emotion- and Personality-Driven Response Generation

Emotional expression is an indispensable aspect of game interaction. The GERP model incorporates the OCEAN Big Five personality traits (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) and an emotion prediction mechanism to introduce personalization and emotional dynamics into dialogue generation. By dynamically predicting emotional states, the model generates responses aligned with personality traits, thereby enhancing the anthropomorphic quality of virtual characters.

Experimental results show that the GERP model effectively captures conversational contexts and generates appropriate emotional responses. For example, in horror or adventure games, emotional expressions by characters can enhance player immersion. In educational or medical simulation games, emotionally driven dialogues help create empathetic interaction scenarios.

Framework for Preference and Personality Modeling

Characters in games require not only dynamic behavioral capabilities but also evolving preferences and personalities shaped by player behavior and environmental influences. The AFSPP framework creates a virtual sandbox world to simulate the impact of social networks and subjective consciousness on personality and preference formation. Agents within the framework possess memory, planning, reflection, and interaction abilities, allowing them to dynamically adjust their behavior based on environmental inputs and internal states.

The study indicates that environmental factors (e.g., social networks) and cognitive abilities (e.g., planning) play critical roles in shaping preferences. For instance, in strategy games, agents can adapt to players' gaming styles over time, forming personalized interaction strategies. In open-world games, the diverse preferences and personalities of agents enrich the game ecosystem, providing players with a more realistic simulation experience.

Table 2.2 state of your mind

when	where	who	what
Bad	Good	Very Good	Good Good

2.3. Personalized Dialogue Generation

2.3.1 Dialogue Generation

Dialogue generation, as one of the core tasks in the field of Natural Language Processing (NLP), has gradually evolved from the early pattern-matching-based approaches to the modern related approaches based on deep learning and large-scale language models. In the early days, dialog systems mainly relied on predefined dialog patterns, rules, and databases to generate responses by matching user input. For example, dialog systems such as PARRY [11], SHRDLU [12], and LUNAR [13] in the 1970s used pattern matching to search and generate responses based on the context of user input. With the development of the pattern matching technology, such as BORIS [14], ALICE [15], and Intent Detection-based dialog systems [16], these data-driven and information-retrieval conversational systems based on pattern recognition usually achieve good performance in terms of breadth and depth of conversations. However, these systems are often limited by the complexity of manually formulated structural rules and the abundance of the prior datasets, resulting in dialog generation results that are more rigid and do not conform to the natural language habits of humans.

Since the advent of the 21st century, the rise of deep learning algorithms has significantly advanced the field of Natural Language Processing (NLP), particularly in the development of dialogue generation systems. A series of NLP models based on the sequence-to-sequence (seq2seq) [17] paradigm, such as Recurrent Neural Networks (RNNs) [18], Long Short-Term Memory (LSTM) [19] networks, Gated Recurrent Units (GRUs) [20], and Transformer [21], have been proposed. These models not only demonstrate robust capabilities for processing natural language in theory but have also achieved remarkable success in practice. As a result, numerous deep neural network-based dialog generation models have emerged [22–30], bringing innovative breakthroughs in the field of dialog generation. These models combine deep learning techniques with word vector techniques in NLP and utilize multiple rounds of dialog mining, attention mechanisms, and predefined linguistic features (e.g., lexical, syntactic, etc.) to dig deeper into the implicit information in the user's dialog. This fusion of techniques improves the accuracy of text recognition and optimizes the satisfaction of content generation. As a result, the dialog system can generate more natural, fluent, and contextually relevant responses, which significantly improves the quality of the dialog and the user's interaction experience.

Although neural network-based dialogue generation models have demonstrated excellent performance on task-specific datasets, these models often face the problem of insufficient generalization capabilities in practical applications. This limitation is mainly reflected in the models' lack of extensive human-like basic knowledge and contextual adaptation, making them perform less than optimal when dealing with complex and diverse dialog tasks. Compared to humans, these models can usually only infer and generate within the scope of their training data, making it difficult for them to flexibly cope with contexts beyond their training scope or creative and scalable conversational demands. However, in recent years, the rise of Large Language Models (LLMs) [31–33], especially the emergence of models such as ChatGPT [34], has revolutionized the technology of dialogue generation in the field of natural language processing. Pre-trained on linguistic datasets of unprecedented size, LLMs can effectively learn and integrate diverse knowledge systems, and exhibit excellent language generation capabilities. Language models trained on large-scale world knowledge corpora can not only generate syntactically fluent and consistent dialog responses but also adapt to different dialog scenarios and needs based on diverse user inputs. This significant improvement in capability marks the shift of dialogue generation from traditional task-oriented models to more generalized intelligent dialogue systems [35], opening a new era for the future development of natural language processing.

The dialogue capabilities of existing large-scale language models (LLMs) cannot fully meet the requirements of actual interaction scenarios, especially in personalized dialogue generation. Specifically, existing LLMs have deficiencies in accurately understanding and dynamically adapting to the user's personalized preferences, background information, and emotional characteristics of the interaction, making it difficult to generate personalized responses that meet the user's expectations in complex interaction scenarios [36]. In addition, since their generation mechanisms mainly rely on large-scale general corpus training, existing models exhibit limitations when handling dialog tasks with strong contextual dependency [37]. Therefore, personalized dialogue generation is an important research direction in the current NLP field.

2.3.2 Personalized Dialogue Generation based on Deep Learning

Personalized responses can significantly enhance the interaction between the user and the dialogue agent and improve the interaction experience [36, 37]. Although different researchers have different definitions and implementations of "personalization", in essence, personalized response generation can be considered a conditional text generation task. In this task, the generated response not only needs to be based on the given dialog context (i.e., user input) but also needs to be conditional on the role settings explicitly provided by the user or the user attributes implicit in the dialog history. Therefore, deep-learning-based personalized dialogue generation models, are usually classified into two categories, explicit and implicit conditional generation [38], based on how the conditional information is obtained.

- **Explicit condition generation**: Explicit conditional generation methods generate responses that match specific user characteristics by utilizing information explicitly provided by the user (e.g., profile, preferences, emotional state, or role instructions) as a generating condition that is encoded and incorporated into the generation process. For example, [39–47] methods fall into this category. Its main advantage is the direct control over the generated content, which better meets the user's personalization needs. However, the effectiveness of this approach is highly dependent on the adequacy and accuracy of user input and training data, and insufficient or biased conditional information may degrade the generation quality.
- **Implicit conditional generation**: Implicit conditional generation methods allow models to automatically learn potentially high-dimensional features by mining implicit features in users' historical conversations and persona information without relying on predefined human a priori knowledge. For example, [48–53] models can generate responses that are more in line with the user's personality based on these features. Compared to explicit

conditional generation, implicit methods are more robust and generalizable to handle situations where the user does not provide explicit information, but they also face the challenge of rational and effective feature extraction.

However, both explicit and implicit conditional generation methods rely on large-scale preprocessed personalized datasets and require training models from scratch. This imposes high data labeling and computational resource costs. In addition, their personalization capabilities are still limited by the quality and performance of the pre-trained datasets in practical applications.

2.3.3 Personalized Dialogue Generation based on Agent

With the widespread use of Large Language Models (LLMs) such as ChatGPT [34], studying the impact of their capabilities in role-playing and contextual learning on personalized dialogue generation is gaining significant attention. As shown in the study [54–56], LLMs are capable of reflecting on role or personality traits through input prompts (prompts), which is particularly significant in tasks such as personality testing, writing, and reasoning. Therefore, the use of LLMs for multi-intelligence simulation or collaborative generation of diverse personalized dialogues has become a research hotspot. For example, [54–63] have constructed the Agent dialogue system of personalized dialogue generation by fine-tuning LLMs or combining techniques such as retrieval-and-generation (RAG) [64]. However, in emotional interaction scenarios such as otome games, how to satisfy users' needs for emotion and intimacy [65] through personalized dialogue-generated intelligences, and further enhance the immersion and interactivity of the game, is still an area that needs to be explored urgently.

Chapter 3
Survey Design

To quantify and construct the playing characteristics and user profiles of otome game players, this study employed a survey questionnaire to collect data for empirical analysis. The first part of the questionnaire focused on basic demographic information and playing habits, covering aspects including gender, age, education level, occupation, playing frequency, play time, disposable income, and game expenditure. The second part consisted of a scale measuring play motivations. To ensure the quality of the questionnaire design, the scales used in this study were developed based on established questions from previous research. These scales were adapted from existing literature with proven reliability and validity, including studies conducted by Ching-I Teng [7], Miu Shenguang [5]. To ensure reliability, the questionnaire included at least three questions for each dimension. A five-level Likert scale was employed, allowing respondents to select the option that best reflected their gaming experience and attitudes. Responses ranged from "Strongly Disagree" (scored as 1) to "Strongly Agree" (scored as 5). This study places a particular emphasis on the impact of personalization features on the game and players. Accordingly, the questionnaire addressed three components of personalization dimension : preferences, willingness to pay, and the influences on players, with a total of 11 sub-questions. The final questionnaire totally included eight dimensions of play motivations, comprising a total of 35 sub-questions. A complete version of the questionnaire can be found in the appendix.

3.1. Questionnaire Distribution and Collection

The participants in this study were female users who play otome games. The questionnaire targeted and collected samples that met the following two criteria: (1)

having experience playing or watching otome games and (2) being female users. To gather a substantial amount of data, the questionnaire was distributed online via the Wenjuanxing platform (https://www.wjx.cn/). It was primarily disseminated through Chinese social networking platforms, including WeChat, Weibo, and Douyin, using social sharing to collect study samples. To ensure the sample consisted exclusively of female users, a prior gender screening question was set in the questionnaire. A total of 172 responses were collected through questionnaire distribution and social sharing in various communities. After excluding 16 responses from participants whose gender was not female, 156 valid responses were obtained. The valid responses accounted for 90.7 percentages of the total submissions and were used for data analysis.

3.2. Survey Statistics and Analysis

For the 156 valid responses collected, the exported questionnaire data were edited and categorized using Microsoft Excel. Subsequently, the processed sample data were subjected to statistical analysis using the SPSSAU platform. Finally, regression analysis was conducted using STATA to examine the data.

3.2.1 Descriptive Analysis

Demographic Characteristics

Frequency analysis is used to study the demographic characteristics of the respondents, determining the frequencies and percentages, shown as in Table 2-1. Question A2 on gender was excluded as all valid questionnaires selected for this study were from female respondents. For Question A4 on occupation, the categories of "Employee" and "Freelancers" were merged to “Graduate”, to facilitate comparison with the student group.

Table 3.1: Frequency of Demographic Characteristics

Item	Option	Frequency	Percent (%)	Cumulative Percent (%)

Continued on next page

Table 3.1: Frequency of Demographic Characteristics (Continued)

A1 Age	1	17 years old or younger	3	1.92	1.92
	2	18-22 years old	40	25.64	27.56
	3	23-29 years old	100	64.1	91.67
	4	30-39 years old	13	8.33	100
A3 Highest Level of Education	1	Junior high school or below	2	1.28	1.28
	2	High school/vocational school	8	5.13	6.41
	3	Secondary vocational school	9	5.77	12.18
	4	Bachelor's degree	71	45.51	57.69
	5	Master's degree or above	66	42.31	100
A4 Occupation	1	Student	83	53.21	53.21
	2	Graduate	73	46.79	100

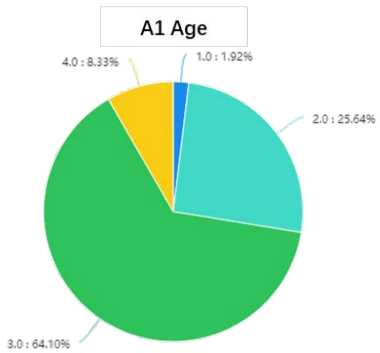

Figure 3.1 Pie chart of percent of Age question

In terms of age, the majority of survey participants fall within the 23-29 age group, accounting for over 64%. This indicates that the questionnaire primarily targets university students and working populations, most of whom have discretionary free time and income or pocket money. The second most represented age group is 18-22 years old.

Regarding education level, the participants are mostly bachelor's and master's

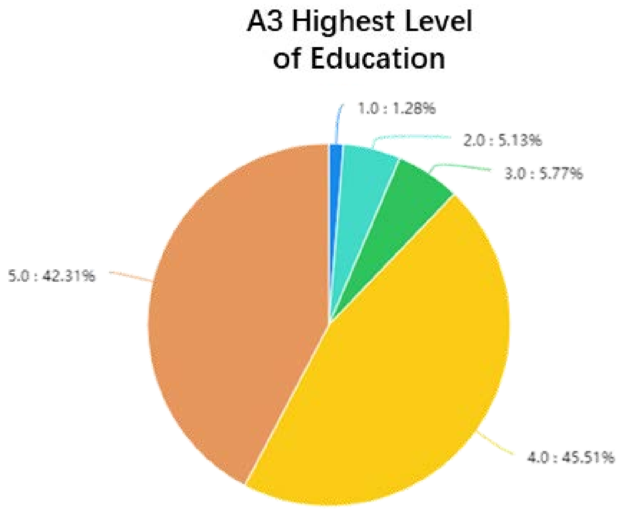

Figure 3.2 Pie chart of percent of Highest Level of Education question

degree holders, collectively exceeding 90%. This aligns with the age data, suggesting that the survey primarily focuses on university students and graduates.

In terms of occupation, the proportion of students and graduates is nearly balanced, with a difference of only about 7%. This balance ensures that the survey encompasses a diverse range of respondents, enabling the collection of comprehensive information and opinions from both student and working populations.

Playing Characteristics

For plaing characteristics of the respondents, frequency table is shown as in Table 2-2. Question A9 on favorite games will be used as a reference for future game design and is not included in the quantitative analysis of this section.

Table 3.2: Frequency of Playing Characteristics

Item	Option	Frequency	Percent (%)	Cumulative Percent (%)

Continued on next page

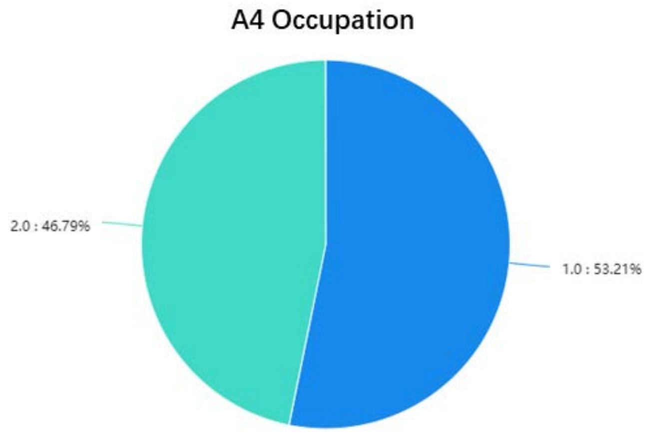

Figure 3.3 Pie chart of percent of Occupation question

Table 3.2: Frequency of Playing Characteristics (Continued)

A5 Playing Frequency	1	Rarely (once every few months)	47	30.13	30.13
	2	Occasionally (a few times a month)	22	14.1	44.23
	3	Sometimes (a few times a week)	23	14.74	58.97
	4	Frequently (at least once a day)	64	41.03	100
A6 Playing Time	1	Less than 15 minutes	52	33.33	33.33
	2	15-30 minutes	52	33.33	66.67
	3	30 minutes to 1 hour	30	19.23	85.9
	4	More than 1 hour	22	14.1	100
A7 Monthly Income	1	Less than ¥2,000 RMB	27	17.31	17.31
	2	¥2,000-¥5,000 RMB	39	25	42.31

Continued on next page

Table 3.2: Frequency of Playing Characteristics (Continued)

	3	¥5,000-¥8,000 RMB	32	20.51	62.82
	4	More than ¥8,000 RMB	58	37.18	100
A8 Monthly Expenditure	1	Less than ¥50 RMB	56	35.9	35.9
	2	¥50-¥100 RMB	22	14.1	50
	3	¥100-¥200 RMB	25	16.03	66.03
	4	¥200-¥500 RMB	27	17.31	83.33
	5	¥500-¥1,000 RMB	10	6.41	89.74
	6	More than ¥1,000 RMB	16	10.26	100

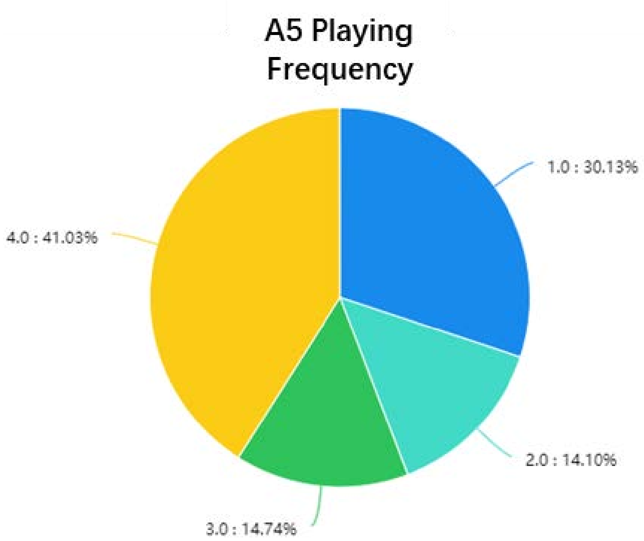

Figure 3.4 Pie chart of percent of Playing Frequency question

In terms of play frequency, participants are mainly distributed at either extreme. Recently, 41% of players reported daily playing otome games, while 30% reported playing very rarely. The large proportion of high-frequency players indicates that this survey effectively captures the gaming experiences of hardcore players. Meanwhile, the lightest players, who may have only tried the game a few times, provide perspectives from people who are not deeply familiar with Otome games.

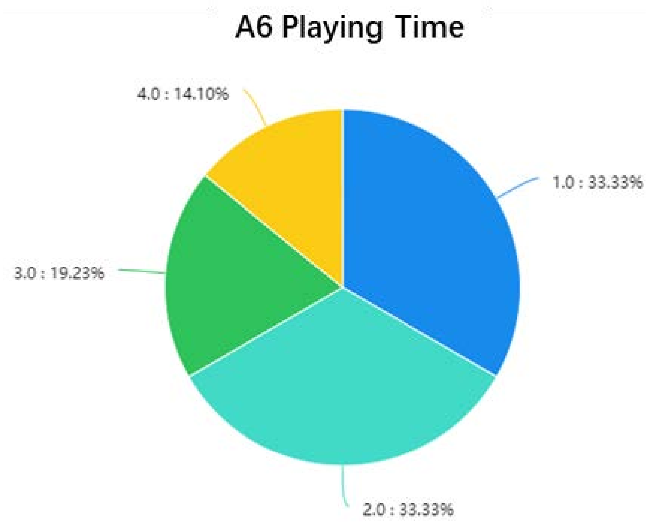

Figure 3.5 Pie chart of percent of Playing Time question

Regarding playtime, the respondents exhibit significant variability in daily play duration during their active periods, ranging from less than 15 minutes to more than 1 hour. Among them, over 66% of players reported spending less than 30 minutes per day on Otome games.

In terms of monthly disposable income, 37% of participants reported earning over 8,000 RMB, which is the highest proportion, while only 17% reported earning less than 2,000 RMB. This indicates that the respondents are relatively independent and possess stable purchasing power.

Regarding monthly expenditure, half of the respondents reported spending 100 RMB or less on otome games, while only 16% spent more than 500 RMB. This suggests that most participants have limited purchase enthusiasm for otome games or rarely buy related merchandise. Comparing these results with the monthly disposable income, it can be inferred that most players exhibit relatively mature spending habits.

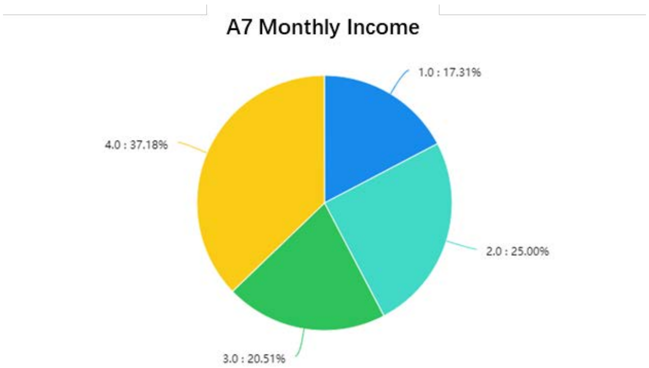

Figure 3.6 Pie chart of percent of Monthly Income

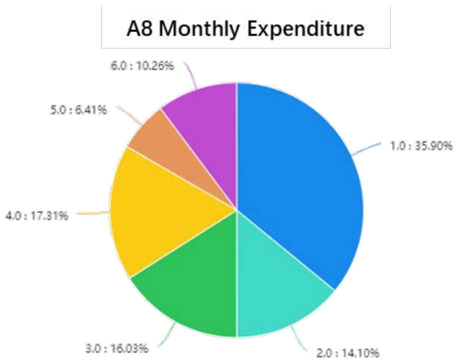

Figure 3.7 Pie chart of percent of Monthly Expenditure

3.2.2 Reliability and Validity Analysis

Reliability Analysis

Reliability Analysis mainly refers to the reliability of the test research, which is used to measure the consistency and stability of the questionnaire items. Consistency refers to whether the content measured by each item of the variable is in the same direction; stability refers to whether the results before and after the observation are consistent after multiple measurements of the same test group at different times/occasions. [66] [67] The reliability analysis in this paper mainly tests and analyzes the internal consistency of the scale. This study uses the Cronbach α value to test the internal consistency of each dimension, which is one of the most essential and pervasive statistics in research involving test construction and use. Devellis proposed that $0.65 < \alpha \leq 0.70$ is the minimum acceptability, $0.70 < \alpha \leq 0.80$ is very reliable, $0.80 < \alpha \leq 0.90$ is very reliable, and $\alpha > 0.90$ is very reliable. The Alpha coefficient should be greater than 0.7 for the whole study to be accepted. [68] This study utilized SPSSAU analysis platform to conduct a reliability test on the survey questionnaire. The results are presented in Table 2-3.

Table 3.3: Reliability Analysis

Dimension	Item	Corrected item-total correlation (CITC)	Alpha if Item Deleted	Cronbach Alpha
B.Romantic Fantasy	B1	0.758	0.804	0.86
	B2	0.659	0.842	
	B3	0.773	0.795	
	B4	0.654	0.848	
C.Character Attachment	C1	0.659	0.745	0.807
	C2	0.577	0.783	
	C3	0.658	0.742	

Continued on next page

Table 3.3: Reliability Analysis (Continued)

	C4	0.61	0.765	
D.Storyline	D1	0.75	0.764	0.849
	D2	0.686	0.819	
	D3	0.723	0.785	
E.Art Style	E1	0.817	0.961	0.939
	E2	0.921	0.877	
	E3	0.891	0.898	
F.Emotional Regulation	F1	0.759	0.811	0.866
	F2	0.688	0.845	
	F3	0.745	0.819	
	F4	0.696	0.838	
G.Social Interaction	G1	0.758	0.804	0.868
	G2	0.737	0.825	
	G3	0.755	0.811	
H.Challenge and Achievement	H1	0.722	0.724	0.827
	H2	0.781	0.669	
	H3	0.566	0.884	
I.Personalization	I1	0.563	0.881	0.888
	I2	0.61	0.878	
	I3	0.523	0.883	
	I4	0.542	0.883	
	I5	0.674	0.875	
	I6	0.653	0.875	
	I7	0.612	0.878	
	I8	0.633	0.877	

Continued on next page

Table 3.3: Reliability Analysis (Continued)

	I9	0.672	0.875
	I10	0.645	0.876
	I11	0.62	0.877
Total			0.936

From the analysis, the overall Cronbach's α coefficient for the questionnaire is 0.936, indicating high reliability and suitability for further analysis. The reliability coefficients for all eight dimensions are greater than 0.8, indicating high reliability. Regarding the 'Alpha if Item Deleted', Deleting any item will not lead to a significant increase in the reliability coefficient, indicating items should not be deleted. Regarding the 'CITC', all items have CITCs greater than 0.4, indicating good relationships between items and suggesting good reliability. [69]

Validity Analysis

Validity is used to measure the degree of data effectiveness, reflecting and examining the fit between the theoretical framework employed and the data collected in this study. It assesses whether the scale adequately represents the research model and the extent of its alignment5). [70] This study employed the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test of sphericity to evaluate the validity of the survey questionnaire. If the KMO value exceeds 0.5 and the significance level of Bartlett's test is less than 0.01, the data is deemed suitable for factor analysis. First, an overall validity test was conducted on the questionnaire scale items. The results indicated that the overall KMO value of the questionnaire was 0.862, with a significance level below 0.01, demonstrating that the questionnaire data was appropriate for factor analysis and exhibited good validity. Subsequently, validity tests were conducted for each dimension as Table 2-4. The results showed that the KMO values for all dimensions exceeded 0.5, and the significance levels were less than 0.01. Therefore, the dimensions are suitable for subsequent data analysis.

Table 3.4: Validity Analysis

Dimension	kmo	p value of Bartlett's Test of Sphericity
B.Romantic Fantasy	0.820	0.000
C.Character Attachment	0.741	0.000
D.Storyline	0.726	0.000
E.Art Style	0.727	0.000
F.Emotional Regulation	0.820	0.000
G.Social Interaction	0.740	0.000
H.Challenge and Achievement	0.659	0.000
I.Personalization	0.822	0.000

Mean Statistics and Analysis

Mean statistics provide an overview of the central tendency of the survey responses, offering insights into participants' general perceptions across different dimensions. The purpose of mean statistics and analysis is to examine the data characteristics of different latent variables, focusing on two key indicators: the mean and standard deviation. This analysis aims to provide a clearer understanding of the Otome game experiences and play motivations of players by analyzing and ranking the dimensions based on their mean scores, as Table 2-5. For the Likert scale scoring system, a higher mean value indicates a stronger tendency toward game experience, and a lower standard deviation suggests greater data stability, with less variability among responses.

Table 3.5: Play Motivation Descriptive Statistic

	Dimension	Sample size	Mean	S.D.	Median
B	Romantic Fantasy	156	3.861	0.814	4
C	Character Attachment	156	3.668	0.854	3.75

Continued on next page

Table 3.5: Play Motivation Descriptive Statistic (Continued)

D	Storyline	156	3.842	0.854	4
E	Art Style	156	4.647	0.601	5
F	Emotional Regulation	156	3.816	0.81	4
G	Social Interaction	156	2.959	1.026	3
H	Challenge and Achievement	156	3.536	0.934	3.667
I	Personalization	156	3.719	0.716	3.727

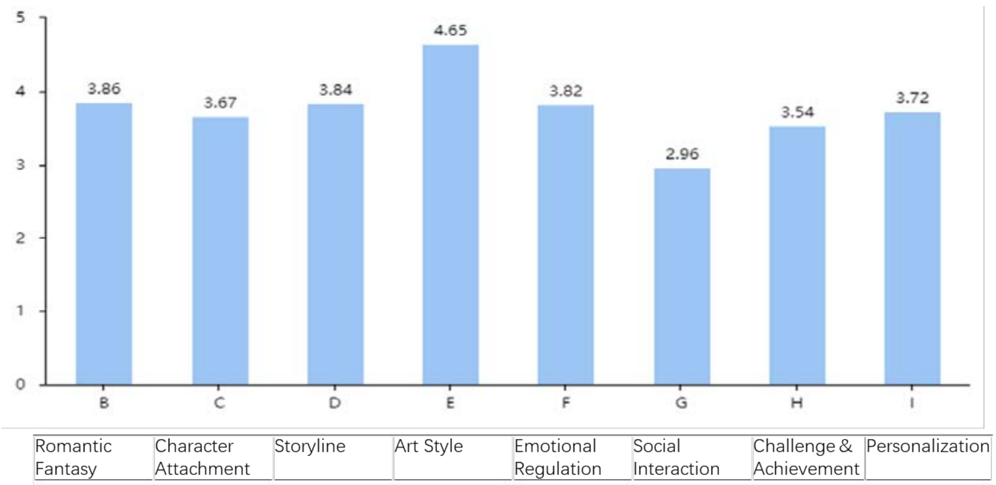

Figure 3.8 Mean Comparison of Play Motivation

Based on the mean scores across dimensions, Art Style received the highest score of 4.65, suggesting that the artistic quality and visual effects of Otome games are critical factors influencing players' engagement. Social Interaction received the lowest score, at 2.96, indicating that most players place relatively little importance on the social functions of Otome games or rarely participate in social discussions. The scores for the remaining six dimensions are relatively close, showing that respondents assign moderate importance to these aspects, including personalization features. Additionally, the small standard deviation for "art style" indicates that players have similar preferences regarding the art style, suggesting that high-quality artistic resources, such as well-designed characters and scenes, are attractive to most players. On the other hand, the large standard deviation for "social interaction" suggests a divergence in players' attitudes towards social engagement. Not all players are enthusiastic about participating in game-related social activities; some prefer to play the game independently.

Descriptive statistic for all the items is listed as Table 2-6.

Table 3.6: Item Descriptive Statistic

Item	n	Mean	S.D.	Median
B1	156	3.923	0.884	4
B2	156	3.699	0.986	4
B3	156	4.006	0.94	4
B4	156	3.814	1.058	4
C1	156	3.571	0.971	4
C2	156	3.423	1.147	4
C3	156	3.865	1.113	4
C4	156	3.814	1.052	4
D1	156	3.846	0.903	4
D2	156	3.654	0.988	4
D3	156	4.026	1.028	4
E1	156	4.641	0.681	5
E2	156	4.667	0.605	5
E3	156	4.635	0.623	5
F1	156	3.744	0.976	4
F2	156	4.038	0.794	4
F3	156	3.603	1.051	4
F4	156	3.878	0.993	4
G1	156	2.859	1.133	3

Continued on next page

Table 3.6: Item Descriptive Statistic (Continued)

G2	156	3.103	1.079	3
G3	156	2.917	1.244	3
H1	156	3.41	1.071	3
H2	156	3.474	1.031	3
H3	156	3.724	1.145	4
I1	156	4.058	0.859	4
I2	156	3.519	1.167	4
I3	156	3.859	0.926	4
I4	156	3.301	1.183	3
I5	156	4.019	0.898	4
I6	156	3.519	1.183	4
I7	156	3.776	1.139	4
I8	156	3.295	1.23	3
I9	156	3.923	0.891	4
I10	156	3.859	0.926	4
I11	156	3.782	0.972	4

Correlation Analysis

Correlation analysis is an essential preliminary step before conducting regression analysis. It is a statistical method used to measure the degree of association between pairs of variables, aiming to assess the strength of relationships among the variables within the model. Correlation analysis determines the degree and direction of relationships based on the size of the correlation coefficient without addressing causal relationships between variables6). Since all the measured items in this study are continuous, Pearson's correlation coefficient is used to examine the correlations among each dimension and demographic characteristics. The Pearson correlation coefficient (r) is a commonly used measure of the strength of the relationship between variables, ranging from -1 to 1. When r is positive, the two variables have a positive correlation; when r is negative, the two variables have a negative correlation. The absolute value of r indicates the strength of the correlation: the larger the absolute value, the stronger the correlation; the smaller the absolute value, the weaker the correlation. [71]

When the absolute value of r is less than 0.40, the variables show a low correlation. If the absolute value of r is between 0.40 and 0.70, the variables exhibit a moderate correlation. An absolute value of r greater than 0.70 indicates a high correlation between the variables. However, the absolute value of the correlation coefficient alone is not sufficient to determine the significance of the correlation. It is also necessary to consider the p-value from the significance test of the correlation coefficient to make valid inferences. The correlation coefficients and Pearson coefficients between the dimensions and demographic characteristics in this study are shown in Table 2-7.

Table 3.7: Pearson Correlation

	A1 Age	A3 Education	A4 Occupation	A5 Frequency	A6 Time Spent	A7 Income	A8 Expenditure
Age	1						
Education	0.526**	1					
Occupation	0.536**	0.232**	1				
Frequency	-0.049	-0.106	-0.007	1			
Time Spent	-0.136	-0.128	-0.053	0.378**	1		
Monthly Income	0.510**	0.420**	0.416**	-0.043	-0.116	1	
Monthly Expenditure	0.005	-0.058	-0.089	0.475**	0.320**	0.136	1
Romantic Fantasy	-0.024	-0.065	-0.104	0.387**	0.245**	-0.047	0.259**

Continued on next page

Table 3.7: Pearson Correlation (Continued)

Character Attachment	-0.104	-0.097	-0.129	0.442**	0.262**	-0.011	0.406**
Storyline	0.08	0.025	-0.012	0.292**	0.246**	0.092	0.256**
Art Style	-0.029	-0.058	-0.134	0.253**	0.173*	0.054	0.233**
Emotional Regulation	-0.069	-0.094	-0.092	0.181*	0.223**	-0.102	0.166*
Social Interaction	-0.086	0.018	-0.047	0.042	0.123	-0.071	0.152
Challenge and Achievement	-0.03	0.081	-0.108	0.028	0.192*	-0.003	0.112
Personalizati on	-0.155	-0.087	-0.217**	0.161*	0.163*	-0.017	0.259**

Table 3.8: Pearson Correlation (continued)

	B	C	D	E	F	G	H	I
	B.Romantic Fantasy, C.Character Attachment, D.Storyline, E.Art Style, F.Emotional Regulation, G.Social Interaction, H.Challenge and Achievement, I.Personalization							
Romantic Fantasy	1							
Character Attachment	0.669**	1						
Storyline	0.577**	0.483**	1					
Art Style	0.383**	0.369**	0.231**	1				
Emotional Regulation	0.680**	0.443**	0.352**	0.301**	1			
Social Interaction	0.277**	0.453**	0.199*	0.157	0.308**	1		

Challenge and Achievement	0.267**	0.304**	0.255**	0.329**	0.284**	0.499**	1	
Personalization	0.412**	0.537**	0.230**	0.399**	0.458**	0.423**	0.417**	1

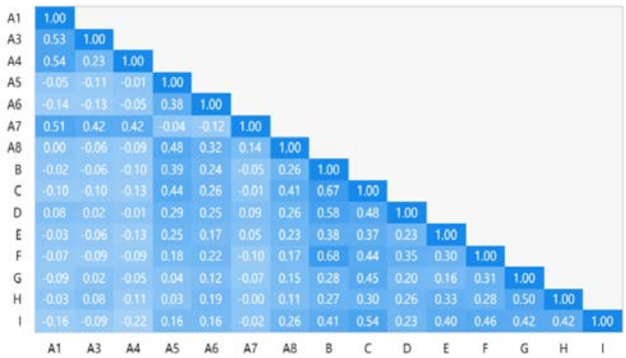

Figure 3.9 Pearson Correlation Visualization

The Pearson correlation matrix reveals that the eight dimensions of gameplay motivation are highly positively correlated, with all but the Art Style-Social Interaction pair showing significant relationships, most of which exhibit strong correlations at the 1% level. This suggests that the experiences players get from otome games in one dimension influence other dimensions and eventually shape their overall preference for the game. The strongest correlations in the matrix are between Romantic Fantasy-Character Attachment (0.67) and Romantic Fantasy-Emotional Regulation (0.68) pairs. This indicates that players who experience intense romantic love feelings within the game are more likely to invest their emotions into game characters and receive greater emotional feedback from the gameplay experience. On the other hand, when examining the relationship between the eight dimensions of play motivation and respondents' basic information, play frequency, playtime, and monthly spending show significant correlations with all dimensions except Social Interaction and Achievement and Challenge. This suggests that when players are immersed and motivated by Otome games, their play frequency, playtime, and spending tend to increase accordingly. These findings align with the expectations of game developers aiming to attract players and encourage playing and in-game purchases through high-quality content.

Regression Analysis for Personalization

One focus of this study is to explore which player characteristics and related dimensions influence players' preferences for personalization content. To achieve this, regression analysis of the factors affecting the personalization dimension was conducted using STATA software. [72] In addition to analyzing the overall personalization dimension, we also examined three components of personalization content: preferences, willingness to pay, and impact on players. The specific composition of these components is shown in Table 2-8. The overall dimension and its three components are used as dependent variables, while players' demographic information and other dimensions are employed as independent variables for multiple regression analysis.

Table 3.9: Components of Personalization Dimension

Component	Item ID	Item question
Personalization -preference	I1	The character personalization features of appearance (such as clothing, hairstyles, makeup, etc.) enhance my gaming experience.
	I3	The personalization features of scene (such as scene backgrounds, background music, items, etc.) enhance my gaming experience.
	I5	Having personalized dialogues and interactions with game characters (such as real-time calls, etc.) enhances my gaming experience.
	I7	The personalized practical services in the game (such as study companionship, fitness companionship, emergency calls, and menstrual tracking, etc.) enhance my daily life experience.

Continued on next page

Table 3.9: Components of Personalization Dimension (Continued)

Personalizat ion -payment	I2	I am willing to pay for character personalization services of appearance (such as clothing, hairstyles, makeup, etc.) in the game.
	I4	I am willing to pay for personalization services of scene (such as scene backgrounds, background music, items, etc.) in the game.
	I6	I am willing to pay for personalized dialogues and interactions services with game characters (such as real-time calls, etc.).
	I8	I am willing to pay for the personalized practical services in the game (such as study companionship, fitness companionship, emergency calls, and menstrual tracking, etc.).
Personalizat ion -influence	I9	The personalization features provided in the game give me sense of joy and satisfaction.
	I10	The personalization features provided in the game give me sense of belonging and immersion.
	I11	The personalization features provided in the game allow me to achieve self-expression.

Whole Dimension Regression Using the Whole dimension of personalization as the dependent variable, the regression result is as Table 2-9.

Table 3.10: Whole dimension regression

Dependent variable: Personalization		
Variables	Coefficient	Standard errors
A1.Age	-0.0653	(0.103)
A3. Highest Level of Education	-0.0231	(0.0630)
A4. Occupation	-0.163	(0.112)
A5. Recent Frequency of Playing Otome Games	-0.0275	(0.0448)

A6. Average Daily Time Spent on Each Otome Game	-0.0155	(0.0488)
A7. Average Monthly Disposable Income	0.0589	(0.0504)
A8. Average Monthly Expenditure on Otome Games	0.0229	(0.0323)
B.Romantic Fantasy	-0.0715	(0.0991)
C.Character Attachment	0.283***	(0.0829)
D.Storyline	-0.0644	(0.0663)
E.Art Style	0.184**	(0.0854)
F.Emotional Regulation	0.230***	(0.0775)
G.Social Interaction	0.0818	(0.0557)
H.Challenge and Achievement	0.123**	(0.0591)
Constant	1.192**	(0.473)
Observations	156	
R-squared	0.464	

Notes: The standard errors are reported in parentheses. *, **, and *** indicate statistical significance at 10%, 5%, and% levels, respectively.

From the regression result, Character Attachment and Emotional Regulation have a significant and strong positive impact on the personalization dimension at the 1% level. This indicates that players who deeply invest their emotions in-game characters and receive greater emotional feedback from the games are more likely to value personalized content within the game.

Art Style and Challenge and Achievement have a significant positive impact on personalization at the 5% level. This suggests that high-quality artistic design and a well-developed achievement and challenge system within the game can encourage players to engage more with personalized content.

Preference for Personalization Using the Preference for personalization as the dependent variable, the regression result is as Table 2-10.

Table 3.11: Preference for personalization regression

Dependent variable: Preference for personalization		
Variables	Coefficient	Standard errors

A1.Age	-0.0407	(0.110)
A3. Highest Level of Education	-0.130*	(0.0672)
A4. Occupation	-0.0773	(0.120)
A5. Recent Frequency of Playing Otome Games	0.0168	(0.0478)
A6. Average Daily Time Spent on Each Otome Game	0.0328	(0.0520)
A7. Average Monthly Disposable Income	0.0286	(0.0538)
A8. Average Monthly Expenditure on Otome Games	0.0169	(0.0345)
B.Romantic Fantasy	0.0132	(0.106)
C.Character Attachment	0.194**	(0.0885)
D.Storyline	-0.0953	(0.0707)
E.Art Style	0.214**	(0.0911)
F.Emotional Regulation	0.266***	(0.0827)
G.Social Interaction	0.0145	(0.0594)
H.Challenge and Achievement	0.119*	(0.0630)
Constant	1.598***	(0.505)
Observations	156	
R-squared	0.439	

Notes: The standard errors are reported in parentheses. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

From the regression result, Emotional Regulation has a significant and strong positive impact on players' preferences for personalization content at the 1% level. This indicates that players who receive more emotional feedback from the game are more inclined to favor personalized features within the game. Character Attachment, Art Style, and Challenge and Achievement have significant positive impacts on personalization at the 5% to 10% level. This suggests that high-quality artistic design, well-developed achievement and challenge systems, and players' deep emotional investment in-game characters enhance their preference for personalized content.

The highest Level of Education has a weak significant negative impact on personalization. Players with higher levels of education show lower preferences for personalization game content. It is hypothesized that as education and life experience increase, the novelty and anticipation for personalization features may diminish.

Willingness to Pay for Personalization Using the Willingness to pay for personalization as the dependent variable, the regression result is as Table 2-11.

Table 3.12: Willingness to pay for personalization regression

Dependent variable: Willingness to pay for personalization		
Variables	Coefficient	Standard errors
A1.Age	-0.0315	(0.153)
A3. Highest Level of Education	0.0563	(0.0935)
A4. Occupation	-0.137	(0.167)
A5. Recent Frequency of Playing Otome Games	-0.113*	(0.0664)
A6. Average Daily Time Spent on Each Otome Game	-0.01	(0.0724)
A7. Average Monthly Disposable Income	0.0857	(0.0748)
A8. Average Monthly Expenditure on Otome Games	0.101**	(0.0479)
B.Romantic Fantasy	-0.256*	(0.147)
C.Character Attachment	0.456***	(0.123)
D.Storyline	-0.0113	(0.0984)
E.Art Style	0.254**	(0.127)
F.Emotional Regulation	0.203*	(0.115)
G.Social Interaction	0.147*	(0.0826)
H.Challenge and Achievement	0.114	(0.0877)
Constant	-0.165	(0.702)
Observations	156	
R-squared	0.393	

Notes: The standard errors are reported in parentheses. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

From the regression result, Character Attachment has a significant and strong positive impact on players' willingness to pay for personalized content at the 1 p level. This indicates that players who deeply invest their emotions in-game characters are more likely to spend money on personalized features within the game. Monthly Expenditure and Art Style have significant positive impacts at the 5% level. Players who used to spend more on otome games are also more willing to pay for personalized content. Additionally, high-quality artistic design in the game significantly enhances players' willingness to pay for personalization features. Emotional Regulation and Social Interaction have significant positive impacts at the 10% level. Players who receive greater emotional feedback from the game and who actively participate in social activities related to Otome games are more inclined to pay for personalized content. Game Frequency shows a weakly significant negative impact. This may be because high-frequency players, who are already accustomed to the game's existing content, exhibit lower interest in personalized features.

Influence of personalization Using the Influence of personalization as the dependent variable, the regression result is as Table 2-12.

Table 3.13: Influence of personalization regression

Dependent variable: Influence of personalization		
Variables	Coefficient	Standard errors
A1.Age	-0.152	(0.147)
A3. Highest Level of Education	0.00436	(0.0899)
A4. Occupation	-0.296*	(0.160)
A5. Recent Frequency of Playing Otome Games	0.0422	(0.0639)
A6. Average Daily Time Spent on Each Otome Game	-0.102	(0.0696)
A7. Average Monthly Disposable Income	0.0748	(0.0719)
A8. Average Monthly Expenditure on Otome Games	-0.0689	(0.0461)
B.Romantic Fantasy	0.0839	(0.141)
C.Character Attachment	0.135	(0.118)
D.Storyline	-0.103	(0.0945)
E.Art Style	0.0651	(0.122)
F.Emotional Regulation	0.213*	(0.110)
G.Social Interaction	0.106	(0.0794)
H.Challenge and Achievement	0.132	(0.0842)

Constant	2.464***	(0.675)
Observations	156	
R-squared	0.267	

Notes: The standard errors are reported in parentheses. *, **, and *** indicate statistical significance at 10%, 5%, and 1% levels, respectively.

From the regression result, Players' occupations have a weakly significant impact on their perception of personalization. Students tend to get greater satisfaction, immersion, and self-expression from personalization content, while those who have graduated exhibit relatively less sensitivity to such features. Emotional Regulation has a weakly significant positive impact on the perception of personalization. This indicates that players who generally receive greater emotional feedback from the game are also more likely to experience enhanced satisfaction, immersion, and self-expression from personalization content. Based on the statistical analysis, it can be found that positive correlations were observed among the motivational dimensions, with Romantic Fantasy showing strong links to Character Attachment (r=0.67) and Emotional Regulation (r=0.68). This indicates that emotional and narrative elements significantly shape player preferences. Demographic factors such as play frequency and expenditure correlated positively with motivational dimensions, except for Social Interaction and Challenge/Achievement. And the regression analysis revealed that personalization preferences were positively influenced by Character Attachment, Emotional Regulation, Art Style, and Challenge/Achievement. Players with strong emotional connections to characters and immersive experiences showed a higher appreciation for personalization features. This analysis provides a comprehensive understanding of player behavior, highlighting the importance of emotional and artistic elements in otome game design while identifying opportunities for enhancing personalization to deepen engagement.

3.3. Personalized Dialogue System

3.3.1 Concept

Here is the figure 3.10 of the concept of the Personalized dialogue:

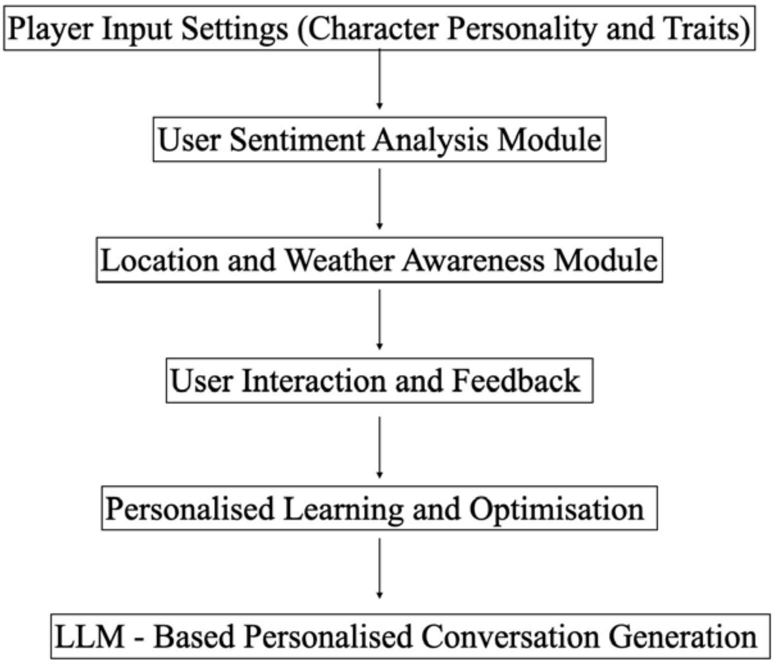

Figure 3.10 Concept of Personalized Dialogue System

3.3.2 Personalized Dialogue System based on LLMs Agent

The questionnaire data highlights the significance of personalized interaction features in enhancing players' gaming experience. We have designed a multi-agent framework based on LLMs to improve emotional engagement and personalization in dialogue generation for otome games. This framework consists of three auxiliary analytical agents, a primary agent for generating personalized content, and a Memory Block. By incrementally analyzing information provided by the auxiliary agents, the primary agent progressively understands the emotions, context, and location inferred from the user's input. It then combines this understanding with the desired character persona and historical information stored in the memory module to generate responses that offer deeper emotional and relational support. The detailed framework is as follows:

- **Personalized Character Agent:** This agent analyzes the latent emotions inherent in the character's personality, denoted as C_e.
- **Personalized Sentiment Analysis Agent:** This agent evaluates the emotions embedded in the user's input, denoted as U_e.
- **Personalized Location and Weather Agent:** This agent identifies potential user location details and related weather information, denoted as U_l, w.
- **Memory Block:** The Memory Block serves as a repository for dialogue history throughout the dialogue process, donated as M.
- **Personalized Dialogue Agent:** This Agent generates personalized responses Y enriched with emotional value and relational support. It integrates insights from the three auxiliary agents and chat memory while adhering to the character persona defined by the user, ensuring emotionally resonant and contextually appropriate replies.

Following the equation 3.1 shows, by integrating emotional analysis, contextual information such as location and weather, and memory-based understanding, Personalized Dialogue Agent will generate dialogue responses Y that align with the predefined character persona with enhanced emotional support and personalized interaction. x is the user's dialog input for this round.

$$Y = \text{Dialogue\_Agent}(x \mid C_e, U_e, U_{l,w}, M) \tag{3.1}$$

Personalized Character Agent

The Personalized Character Agent analyzes the personality traits set by the user to uncover potential emotions, laying the foundation for personalized dialogue generation. Specifically, this Agent maps the user's preset Big Five personality model [73] (OCEAN, encompassing Openness (O), Conscientiousness (C), Extraversion (E), Agreeableness (A), and Neuroticism (N)) onto the PAD emotional model proposed by Russell et al. [74, 75], which includes Pleasure (P), Arousal (A), and Dominance (D) dimensions. This mapping is based on the quantitative relationships between the Five-Factor personality traits and PAD emotions as proposed by Mehrabian et al [76, 77]. Through this equation 3.2, the scores of the character on the Big Five personality dimensions (O, C, E, A, N $\in$ [-1,1]) are transformed into the corresponding emotional vector V_e = (P, A, D), which represents the implicit emotional state of the character.

$$\begin{aligned} P &= 0.21E + 0.59A + 0.19N, \\ A &= 0.15O + 0.30A - 0.57N, \\ D &= 0.25O + 0.17C + 0.60E - 0.32A \end{aligned} \tag{3.2}$$

Subsequently, by analyzing the positivity or negativity of the V_e's PAD tridimensionality, potential emotions from a preset emotional library 3.14 can be selected for the character based on research by Navarretta [1], providing a clear benchmark for generating emotionally nuanced dialogue and ensuring that responses align with the character's traits and the user's desired emotional expression. Based on the above process, the Personalized Character Agent will output the implicit emotion of the preset character C_e.

Table 3.14 List of emotions and their PAD values [1]

P	**A**	**D**	**Emotions**
+	+	+	Amused, Excited, Happy, Interested, Ironic, Joking, Proud, Satisfied, Self-Confident, Supportive
−	−	−	Disappointed, Hesitant, Unconfident, Uncomfortable, Uninterested
+	−	−	Certain, Friendly
−	+	−	Awkward, Embarrassed, Puzzled, Shy, Uncertain, Uneasy
+	+	−	Engaged, Surprised
+	−	−	Docile, Thoughtful
−	+	−	Irritated

Personalized Sentimental analysis Agent

Firstly, construct a prompt $Prompt\_pad$ based on the descriptions of the Pleasure (P), Arousal (A), and Dominance (D) dimensions of the PAD emotional model proposed by research [1, 74, 75]. Subsequently, as illustrated in the equation 3.3, the Personalized Sentimental analysis Agent analyzes the emotions implied in the user's input statement based on this prompt, thereby obtaining the potential emotional information U_e within the user's input statement x.

$$U_e = \text{SentimentalAnalysis\_Agent}(x \mid \text{Prompt\_pad}) \tag{3.3}$$

Personalized Location and Weather Agent

The Personalized Location and Weather Agent initially analyzes the user's input x to determine whether it contains location information. If a location mention is detected, the agent will invoke the Google Maps API and the OpenWeather API to retrieve detailed information about the user's specified location and the local weather conditions. Specifically, the Google Maps API is utilized to parse the location mentioned by the user and obtain nearby map-related information; the OpenWeather API, based on these coordinates, retrieves real-time weather data for the corresponding location, including temperature, humidity, wind speed, and weather condition descriptions. Subsequently, the agent integrates the aforementioned geographical and meteorological information to generate a concise and clear description U_l, w 3.4. This description serves as part of the contextual information and is used in the subsequent dialogue generation process.

$$U_{l,w} = \text{LocationWeather\_Agent}(x \mid \text{API\_map}, \text{API\_weather}) \tag{3.4}$$

Memory Block

Memory Block is used to store all historical information during a dialog M, providing the basis for contextual understanding and generation of subsequent dialogs.

Personalized Dialogue Agent

The Personalized Dialogue Agent, serving as the primary dialogue agent, synthesizes information from multiple auxiliary agents, including those for character latent emotions C_e, user sentiment analysis U_e, and location and weather analysis U_l, w, to generate emotionally resonant and user-tailored dialogue responses Y. Leveraging pre-defined character personality and real-time user emotional states, in conjunction with historical dialogue data M, the agent generates contextually relevant responses Y that prioritize emotional consistency, thematic coherence, and the provision of emotional support and relational intimacy. Ultimately, the output manifests as personalized dialogue exchanges. Through this multi-dimensional information fusion process, the agent endeavors to enhance user emotional engagement and interactive immersion within the otome game experience.

Chapter 4
Experimentation on Interactive system

4.1. Experiment

4.1.1 Objective

The primary aim is to evaluate whether the personalized dialogue system outperforms existing dialogue systems in popular otome games in enhancing player immersion and optimizing personalized gaming experiences based on real player feedback. Additionally, the study seeks to gather firsthand usability data from players' experiences with this system, obtaining suggestions and perspectives on various components of the personalized dialogue system. This information will support future improvements and iterations.

4.1.2 Methodology

A mixed-methods approach combining quantitative and qualitative research methods was adopted. Eleven players, all with prior experience and familiarity with otome games, were selected to participate in gameplay tests, questionnaire surveys, and in-depth playtesting followed by interviews.

4.1.3 Procedure

The testing process involved gameplay sessions on an iPad with the dialogue interaction systems of two popular otome games: Game A–*Light and Night*, Game B–*Love and Deepspace*, and Game C–the personalized dialogue system. During the tests, the basic operation flow and interaction duration with each dialogue system were recorded. Interaction with each game was limited to 5 minutes, for a total playtest time of 15 minutes. Based on the dialogue topics of the existing games, the 'eating' topic was selected for interaction to facilitate comparison of the experiments. After the gameplay session, participants completed a questionnaire and participated in a brief interview, with the post-play phase lasting approximately 30 minutes.

During the test, we closely monitored the basic operation flow, as well as the details and frequency of interactions within the gameplay experience. Particular attention was given to usability issues within the interaction systems, including Response Issues, Instances where a click failed to trigger any response; Latency Issues: Cases where a response occurred with noticeable delay after a click.

4.1.4 Content

Dialogue System of Game A

Game A incorporates a multimodal dialogue system that integrates images, audio, and text. The dialogue takes the form of video conversations between the player and in-game characters.

There are five characters in the game for the player to choose from to interact with.

Players are presented with dialogue options such as "Have you decided what to eat?" and "Recommend something for me." If the player selects "Recommend something for me," they have the opportunity to request an alternative by choosing "Show me something else" if they are dissatisfied with the suggested food. Alternatively, selecting options like "I'll go with your choice" or "I've already decided" will lead directly to a scenario where the character and the player agree to have a meal together, concluding the dialogue interaction. Below is a selection of dialogue interactions recorded from Game A:
Player: "Join me for a meal."
Character: "What are you in the mood for? The cafeteria, takeout, or cooking at home? If you're not sure what to eat, want me to give you some ideas?"
Player: "Recommend something."
Character: "How about some braised fish? You can compare the taste between

(Source: Otome game: Light and Night

Figure 4.1 Characters in Game A

(Source: Otome game: Light and Night

Figure 4.2 Dialogue system in Game A(1)

(Source: Otome game: Light and Night

Figure 4.3 Dialogue system in Game A(2)

store-bought and my homemade version."
Player: "Not feeling it. Got anything else?"
Character: "Didn't you say you were craving fried chicken the other day? Let's go with that today."
Player: "Hmm, not quite. Any other ideas?"
Character: "If you want something quick and easy, how about fried rice? It's simple, and as long as the seasoning's right, it's hard to mess up." Player: "Alright, I'll go with that."
Character: "Perfect. I've already ordered my takeout, too. Let's wait for our food to arrive and eat together."
Player: "The food's here. Let's dig in."
Character: "Ready to eat? Mine just got here, too. I'll enjoy your 'mukbang' while I eat."
Player: "All done."
Character: "You little foodie, you've managed to get food on your face. Here, let me wipe it off for you."

Dialogue System of Game B

Game B incorporates a multimodal dialogue system that integrates images and text. The dialogue takes the form of face-to-face conversations between the player and in-game characters.

There are four characters in the game for the player to choose from to interact with.

Below is a selection of dialogue interactions recorded from Game B:
Player: "What should I eat?" Character: "At this hour, it's not a meal, is it?" Player: "What are you having? Give me some ideas." Character: "I usually go for something sweet in the afternoon. Lately, I've been ordering fruit tea a lot." Player: "Not in the mood. Got anything else?" Character: "How about mango sago? A little sugar can help lift your spirits." Player: "Sounds good. I'll go with that." Character: "Great choice! After you eat, why not stretch your legs a bit? We could go for a walk together."

(Source: Otome game: Love and Deepspace

Figure 4.4 Characters in Game A

(Source: Otome game: Love and Deepspace

Figure 4.5 Dialogue System in Game B(1)

(Source: Otome game: Love and Deepspace

Figure 4.6 Dialogue System in Game B(2)

Personalized Dialogue System

This personalized dialogue system incorporates a single modal dialogue system that only includes text. Before the conversation begins, three characters with distinct personality traits are pre-set in this personalized dialogue system: Apple, Pear, and Grape.

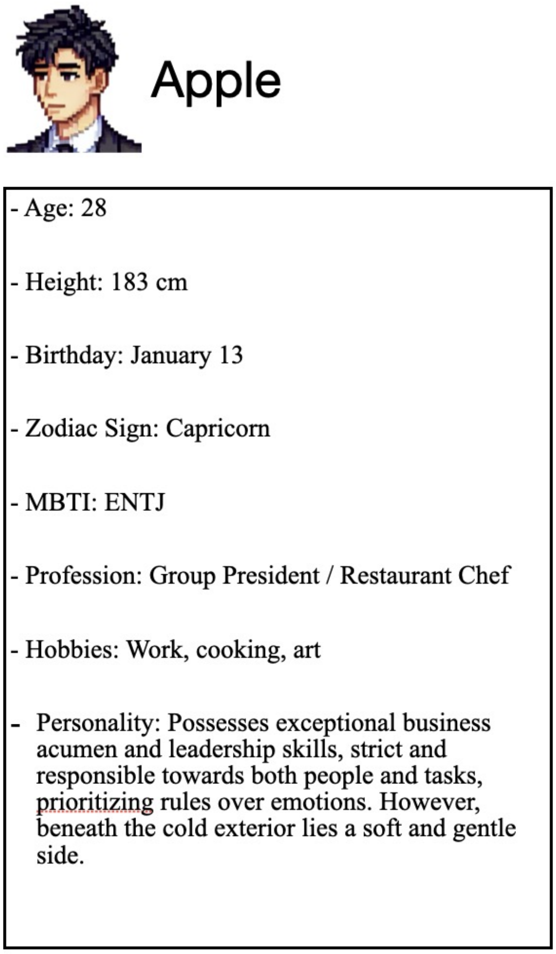

Figure 4.7 Character: Apple

Apple's Big Five personality traits are as follows:

Pear

- Age: 27
- Height: 186 cm
- Birthday: September 5
- Zodiac Sign: Virgo
- MBTI: INTJ
- Profession: Cardiothoracic Surgeon
- Traits: Dislikes carrots, doesn't drink alcohol, loves sweets even with toothaches, has a background as a military doctor
- Skills: Proficient in skiing, excellent at billiards, can peel an apple without breaking the peel
- Personality: Precise, restrained, rational, steady, introverted, workaholic, planner, detail-oriented, tells dry jokes, exhibits contrasting charm

Figure 4.8 Character: Pear

Grape

- Age: 24
- Height: 183 cm
- Birthday: March 6
- Zodiac Sign: Pisces
- MBTI: INFP
- Profession: Painter
- Traits: Likes seafood, enjoys bathing, afraid of cats, has a fear of heights, gets seasick, and cannot ride a bicycle
- Skills: Plays the harmonica, speaks multiple languages
- Personality: Romantic, sensitive, clingy, innocent, lively, and enjoys acting spoiled

Figure 4.9 Character: Grape

- Extraversion: Moderate to low
- Agreeableness: Moderate
- Openness: Moderate to high
- Conscientiousness: Very hight
- Emotional Stability: Moderate to high

Pear's Big Five personality traits are as follows:

- Extraversion: Moderate to low
- Agreeableness: Moderate to high
- Openness: Moderate to high
- Conscientiousness: Very high
- Emotional Stability: Moderate

Grape's Big Five personality traits are as follows:

- Extraversion: High
- Agreeableness: Moderate to high
- Openness: Very high
- Conscientiousness: Moderate
- Emotional Stability: Moderate to low

Below is a selection of dialogue interactions recorded from the personalized dialogue system:
Player: "Good afternoon."
Character: "Good afternoon, my princess! Is there something that's making you unhappy?"

Player: "Not at all! The sunlight in Tokyo is lovely today, and I feel great. "
Character: "The sunlight in Tokyo sounds wonderful! I'm happy too and would love to take a walk with you under the sun, soaking in Tokyo's vibrant energy."
Player: "How about we go for a walk this afternoon?"
Character: "Sure! Where shall we go for a walk, my cutie?"
Player: "Let's decide what to eat for dinner first, then take a walk around the area near the restaurant."
Character: "Hmm… I'd love to have your favorite seafood! Then we can take a leisurely walk. How does that sound?"
Player: "Sounds great! Are there any good seafood restaurants in Tokyo?"
Character: "I know a fantastic sushi restaurant! I'll take you there tonight. It's by the seaside, and the view is absolutely stunning."

4.2. Experiment Results

4.2.1 Anova Analysis

Analysis of Variance (ANOVA) is a statistical technique used to examine whether there are statistically significant differences between the means of three or more independent groups, by analyzing the variance within groups and between groups1). In this study, Anova Analysis will be used to compare the three experiment groups simultaneously to determine if at least one experiment group differs significantly from the others.

Interaction Time and Interaction Frequency Analysis

The experiments focused on the interaction time and interaction frequency of the eleven participants across three games, serving as indicators of player engagement and preference for the games. The ANOVA analysis of the playing data from the three games was conducted using the SPSSAU platform. The results are presented in Table 1-1.

Table 4.1: One-Way ANOVA Analysis for Interaction

Item	Option	n	Mean	S.D.	F	p Value

Time	A	11	53	14.91	415.719	0.000**
	B	11	49	9.49		
	C	11	275.27	31.92		
	Total	33	125.76	109.3		
Frequency	A	11	5	1	32.5	0.000**
	B	11	4.55	0.82		
	C	11	10.27	2.94		
	Total	33	6.61	3.19		

Notes: * and ** indicate statistical significance at 5% and 1% levels, respectively.

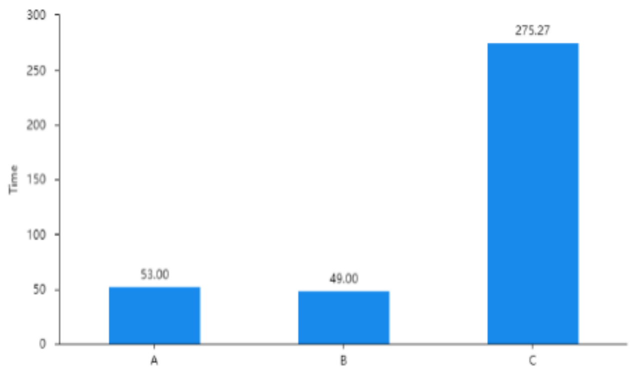

Figure 4.10 Mean of Interaction time

As a result, interaction time shows significance at the 0.01 level (F = 415.719, p = 0.000), indicating strong and significant interaction time differences among the three games, and the specific comparison of differences indicates that the comparison with significant differences is C ¿ A; C ¿ B. It can be inferred that the participants played more time on Game C compared with Game A and Game B. Similarly, interaction frequency shows significance at the 0.01 level (F = 32.500, p = 0.000), indicating strong and significant interaction frequency differences among the three games, and the specific comparison of differences indicates that the comparison with significant differences is C ¿ A; C ¿ B. It can be inferred that

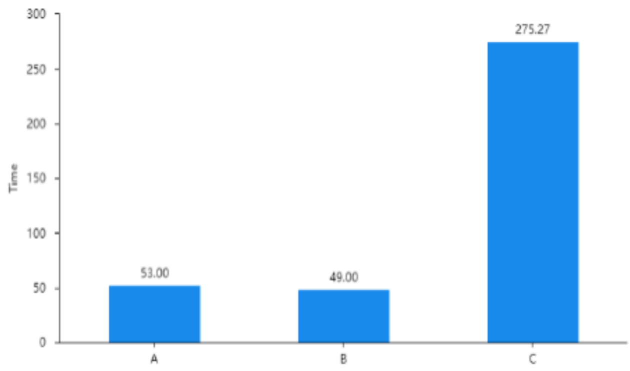

Figure 4.11 Mean of Interaction time

the participants had more interactions in Game C compared with Game A and Game B.

4.2.2 Questionnaire Analysis

As part of the experiment, a questionnaire was administered after each gameplay session. An ANOVA analysis was conducted on the questionnaire results from the three experiment groups to examine whether significant differences existed between groups. Questions 6, 8, 9, 10, and 18 were reverse-coded and have been standardized for scoring. The analysis results are presented in Table 1-2.

Table 4.2: One-Way ANOVA Analysis for Questionnaire

	Game (Mean ± S.D.)			F	p Value
	A (n = 11)	B (n = 11)	C (n = 11)		
Q1	3.09 ± 0.94	3.18 ± 1.17	4.36 ± 0.50	6.63	0.004**
Q2	2.91 ± 0.94	3.00 ± 0.89	4.27 ± 0.65	9.095	0.001**
Q3	2.27 ± 1.10	2.55 ± 1.13	3.91 ± 0.94	7.5	0.002**
Q4	2.73 ± 0.90	2.45 ± 1.37	4.09 ± 0.83	7.5	0.002**
Q5	2.55 ± 1.21	2.27 ± 1.10	4.00 ± 0.77	8.646	0.001**
Q6	2.73 ± 1.27	2.09 ± 0.83	3.36 ± 0.81	4.509	0.019*
Q7	2.73 ± 1.49	2.55 ± 1.04	2.64 ± 1.21	0.057	0.944
Q8	2.55 ± 1.13	2.36 ± 0.92	3.64 ± 0.81	5.621	0.008**

Q9	2.64 ± 1.21	2.64 ± 0.81	3.82 ± 0.75	5.748	0.008**
Q10	3.55 ± 1.57	3.27 ± 1.35	4.09 ± 1.14	1.026	0.371
Q11	2.82 ± 0.75	3.27 ± 1.01	4.00 ± 0.77	5.375	0.010*
Q12	2.00 ± 1.00	2.09 ± 0.94	3.36 ± 1.03	6.512	0.004**
Q13	3.18 ± 1.17	2.91 ± 1.14	4.27 ± 0.65	5.592	0.009**
Q14	2.73 ± 1.10	2.45 ± 1.04	3.82 ± 0.87	5.625	0.008**
Q15	2.18 ± 1.17	2.18 ± 0.87	3.55 ± 1.04	6.392	0.005**
Q16	3.09 ± 1.22	3.27 ± 1.10	4.55 ± 0.69	6.514	0.004**
Q17	1.45 ± 0.69	1.45 ± 0.93	2.82 ± 1.33	6.579	0.004**
Q18	3.82 ± 1.25	3.45 ± 1.21	4.18 ± 1.08	1.039	0.366
Q19	2.82 ± 0.98	2.91 ± 0.94	3.91 ± 0.70	5.155	0.012*
Q20	4.73 ± 0.65	4.45 ± 0.82	3.73 ± 0.90	4.619	0.018*
Q21	2.82 ± 1.47	3.00 ± 1.10	3.36 ± 0.67	0.667	0.521
Q22	3.45 ± 1.04	3.55 ± 1.13	3.82 ± 0.87	0.38	0.687
Q23	2.55 ± 1.21	2.73 ± 1.10	3.00 ± 1.18	0.422	0.659
Q24	3.45 ± 0.93	2.64 ± 1.03	4.00 ± 1.10	4.971	0.014*
Q25	2.45 ± 1.37	2.36 ± 1.36	3.55 ± 1.29	2.643	0.088
Q26	2.36 ± 1.03	2.64 ± 1.03	4.27 ± 0.65	13.921	0.000**
Q27	2.64 ± 1.21	2.64 ± 0.81	4.27 ± 0.79	10.8	0.000**
Q28	2.73 ± 1.10	3.18 ± 0.98	4.27 ± 1.10	6.123	0.006**

Notes: * and ** indicate statistical significance at 5% and 1% levels, respectively.

The analysis results indicate that most questions showed significant differences, with Game C outperforming the other two games. Among the 28 questions, 20 exhibited significant differences where Game C had the highest mean score, 7 showed no significant group differences, and 1 question showed a significant difference where Game C had the lowest mean score. These findings suggest that the framework designed for Game C in this study is significantly better than the other two available games in terms of gameplay and interaction, the sense of immersion and motivation it provides to players, and its ability to establish emotional connections with users. This validates the effectiveness of the game design proposed in this study. Question 20 revealed a significant difference with Game C having the lowest mean score, indicating a substantial experiential gap between unimodal and multimodal interactions. This highlights the need for future design optimizations to improve the functions of visual and auditory elements in enhancing text-based interactions while refining algorithmic improvements.

Chapter 5
Conclusion

5.1. Results

This study highlights the pivotal role of personalized dialogue systems in enhancing player engagement, emotional resonance, and retention within mobile otome games. Through the integration of advanced technologies, particularly Large Language Models, the personalized interaction framework demonstrated superior adaptability and interactivity. The experimental results indicate that personalized character agents, sentimental analysis mechanisms, and contextual features such as location and weather-based adaptations significantly elevate the gaming experience, fostering deeper connections between players and virtual characters.

5.2. Limitation

Firstly, the current personalized dialogue system is constrained to unimodal interactions, focusing primarily on text-based dialogues. This limitation restricts the exploration of multimodal engagement, such as integrating voice, visual cues, or haptic feedback, which could offer a more immersive user experience. Secondly, the dataset and experimental scope were limited to a specific demographic group, primarily targeting female otome game players. This may affect the generalizability of the findings to broader gaming audiences or different cultural contexts. Additionally, the reliance on pre-trained models could introduce biases inherent in the original datasets, potentially impacting the authenticity of personalized interactions.

5.3. Future Works

Future research will be aim to address these limitations by exploring multimodal interaction systems. Incorporating auditory and visual elements, alongside haptic feedback, could significantly enhance the immersive quality of personalized dialogues. Moreover, expanding the dataset to include diverse player demographics and cross-cultural contexts will improve the robustness and applicability of the findings.

Another promising direction is the refinement of dynamic personalization algorithms. By integrating real-time user feedback and adaptive learning mechanisms, future systems can evolve based on individual player behaviors, preferences, and emotional responses. Additionally, the development of hybrid models that combine rule-based and machine learning approaches may help mitigate the biases of pre-trained LLMs, ensuring more authentic and contextually relevant interactions.

The results of this study contribute to future works by establishing a foundational framework for personalized dialogue systems in gaming environments. This framework can serve as a reference for designing interactive systems beyond otome games, including educational applications, virtual assistants, and therapeutic tools. The insights gained from the experimental analysis provide critical benchmarks for evaluating user engagement metrics, emotional connectivity, and the effectiveness of AI-driven personalization strategies in interactive media.

References

[1] Costanza Navarretta. Speech, emotions and facial expressions in dyadic spontaneous conversations. *Intelligent Decision Technologies*, 8:255–263, 2014. doi:10.3233/IDT-140194.

[2] Kun Huang. Online mobile game design based on the emotional characteristics of female players, 2015. master's thesis.

[3] Yiqun Liu. Research on the influence factors of user's contiuance intention of domestic "maiden" mobile games, 2022. master's thesis.

[4] Jialin Wang. Research on the motivation and behaviour of "female-oriented" game players from the perspective of using and satisfying–ashes of the kingdom as an example, 2024. master's thesis.

[5] Miu Shenguang. Study on the influencing factors of consumption intention of female-oriented games–taking b-girl games as an example. Master's thesis, Nanjing University of the Arts, 2023. URL: https://kns.cnki.net/kcms/detail/detail.aspx?dbcode=CMFD&filename=1024381692.nh.

[6] Du Ruiying. Design optimization research of otome mobile games based on the influencing factors of user satisfaction. 2022. master's thesis.

[7] Ching-I Teng. Customization, immersion satisfaction, and online gamer loyalty. *Computers in Human Behavior*, 26(6):1547–1554, 2010.

[8] Emily Brown and Paul Cairns. A grounded investigation of game immersion. In *CHI'04 extended abstracts on Human factors in computing systems*, pages 1297–1300. ACM, 2004.

[9] Mihaly Csikszentmihalyi. *Flow: The psychology of optimal experience.* Harper and Row, 1990.

[10] Laura Ermi and Frans Mäyrä. Fundamental components of the gameplay experience: Analysing immersion. In *DiGRA Conference Proceedings*, 2005.

[11] Kenneth Mark Colby, Sylvia Weber, and Franklin Dennis Hilf. Artificial paranoia. *Artif. Intell.*, 2(1):1–25, January 1971. URL: `https://doi.org/10.1016/0004-3702(71)90002-6`, `doi:10.1016/0004-3702(71)90002-6`.

[12] Terry Winograd. Understanding natural language. *Cognitive Psychology*, 3(1):1–191, 1972. URL: `https://www.sciencedirect.com/science/article/pii/0010028572900023`, `doi:https://doi.org/10.1016/0010-0285(72)90002-3`.

[13] W. A. Woods. Progress in natural language understanding: an application to lunar geology. In *Proceedings of the June 4-8, 1973, National Computer Conference and Exposition*, AFIPS '73, page 441–450, New York, NY, USA, 1973. Association for Computing Machinery. URL: `https://doi.org/10.1145/1499586.1499695`, `doi:10.1145/1499586.1499695`.

[14] Wendy G. Lehnert, Michael G. Dyer, Peter N. Johnson, C.J. Yang, and Steve Harley. Boris—an experiment in in-depth understanding of narratives. *Artificial Intelligence*, 20(1):15–62, 1983. URL: `https://www.sciencedirect.com/science/article/pii/0004370283900140`, `doi:https://doi.org/10.1016/0004-3702(83)90014-0`.

[15] Bayan Shawar and Eric Atwell. A chatbot system as a tool to animate a corpus. *ICAME Journal: International Computer Archive of Modern and Medieval English Journal*, 29:5–24, 10 2005.

[16] Jiao Liu, Yanling Li, and Min Lin. Review of intent detection methods in the human-machine dialogue system. *Journal of Physics: Conference Series*, 1267:012059, 07 2019. `doi:10.1088/1742-6596/1267/1/012059`.

[17] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27. Curran Associates,

Inc., 2014. URL: https://proceedings.neurips.cc/paper_files/paper/2014/file/a14ac55a4f27472c5d894ec1c3c743d2-Paper.pdf.

[18] J. L. Elman. Finding structure in time. *Cognitive Science*, 14:179–211, 1990.

[19] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[20] Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using RNN encoder–decoder for statistical machine translation. In Alessandro Moschitti, Bo Pang, and Walter Daelemans, editors, *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734, Doha, Qatar, October 2014. Association for Computational Linguistics. URL: https://aclanthology.org/D14-1179, doi:10.3115/v1/D14-1179.

[21] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Ł ukasz Kaiser, and Illia Polosukhin. Attention is all you need. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

[22] Tsung-Hsien Wen, Milica Gašić, Dongho Kim, Nikola Mrkšić, Pei-Hao Su, David Vandyke, and Steve Young. Stochastic language generation in dialogue using recurrent neural networks with convolutional sentence reranking. In Alexander Koller, Gabriel Skantze, Filip Jurcicek, Masahiro Araki, and Carolyn Penstein Rose, editors, *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 275–284, Prague, Czech Republic, September 2015. Association for Computational Linguistics. URL: https://aclanthology.org/W15-4639, doi:10.18653/v1/W15-4639.

[23] Tsung-Hsien Wen, Milica Gašić, Nikola Mrkšić, Pei-Hao Su, David Vandyke, and Steve Young. Semantically conditioned LSTM-based natural language

generation for spoken dialogue systems. In Lluís Màrquez, Chris Callison-Burch, and Jian Su, editors, *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1711–1721, Lisbon, Portugal, September 2015. Association for Computational Linguistics. URL: `https://aclanthology.org/D15-1199`, `doi:10.18653/v1/D15-1199`.

[24] Tsung-Hsien Wen, Milica Gašić, Nikola Mrkšić, Lina M. Rojas-Barahona, Pei-Hao Su, David Vandyke, and Steve Young. Multi-domain neural network language generation for spoken dialogue systems. In Kevin Knight, Ani Nenkova, and Owen Rambow, editors, *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 120–129, San Diego, California, June 2016. Association for Computational Linguistics. URL: `https://aclanthology.org/N16-1015`, `doi:10.18653/v1/N16-1015`.

[25] Hao Zhou, Minlie Huang, and Xiaoyan Zhu. Context-aware natural language generation for spoken dialogue systems. In Yuji Matsumoto and Rashmi Prasad, editors, *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 2032–2041, Osaka, Japan, December 2016. The COLING 2016 Organizing Committee. URL: `https://aclanthology.org/C16-1191`.

[26] Ondřej Dušek and Filip Jurčíček. Sequence-to-sequence generation for spoken dialogue via deep syntax trees and strings. In Katrin Erk and Noah A. Smith, editors, *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 45–51, Berlin, Germany, August 2016. Association for Computational Linguistics. URL: `https://aclanthology.org/P16-2008`, `doi:10.18653/v1/P16-2008`.

[27] Ryan Lowe, Nissan Pow, Iulian Serban, and Joelle Pineau. The Ubuntu dialogue corpus: A large dataset for research in unstructured multi-turn dialogue systems. In Alexander Koller, Gabriel Skantze, Filip Jurcicek, Masahiro Araki, and Carolyn Penstein Rose, editors, *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 285–294, Prague, Czech Republic, September 2015. Association for

Computational Linguistics. URL: https://aclanthology.org/W15-4640, doi:10.18653/v1/W15-4640.

[28] Janarthanan Rajendran, Jatin Ganhotra, and Lazaros C. Polymenakos. Learning end-to-end goal-oriented dialog with maximal user task success and minimal human agent use. *Transactions of the Association for Computational Linguistics*, 7:375–386, 2019. URL: https://aclanthology.org/Q19-1024, doi:10.1162/tacl_a_00274.

[29] Oriol Vinyals and Quoc V. Le. A neural conversational model. *CoRR*, abs/1506.05869, 2015. URL: http://dblp.uni-trier.de/db/journals/corr/corr1506.html#VinyalsL15.

[30] Baotian Hu, Zhengdong Lu, Hang Li, and Qingcai Chen. Convolutional neural network architectures for matching natural language sentences. NIPS'14, page 2042–2050, Cambridge, MA, USA, 2014. MIT Press.

[31] Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, Yifan Du, Chen Yang, Yushuo Chen, Zhipeng Chen, Jinhao Jiang, Ruiyang Ren, Yifan Li, Xinyu Tang, Zikang Liu, Peiyu Liu, Jian-Yun Nie, and Ji-Rong Wen. A survey of large language models, 2024. URL: https://arxiv.org/abs/2303.18223, arXiv:2303.18223.

[32] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. Llama: Open and efficient foundation language models, 2023. URL: https://arxiv.org/abs/2302.13971, arXiv:2302.13971.

[33] Jinze Bai, Shuai Bai, Yunfei Chu, Zeyu Cui, Kai Dang, Xiaodong Deng, Yang Fan, Wenbin Ge, Yu Han, Fei Huang, Binyuan Hui, Luo Ji, Mei Li, Junyang Lin, Runji Lin, Dayiheng Liu, Gao Liu, Chengqiang Lu, Keming Lu, Jianxin Ma, Rui Men, Xingzhang Ren, Xuancheng Ren, Chuanqi Tan, Sinan Tan, Jianhong Tu, Peng Wang, Shijie Wang, Wei Wang, Shengguang Wu, Benfeng Xu, Jin Xu, An Yang, Hao Yang, Jian Yang, Shusheng Yang, Yang

Yao, Bowen Yu, Hongyi Yuan, Zheng Yuan, Jianwei Zhang, Xingxuan Zhang, Yichang Zhang, Zhenru Zhang, Chang Zhou, Jingren Zhou, Xiaohuan Zhou, and Tianhang Zhu. Qwen technical report, 2023. URL: https://arxiv.org/abs/2309.16609, arXiv:2309.16609.

[34] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. Language models are few-shot learners, 2020. URL: https://arxiv.org/abs/2005.14165, arXiv:2005.14165.

[35] Hongshen Chen, Xiaorui Liu, Dawei Yin, and Jiliang Tang. A survey on dialogue systems: Recent advances and new frontiers. *SIGKDD Explor. Newsl.*, 19(2):25–35, November 2017. URL: https://doi.org/10.1145/3166054.3166058, doi:10.1145/3166054.3166058.

[36] Deuksin Kwon, Sunwoo Lee, Ki Hyun Kim, Seojin Lee, Taeyoon Kim, and Eric Davis. What, when, and how to ground: Designing user persona-aware conversational agents for engaging dialogue. In Sunayana Sitaram, Beata Beigman Klebanov, and Jason D Williams, editors, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 5: Industry Track)*, pages 707–719, Toronto, Canada, July 2023. Association for Computational Linguistics. URL: https://aclanthology.org/2023.acl-industry.68, doi:10.18653/v1/2023.acl-industry.68.

[37] Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. Personalizing dialogue agents: I have a dog, do you have pets too? In Iryna Gurevych and Yusuke Miyao, editors, *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2204–2213, Melbourne, Australia, July 2018. Association for Computational Linguistics. URL: https://aclanthology.

org/P18-1205, doi:10.18653/v1/P18-1205.

[38] Yi-Pei Chen, Noriki Nishida, Hideki Nakayama, and Yuji Matsumoto. Recent trends in personalized dialogue generation: A review of datasets, methodologies, and evaluations. In Nicoletta Calzolari, Min-Yen Kan, Veronique Hoste, Alessandro Lenci, Sakriani Sakti, and Nianwen Xue, editors, *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024)*, pages 13650–13665, Torino, Italia, May 2024. ELRA and ICCL. URL: https://aclanthology.org/2024.lrec-main.1192.

[39] Haoyu Song, Yan Wang, Kaiyan Zhang, Wei-Nan Zhang, and Ting Liu. BoB: BERT over BERT for training persona-based dialogue models from limited personalized data. In Chengqing Zong, Fei Xia, Wenjie Li, and Roberto Navigli, editors, *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 167–177, Online, August 2021. Association for Computational Linguistics. URL: https://aclanthology.org/2021.acl-long.14, doi:10.18653/v1/2021.acl-long.14.

[40] Yoonna Jang, Jung Hoon Lim, Yuna Hur, Dongsuk Oh, Suhyune Son, Yeonsoo Lee, Donghoon Shin, Seungryong Kim, and Heuiseok Lim. Call for customized conversation: Customized conversation grounding persona and knowledge. In *AAAI Conference on Artificial Intelligence*, 2021. URL: https://api.semanticscholar.org/CorpusID:245218804.

[41] Jungwoo Lim, Myugnhoon Kang, Yuna Hur, Seung Won Jeong, Jinsung Kim, Yoonna Jang, Dongyub Lee, Hyesung Ji, DongHoon Shin, Seungryong Kim, and Heuiseok Lim. You truly understand what I need : Intellectual and friendly dialog agents grounding persona and knowledge. In Yoav Goldberg, Zornitsa Kozareva, and Yue Zhang, editors, *Findings of the Association for Computational Linguistics: EMNLP 2022*, pages 1053–1066, Abu Dhabi, United Arab Emirates, December 2022. Association for Computational Linguistics. URL: https://aclanthology.org/2022.findings-

emnlp.75, doi:10.18653/v1/2022.findings-emnlp.75.

[42] Hanxun Zhong, Zhicheng Dou, Yutao Zhu, Hongjin Qian, and Ji-Rong Wen. Less is more: Learning to refine dialogue history for personalized dialogue generation. In Marine Carpuat, Marie-Catherine de Marneffe, and Ivan Vladimir Meza Ruiz, editors, *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 5808–5820, Seattle, United States, July 2022. Association for Computational Linguistics. URL: https://aclanthology.org/2022.naacl-main.426, doi:10.18653/v1/2022.naacl-main.426.

[43] Xinchao Xu, Zhibin Gou, Wenquan Wu, Zheng-Yu Niu, Hua Wu, Haifeng Wang, and Shihang Wang. Long time no see! open-domain conversation with long-term persona memory. In Smaranda Muresan, Preslav Nakov, and Aline Villavicencio, editors, *Findings of the Association for Computational Linguistics: ACL 2022*, pages 2639–2650, Dublin, Ireland, May 2022. Association for Computational Linguistics. URL: https://aclanthology.org/2022.findings-acl.207, doi:10.18653/v1/2022.findings-acl.207.

[44] Ruijun Chen, Jin Wang, Liang-Chih Yu, and Xuejie Zhang. Learning to memorize entailment and discourse relations for persona-consistent dialogues. In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence*, AAAI'23/IAAI'23/EAAI'23. AAAI Press, 2023. URL: https://doi.org/10.1609/aaai.v37i11.26489, doi:10.1609/aaai.v37i11.26489.

[45] Fucheng Wang, Yunfei Yin, Faliang Huang, and Kaigui Wu. Please don't answer out of context: Personalized dialogue generation fusing persona and context. In *2023 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, 2023. doi:10.1109/IJCNN54540.2023.10191649.

[46] Chen Henry Wu, Yinhe Zheng, Xiaoxi Mao, and Minlie Huang. Transferable persona-grounded dialogues via grounded minimal edits. In Marie-

Francine Moens, Xuanjing Huang, Lucia Specia, and Scott Wen-tau Yih, editors, *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 2368–2382, Online and Punta Cana, Dominican Republic, November 2021. Association for Computational Linguistics. URL: https://aclanthology.org/2021.emnlp-main.183, doi:10.18653/v1/2021.emnlp-main.183.

[47] Ziyi Zhou, Ying Shen, Xuri Chen, and Dongqing Wang. Gerp: A personality-based emotional response generation model. *Applied Sciences*, 2023. URL: https://api.semanticscholar.org/CorpusID:258267089.

[48] Jing Yang Lee, Kong Aik Lee, and Woon Seng Gan. Improving contextual coherence in variational personalized and empathetic dialogue agents. In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 7052–7056, 2022. doi:10.1109/ICASSP43922.2022.9747458.

[49] Andrea Madotto, Zhaojiang Lin, Chien-Sheng Wu, and Pascale Fung. Personalizing dialogue agents via meta-learning. In Anna Korhonen, David Traum, and Lluís Màrquez, editors, *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5454–5459, Florence, Italy, July 2019. Association for Computational Linguistics. URL: https://aclanthology.org/P19-1542, doi:10.18653/v1/P19-1542.

[50] Zhengyi Ma, Zhicheng Dou, Yutao Zhu, Hanxun Zhong, and Ji-Rong Wen. One chatbot per person: Creating personalized chatbots based on implicit user profiles. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '21, page 555–564, New York, NY, USA, 2021. Association for Computing Machinery. URL: https://doi.org/10.1145/3404835.3462828, doi:10.1145/3404835.3462828.

[51] Qiushi Huang, Yu Zhang, Tom Ko, Xubo Liu, Bo Wu, Wenwu Wang, and H Tang. Personalized dialogue generation with persona-adaptive attention. In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications*

of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence, AAAI'23/IAAI'23/EAAI'23. AAAI Press, 2023. URL: https://doi.org/10.1609/aaai.v37i11.26518, doi:10.1609/aaai.v37i11.26518.

[52] Itsugun Cho, Dongyang Wang, Ryota Takahashi, and Hiroaki Saito. A personalized dialogue generator with implicit user persona detection. In Nicoletta Calzolari, Chu-Ren Huang, Hansaem Kim, James Pustejovsky, Leo Wanner, Key-Sun Choi, Pum-Mo Ryu, Hsin-Hsi Chen, Lucia Donatelli, Heng Ji, Sadao Kurohashi, Patrizia Paggio, Nianwen Xue, Seokhwan Kim, Younggyun Hahm, Zhong He, Tony Kyungil Lee, Enrico Santus, Francis Bond, and Seung-Hoon Na, editors, *Proceedings of the 29th International Conference on Computational Linguistics*, pages 367–377, Gyeongju, Republic of Korea, October 2022. International Committee on Computational Linguistics. URL: https://aclanthology.org/2022.coling-1.29.

[53] Yihong Tang, Bo Wang, Miao Fang, Dongming Zhao, Kun Huang, Ruifang He, and Yuexian Hou. Enhancing personalized dialogue generation with contrastive latent variables: Combining sparse and dense persona. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 5456–5468, Toronto, Canada, July 2023. Association for Computational Linguistics. URL: https://aclanthology.org/2023.acl-long.299, doi:10.18653/v1/2023.acl-long.299.

[54] Cheng Li, Ziang Leng, Chenxi Yan, Junyi Shen, Hao Wang, Weishi MI, Yaying Fei, Xiaoyang Feng, Song Yan, HaoSheng Wang, Linkang Zhan, Yaokai Jia, Pingyu Wu, and Haozhen Sun. Chatharuhi: Reviving anime character in reality via large language model, 2023. URL: https://arxiv.org/abs/2308.09597, arXiv:2308.09597.

[55] Leonard Salewski, Stephan Alaniz, Isabel Rio-Torto, Eric Schulz, and Zeynep Akata. In-context impersonation reveals large language models' strengths and biases. In *Thirty-seventh Conference on Neural Information Processing Systems*, 2023. URL: https://openreview.net/forum?id=CbsJ53LdKc.

[56] Hang Jiang, Xiajie Zhang, Xubo Cao, Cynthia Breazeal, Deb Roy, and Jad Kabbara. Personallm: Investigating the ability of large language models to express personality traits, 2024. URL: https://arxiv.org/abs/2305.02547, arXiv:2305.02547.

[57] Yifan Liu, Wei Wei, Jiayi Liu, Xianling Mao, Rui Fang, and Dangyang Chen. Improving personality consistency in conversation by persona extending. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, CIKM '22, page 1350–1359, New York, NY, USA, 2022. Association for Computing Machinery. URL: https://doi.org/10.1145/3511808.3557359, doi:10.1145/3511808.3557359.

[58] Junkai Zhou, Liang Pang, Huawei Shen, and Xueqi Cheng. SimOAP: Improve coherence and consistency in persona-based dialogue generation via over-sampling and post-evaluation. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 9945–9959, Toronto, Canada, July 2023. Association for Computational Linguistics. URL: https://aclanthology.org/2023.acl-long.553, doi:10.18653/v1/2023.acl-long.553.

[59] Weize Chen, Yusheng Su, Jingwei Zuo, Cheng Yang, Chenfei Yuan, Chi-Min Chan, Heyang Yu, Yaxi Lu, Yi-Hsin Hung, Chen Qian, Yujia Qin, Xin Cong, Ruobing Xie, Zhiyuan Liu, Maosong Sun, and Jie Zhou. Agentverse: Facilitating multi-agent collaboration and exploring emergent behaviors. In *The Twelfth International Conference on Learning Representations*, 2024. URL: https://openreview.net/forum?id=EHg5GDnyq1.

[60] Quan Tu, Chuanqi Chen, Jinpeng Li, Yanran Li, Shuo Shang, Dongyan Zhao, Ran Wang, and Rui Yan. Characterchat: Learning towards conversational ai with personalized social support, 2023. URL: https://arxiv.org/abs/2308.10278, arXiv:2308.10278.

[61] Liang Chen, Hongru Wang, Yang Deng, Wai Chung Kwan, Zezhong Wang, and Kam-Fai Wong. Towards robust personalized dialogue generation

via order-insensitive representation regularization. In Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki, editors, *Findings of the Association for Computational Linguistics: ACL 2023*, pages 7337–7345, Toronto, Canada, July 2023. Association for Computational Linguistics. URL: `https://aclanthology.org/2023.findings-acl.462`, `doi:10.18653/v1/2023.findings-acl.462`.

[62] Hana Kim, Kai Ong, Seoyeon Kim, Dongha Lee, and Jinyoung Yeo. Commonsense-augmented memory construction and management in long-term conversations via context-aware persona refinement. In Yvette Graham and Matthew Purver, editors, *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 104–123, St. Julian's, Malta, March 2024. Association for Computational Linguistics. URL: `https://aclanthology.org/2024.eacl-short.11`.

[63] Young-Jun Lee, Chae-Gyun Lim, Yunsu Choi, Ji-Hui Lm, and Ho-Jin Choi. PERSONACHATGEN: Generating personalized dialogues using GPT-3. In Heuiseok Lim, Seungryong Kim, Yeonsoo Lee, Steve Lin, Paul Hongsuck Seo, Yumin Suh, Yoonna Jang, Jungwoo Lim, Yuna Hur, and Suhyune Son, editors, *Proceedings of the 1st Workshop on Customized Chat Grounding Persona and Knowledge*, pages 29–48, Gyeongju, Republic of Korea, October 2022. Association for Computational Linguistics. URL: `https://aclanthology.org/2022.ccgpk-1.4`.

[64] Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yi Dai, Jiawei Sun, Meng Wang, and Haofen Wang. Retrieval-augmented generation for large language models: A survey, 2024. URL: `https://arxiv.org/abs/2312.10997`, `arXiv:2312.10997`.

[65] Hao Gao, Ruoqing Guo, and Qingqing You. Parasocial interactions in otome games: Emotional engagement and parasocial intimacy among chinese female players. *Media and Communication*, 13, 09 2024. `doi:10.17645/mac.8662`.

[66] D. George and P. Mallery. Reliability analysis. In *IBM SPSS statistics 25 step by step*, pages 249–260. Routledge, 2018.

[67] R. F. DeVellis and C. T. Thorpe. *Scale development: Theory and applications*. Sage Publications, 2021.

[68] J. D. Brown. The cronbach alpha reliability estimate. *JALT Testing & Evaluation SIG Newsletter*, 6(1), 2002.

[69] R. Eisinga, M. T. Grotenhuis, and B. Pelzer. The reliability of a two-item scale: Pearson, cronbach, or spearman-brown? *International Journal of Public Health*, 58:637–642, 2013.

[70] R. H. Gim Chung, B. S. Kim, and J. M. Abreu. Asian american multidimensional acculturation scale: Development, factor analysis, reliability, and validity. *Cultural Diversity and Ethnic Minority Psychology*, 10(1):66, 2004.

[71] J. Hauke and T. Kossowski. Comparison of values of pearson's and spearman's correlation coefficients on the same sets of data. *Quaestiones Geographicae*, 30(2):87–93, 2011.

[72] J. S. Long and J. Freese. *Regression models for categorical dependent variables using Stata*, volume 7. Stata Press, 2006.

[73] Donald Winslow Fiske. Consistency of the factorial structures of personality ratings from different sour sources. *Journal of abnormal psychology*, 44 3:329–44, 1949. URL: `https://api.semanticscholar.org/CorpusID:25439248`.

[74] James Russell and Albert Mehrabian. Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11:273–294, 09 1977. `doi:10.1016/0092-6566(77)90037-X`.

[75] Iris Bakker, Theo Van der Voordt, Jan Boon, and Peter Vink. Pleasure, arousal, dominance: Mehrabian and russell revisited. *Current Psychology*, 33:405–421, 10 2014. `doi:10.1007/s12144-014-9219-4`.

[76] Albert Mehrabian. Analysis of the big-five personality factors in terms of the pad temperament model. *Australian Journal of Psychology*, 48:86–92, 1996. URL: `https://api.semanticscholar.org/CorpusID:145137143`.

[77] Albert Mehrabian. Relationships among three general approaches to personality description. *The Journal of Psychology*, 129(5):565–581, 1995. PMID: 7473304. doi:10.1080/00223980.1995.9914929.

Appendices

A. Programming for Multi-Agent Personalized Dialogue Generation for Otome Game

```
import streamlit as st
from otome_game_agent import (
    CharacterAgent,
    SentimentAnalysisAgent,
    LocationWeatherAgent,
    MemoryBlock,
    DialogueAgent,
)

character_agent = CharacterAgent()
sentiment_agent = SentimentAnalysisAgent()
location_weather_agent = LocationWeatherAgent()
memory_block = MemoryBlock()
dialogue_agent = DialogueAgent()

st.title("Interactive␣Otome␣Game␣Agent␣Platform")

st.header("Personalized␣Dialogue␣System")
user_input = st.text_area("Enter␣your␣message:")

if st.button("Send␣Message"):
    if user_input:
        memory_block.add_to_memory(user_input, "user")

        character_emotion = character_agent.analyze_emotion()
        memory_block.add_to_memory(character_emotion, "
            character_emotion")
```

```
        user_emotion = sentiment_agent.analyze_emotion(user_input
            )
        memory_block.add_to_memory(user_emotion, "user_emotion")

        location_weather = location_weather_agent.
            get_location_weather(user_input)
        memory_block.add_to_memory(location_weather, "
            location_weather")

        personalized_response = dialogue_agent.generate_response(
            user_input=user_input,
            user_emotion=user_emotion,
            character_emotion=character_emotion,
            location_weather=location_weather,
            memory_block=memory_block.get_memory()
        )

        memory_block.add_to_memory(personalized_response, "agent"
            )
        st.write("Character␣␣␣Response:")
        st.write(personalized_response)
    else:
        st.warning("Please␣enter␣a␣message␣to␣continue.")

st.sidebar.header("Dialogue␣History")
for i, entry in enumerate(memory_block.get_memory()):
    st.sidebar.write(f"{entry['role'].capitalize()}:␣{entry['
        message']}")

st.sidebar.markdown("---")
st.sidebar.write("Powered␣by␣Interactive␣Otome␣Game␣Agents")
```