Chatify:

An AI-assisted Platform Offering Gamified Multiple-choice Activities Designed Specifically for ASD Children to Improve Their Empathy Ability and Social Skills

Antong Dai,

Chufei Wang,

Fuyu Qiu,

Haoqing Zhang,

Ziyi Yang

* Ranked in alphabetical order

Abstract:

As a special group, ASD teenagers have difficulty understanding other people's thoughts, developing empathy, and communicating smoothly. Studies have shown that constantly presenting empathetic scenarios and practicing continuously can help improve the empathic ability of autistic teenagers with ASD. However, there are currently many gaps in this research, and it is rare. Chatify uses the form of a mini program, combined with the latest science and technology, to provide empathy training and demonstration for a 15-year-old boy through multiple-choice questions. Pre-test, training task, and follow-up assessment results demonstrated maintenance of training gains. Through the promotion of such projects, researchers can help more similar groups improve their empathy and help them move towards a normal life.

作为一个特殊的群体，ASD 青少年在理解他人的想法、发展同理心和顺畅沟通方面存在困难。研究表明，持续呈现共情场景和持续练习有助于提高自闭症青少年的共情能力。然而，目前在这方面的研究还存在很多空白，而且很少。Chatify 以小程序的形式，结合最新的科学技术，通过选择题的形式，为一名 15 岁的男孩提供同理心的训练和示范。测试前、培训任务和后续评估结果表明，培训成果得到了维持。通过这些项目的推广，研究人员可以帮助更多相似的群体提高他们的同理心，帮助他们走向正常的生活。

Content

1. Introduction

Autism Spectrum Disorder (ASD) is a developmental disability caused by differences in the brain. People with ASD often have problems with social communication and interaction, and restricted or repetitive behaviors or interests. According to statistics, the estimated prevalence rate of ASD among Chinese children (6-12 years old) is about 0.7%, which is to say 1 ASD child in every about 143 children.

Today, communication skills play an important role in people's daily social life. However, due to the symptoms of ASD, ASD children have little or no knowledge about empathy or social skills. They do not know how to greet people in public places, what is empathy when facing people who are in trouble, and even how to express their thoughts.

However, after researchers searched for existing ASD training platforms and tried their products, researchers found that most of them focus on very simple tasks, such as training ASD children's generalization skills by showing them different pictures; telling stories to children by playing videos; and enhancing children's logic skills through simple Math problems, and researchers did not find any platform or organization has a training program to help ASD children enhance their empathy or social skills. Thus, researchers came up with the idea to create a platform focused on training empathy and social skills.

Currently, according to our survey, offline institutions and schools face many problems. Firstly, they are rare in suburban and rural areas. Secondly, each site can receive very few children. Based on our survey in an ASD training institution in Shanghai, it can only receive 2-3 students every weekend. Finally, they do not cover a large enough area in the city so that most ASD children can visit them conveniently. Moreover, offline training institution is also a big challenge for ASD children

themselves, since going to a completely unfamiliar place and having lessons there may have negative effects on them, which may decrease the effectiveness of training. Thus, researchers generally want to make an online platform.

In the end, researchers decide to make our platform in the form of a WeChat miniprogram. WeChat is the most popular instant messaging application in China. To use the miniprogram on WeChat, people only need to scan a QR code, and it is free for all people.

In this paper, due to the limitation of getting in touch with enough ASD children, our study will be a case study of a 15-year-old boy Dongdong. The boy met DSM-IV criteria for ASD. Informed consent and written consent were obtained from his mother. The boy's mother agreed to assist him with the use of this mini program for about 3 weeks, through which the effectiveness of our mini program and our training methods were tested through standardized measurements.

2. Preliminary Research

2.1 Empathy in ASD Children

Human social life requires empathy as a fundamental element. It necessitates the capacity to comprehend the mental condition of another and react appropriately in terms of feeling or acting. Atypical empathetic responses have been reported in people with autism spectrum disorder (ASD), which can hinder social relationships and communication.

Theory of mind (ToM) deficiencies are linked to empathy impairment in the ASD group. ToM is the capacity to identify that one's own and other people's mental states (beliefs, intents, wants, emotions, and knowledge) may differ (Premack & Woodruff, 1978). A strongly developed ToM aids in conflict resolution, social skill development,

and behavior prediction (Gweon and Saxe, 2013). It has been proposed that people with ASD lack Theory of Mind (ToM) since many of them struggle to assign mental states to other people.

According to the social motivation hypothesis, underlying brain anomalies in reward processing make social cues less enjoyable for those with ASD.

A critical evaluation that focuses primarily on empathy in ASD is necessary given the growing prevalence of autism spectrum disorders and the significance of empathy in human social functioning. There is currently no known treatment for autism, and while those that are available are successful in enhancing the quality of life and functional independence, few of them concentrate on enhancing social emotion regulation.

2.2 Interpersonal Reactivity Index-Chinese Version

Davis (1980) created 28 items answered on a 5-point Likert scale ranging from "Does not describe me well" to "Describes me very well". The measure has 4 subscales, each made up of 7 different items:

Perspective Taking - the tendency to spontaneously adopt the psychological point of view of others;

Fantasy - taps respondents' tendencies to transpose themselves imaginatively into the feelings and actions of fictitious characters in books, movies, and plays;

Empathic Concern - assesses "other-oriented" feelings of sympathy and concern for unfortunate others;

Personal Distress - measures “self-oriented feelings or personal anxiety and unease intense interpersonal settings”

INTERPERSONAL REACTIVITY INDEX

The following statements inquire about your thoughts and feelings in a variety of situations. For each item, indicate how well it describes you by choosing the appropriate letter on the scale at the top of the page: A, B, C, D, or E. When you have decided on your answer, fill in the letter next to the item number.

READ EACH ITEM CAREFULLY BEFORE RESPONDING. Answer as honestly as you can. Thank you.

A	B	C	D	E
Does not describe me well				Describes very well

When you have decided on your answer, fill in the letter next to the item number. READ EACH ITEM CAREFULLY BEFORE RESPONDING. Answer as honestly as you can. Thank you.

Sample questions:

1. I daydream and fantasize, with some regularity, about things that might happen to me.
2. I often have tender, concerned feelings for people less fortunate than me.
3. I sometimes find it difficult to see things from the "other guy's" point of view.
4. Sometimes I don't feel very sorry for other people when they are having problems.

2.3 Case Manager in ASD Treatment

The case manager is often seen in medical or health care of ASD patients. They arrange and coordinate all appointments and their sequence, rationalizing times and spaces to minimize the ASD patient's exposure to long waits, multiple visits, multiple days in the hospital, and difficult and avoidable situations. This process facilitates the patient's visits with other specialists, as well as the ensuing examinations and follow-up visits.

The case manager also provides the receiving agency with information about the patient's characteristics and, if necessary, goes with the patient and family. A few of the disciplines are extremely popular or highly challenging for patients to access. In these situations, protocols are created to get around particular challenges or traits unique to the field. Providing anesthesia for small procedures and having tools available for visual assessment are two examples of this.

Concluding from this function of case manager, it is hypothesized that case manager can play a similar role in the empathy ability improvement of ASD children

by tailoring individualized training tasks for the patients.

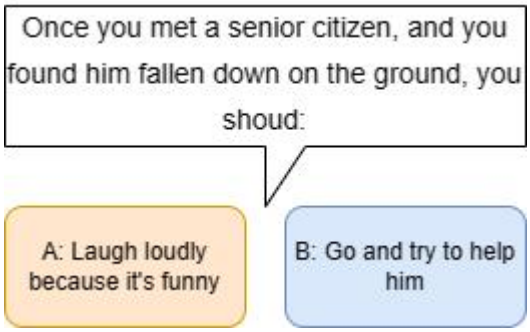
Research has shown that tasks aimed at improving emotion recognition, perspective taking, and emotional affection induced by another person's situation have successful outcomes in improving empathy ability in ASD children aged from 6-17 years old (mean=17.3 years).(Schwenck et al., 2011).

3. Innovative Program

3.1 Training Method:

Researchers will use multiple-choice questions to train ASD children's empathy and social skills.

Each question tells participants a simple scene, usually a short dialogue, of their familiar daily environment. Besides, researchers set up 2 choices below each question’s description. Each choice describes a possible reaction or inner thought corresponding to the question’s description. The following image is an example:



researchers set a correct answer (in this case, it is option B), and an obvious wrong answer. Each of these options demonstrates a different understanding of the question: one shows enough empathy and normal social reactions, while the other does not.

Before the case study, researchers invited about 100 teenagers of Dongdong's age

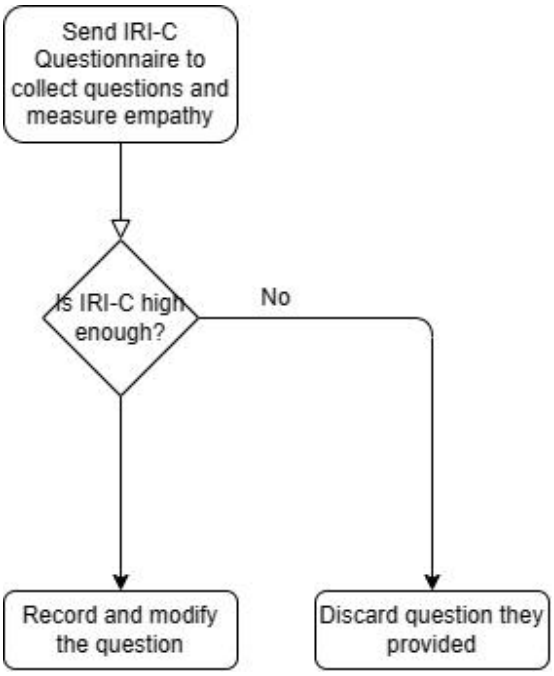
to write these questions and compiled 25 questions from them.

Firstly, researchers will do a pre-test on one of our participants --- Dongdong. researchers asked him to do 10 questions (2 groups) and collected his answers, then calculated the rate of correction.

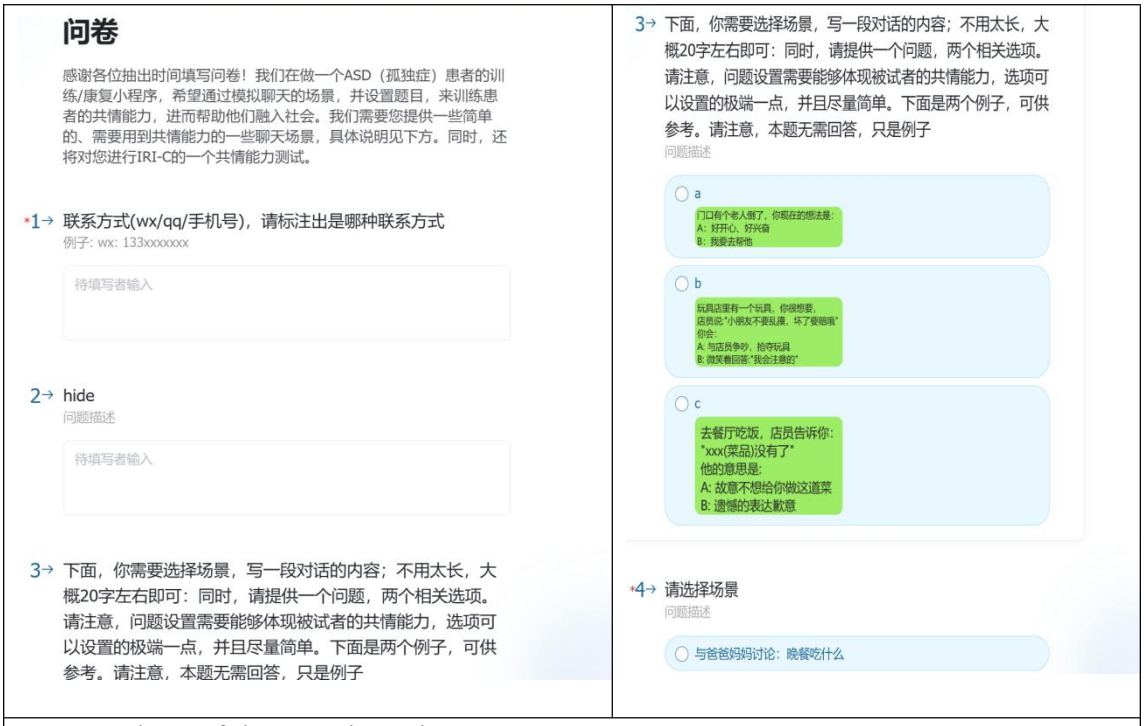
Next, researchers conduct a 2-week training program through our mini program online and invite his mother to assist.

Finally, after the training, researchers do a post-test to validate the effectiveness of our training.

3.2 Questions collection



To collect enough questions, researchers sent out a questionnaire among teenagers of Dongdong's age. In the questionnaire, researchers asked participants to do two tasks: complete an IRI-C test and input one scene or dialogue that demonstrates their empathy from their social experiences.



2 screenshots of the questionnaire

Since researchers need to make questions that can demonstrate high empathy, researchersfilter these questions by participants' IRI-C results.

Researchers stipulate that the IRI-C result > 60 represents that the participant has a relatively higher empathy, which means that his or her questions can be used in our training materials.

In the end, researchers selected about 25 questions from the questionnaire and compiled them into our first training question group.

3.3 Miniprogram

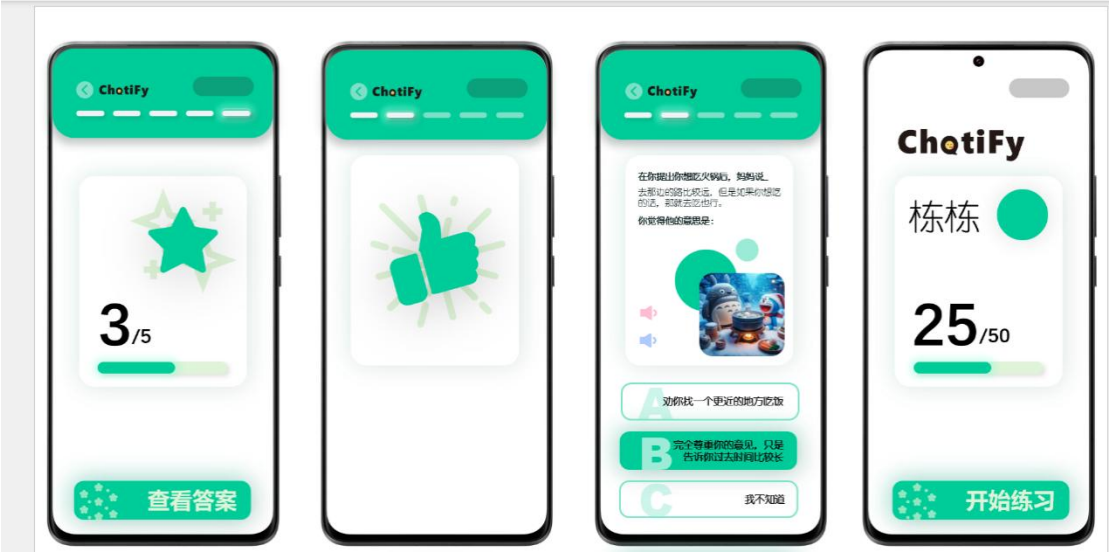
3.3.1 UI Design

On the mini program, it will show the user 5 questions each day. Each question is a single and simple scene of people, especially teenagers, daily life. Usually, they are some simple dialogues. Then, it also shows 2 choices. One of them indicates the right action people need to take, or the right thought people should have based on their

empathy. Besides, each choice contradicts the other obviously, so ASD children can easily distinguish them.

During our communication, researchers found green is Dongdong's favorite color, so researchers decided to use green as the main color of the software, while adding other bright colors to make the interface more vivid. At the same time, green provides a bright and lively color, and researchers also hope that testers' first feeling when seeing this interface is pleasant.

In the choice of design style, researchers realized that as educational software, it must be concise and clear, so the entire user interface design uses large color blocks and text as the main elements of the design, with light and shadow as embellishments. Improve the readability and visual focus of the entire interface. By combining the two styles of skeuomorphism and flatness, and using rounded rectangles, the entire software has a sense of intimacy and can be operated like a fool. In addition, pictures and animations are added to make the entire software more in line with the aesthetics and mentality of autistic teenagers, and vibrations help to improve the interactions with ASD teenagers.



3.3.2 Audio and Pictures

In addition, to help Dongdong understand the question more clearly, researchers provide audio that reads the texts of both questions' descriptions and options. For each audio, researchers provide both a female reader and a male reader since researchers want to find out which voice type Dongdong prefers. Before the experiment, researchers supposed that Dongdong would prefer the female reader because it sounds just like his mother.

Moreover, researchers also display a relevant picture of each scene in the question to help Dongdong understand.

Screenshot of the mini program (questions)

3.3.3 Buttons set

In addition, researchers found that Dongdong might say words like "I don't know" when researcherspretest offline, so researchersadded an option: "I don't know" to the program. To encourage Dongdong thinking, researchers set the "I don't know" button to appear 30 seconds after the other options are displayed.

3.4 Encouragement

After Dongdong makes his choice, a new page with some encouraging words will

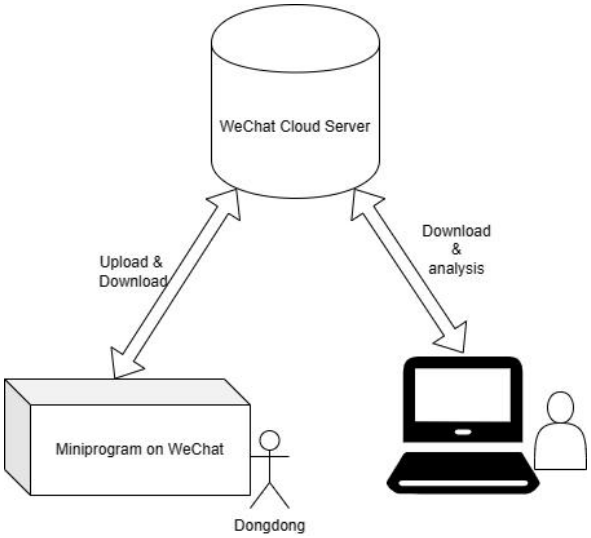
pop up.

3.5 Auto-updating Question Library

Each time the mini program runs, it will first check the update of the question library from the cloud server. If it finds a newer version, it will download that and replace the old one automatically.

3.6 Results collection

All the results, including the correction rate, index of questions Dongdong did, time spent on each question, and which voice (female or male) Dongdong listened to are recorded and uploaded to a cloud server automatically by the miniprogram after Dongdong completes every 5 questions per day. After Dongdong finishes his training, researcherswill download data from the server and analyze them.



3.7 Post-training test

One female researcher meets Dongdong in the office where Dongdong usually has extracurricular class to do the post test.

Post test is a structured interview with 30 questions. Each question comprises of a 10-second video clip, a 15-second asking time and a one-minute choosing time. Video clips (average length 10s) regarding contexts the questions have covered in MCQ were played to Dongodng.

After the completeness of video, the researcher asks questions with similar pattern of the MCQ. Dongdong is asked to choose from the 3 choices (including "I don't know") given orally.

4. Results

4.1 Question Set Validation:

After completion of the miniprogram, all the team members used the program and answered all the questions on it. The result of tests is shown below

member	score
Mike	80%
Erin	100%
Kaykay	100%

Since Tony and Arcadia participated in compiling the questions and making the miniprogram, they were not involved in our validation process.

It is also confirmed that all these members' IRI-C scores reach the predetermined HIGH criteria(>60). According to the result, it can be concluded that our question sets are valid since teenagers with HIGH empathy levels can easily get high correction rates(>90%).

4.2 Pre-test

Researchers spent a weekend meeting with Dongdong. In the pre-test, a printed version of the question set (15 questions in total) is used to test Dongdong's preliminary empathy and social ability. This diagram records the correction for each question (C represents correct, W represents wrong):

1	2	3	4	5
C	W	C	W	W
W	W	W	C	C
W	W	W	W	W

The correction rate is calculated as around 26.77%

4.3Training & Post-test

4.3.1Raw data

In the training process, a 2-week case study is planned to conduct on Dongdong, and 10 groups of data are expected data size; however, due to the schedule conflicts and many other factors, 6 groups of data from the miniprogram during the training process were obtained. Raw data is shown below.

4.3.2 Variables & Description

Notation	Description
N	the serial number of the question answered (0-24)
T	the total time spent answering the question
red block	wrong answer
green block	correct answer

Day	Question										Score	Frequency of male voice	Frequency of female voice	Date (All from 2024)
	Question No.1		Question No.2		Question No.3		Question No.4		Question No.5					
	N	T	N	T	N	T	N	T	N	T				
1	5	90	16	45	17	34	8	33	23	80	40%	10	20	1/13
2	20	93	10	43	12	95	5	37	15	46	20%	11	12	1/15
3	9	67	3	49	16	34	18	61	19	36	20%	8	39	1/17
4	23	77	17	41	13	50	11	43	6	39	60%	13	69	1/21
5	1	50	13	55	19	49	0	41	2	60	80%	12	99	1/22
6	4	55	21	57	8	49	20	29	18	44	60%	20	129	1/23

In the data, these variables are collected:

1.Question index, correction for each single question:

They are used to calculate the overall correction rate and further category.

2.Question remain time:

This value indicates how much time Dongdong spent on each question. Researchers want to use this value to trace the answering speed. Besides, to certain extents, it indicates whether Dongdong reads and answers the question seriously.

3.The times each audio is played:

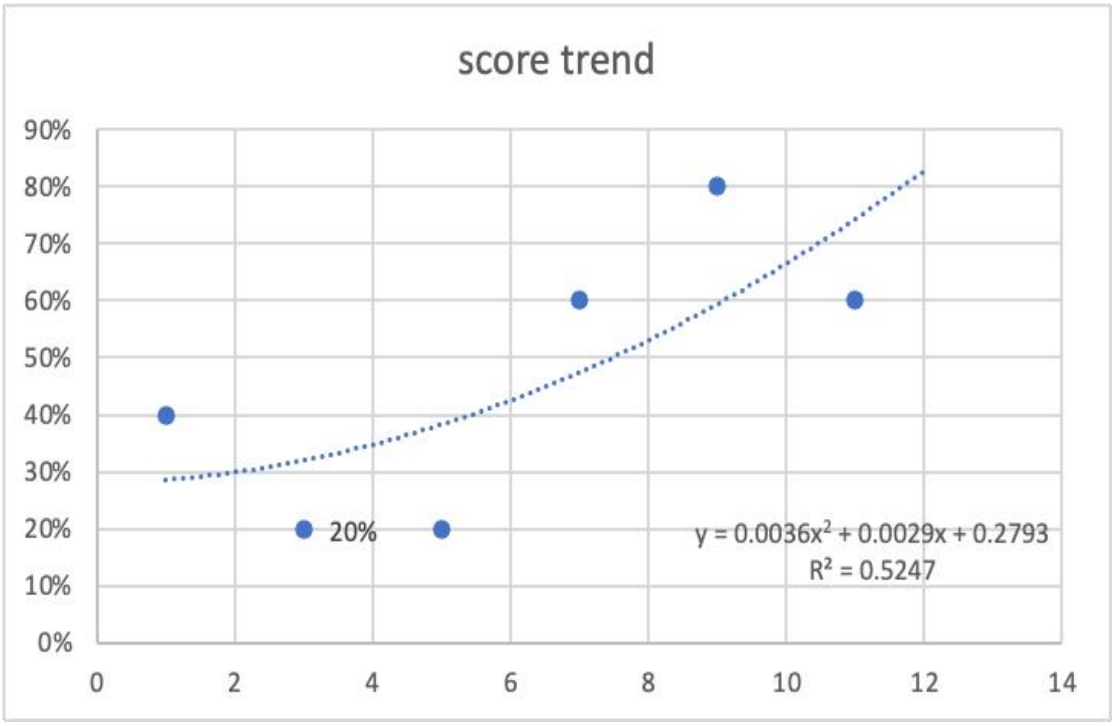
Researchers categorize them into female and male voice. Researchers want to use this value to indicate Dongdong's preference on voices.

4.3.3 Correction Rate Analysis

Then, the correction rates for each day's training is calculated:

day	score
1	40%
2	20%
3	20%
4	60%
5	80%
6	60%

In order to visualize the data and research the potential trends, researchers plot them on a graph:



In this diagram:

X(horizontal)-Axis represents the index of scores (the researchers use odd number 1,3,5,7,9 to scale the whole graph and make it clearer)

Y(vertical)-Axis represents the correction rates(0% to 100%)

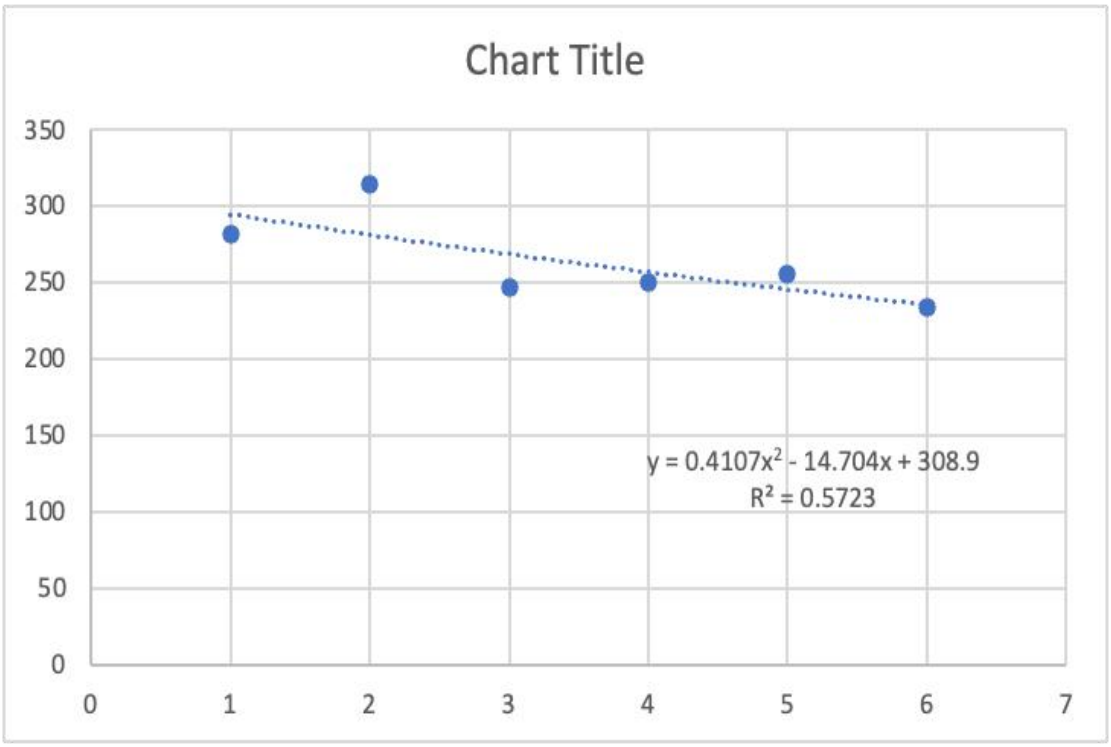
Polynomial regression is used in this estimation because in all regression methods, including linear, exponential and logarithmic regression, polynomial regression has the least R^2 value, indicating higher accuracy.

According to the trend, Dongdong's scores are improving through the training. In detail, the correction rises from the initial about 25% to current about 81%. In the future, the researchers forecast a continuous growth in the correction rate.

4.3.4 Remaining Time Analysis

Next step, the remaining time on each question is analyzed:

day	T total
1	282
2	314
3	247
4	250
5	255
6	234

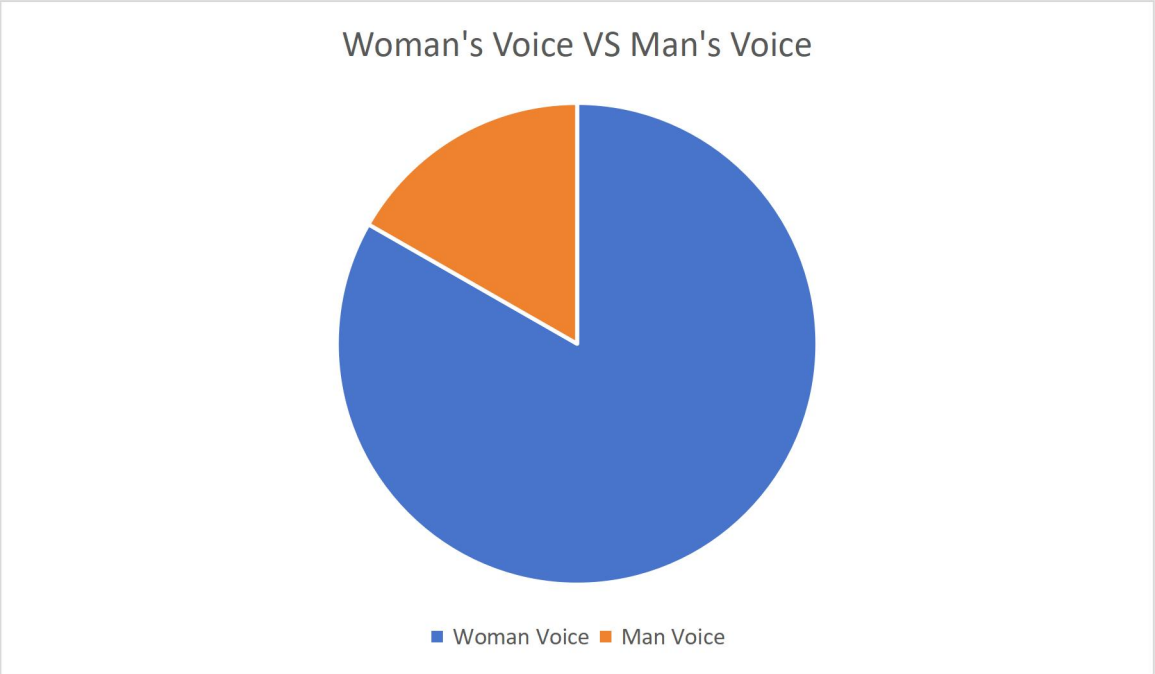


In this table, Y axis represents the summation of time Dongdong spent on each question (Unit: sec)

In this diagram, linear regression is used to estimate the trend. According to the diagram, the study finds that Dongdong spends less time on answering these questions as the training progresses. Therefore, it can be reasonably inferred that

Dongdong becomes more and more proficient in answering the questions. The study can conclude that through the training, Dongdong spends less time reading and understanding each question, which indicates an increase in his reading and thinking ability.

4.3.5 Voice Type Analysis



Also, the researchers sum up the times Dongdong listen to Man voice audio and to Woman voice audio, they are 74:368(man: woman). The diagram above shows the proportion of them. We can conclude that Dongdong prefers Woman voice audio. It is inferred that this is because Dongdong prefers voice like his mother.

4.3.6 Perspective and Behavioral Category

Finally, these questions are classified into 2 classes:

1. Perspective taking: understanding others' attitude and meaning in conversation
2. Behavioral response: awareness of the proper response to others' utterance or current context

Type	Index in Day 1-4	Index in Day 5-6
Perspective	0, 1, 3, 6, 10, 11, 20	7, 8, 9, 10, 11, 12 ,13, 14, 15, 16, 17, 18, 19, 20, 21
Behavioral	2, 4, 5, 7, 8, 9, 12, 13, 14, 15, 16, 17, 18, 19	0, 1, 2 ,3, 4, 5, 6

According to this table, these questions are classified into 2 groups.

Notation	**Description**
N	the serial number of the question answered (0-24)
T	the total time spent answering the question
yellow block	Perceptive
blue block	Behavioral

	Question									
Day	Question No.1		Question No.2		Question No.3		Question No.4		Question No.5	
	N	T	N	T	N	T	N	T	N	T
1	5	90	16	45	17	34	8	33	23	80
2	20	93	10	43	12	95	5	37	15	46
3	9	67	3	49	16	34	18	61	19	36
4	23	77	17	41	13	50	11	43	6	39
5	1	50	13	55	19	49	0	41	2	60
6	4	55	21	57	8	49	20	29	18	44

Based on this category, the correction rate is calculated again.

Type	Perceptive	Behavioral
Correction Rate	66.78%	33.33%

According to the table, Dongdong gets an overall correction rate 66.78% on

perceptive questions and 33.33% on behavioral questions. The study finds that Dongdong always has a higher correction rate on answering those Perceptive questions.

4.4 General Conclusion

In conclusion, our training method is overall effective. This is because, according to the results, Dongdong has an increase in correction rate throughout our training, and it responds to the questions more quickly throughout the training.

5. Conclusion

With fast economic development in China, social welfare is becoming better in modern society. It is good for us to see more and more people start to pay more attention to those ASD children. However, because of many limitations of currently existing ASD training institutions and programs discussed above, researchers believe our innovation of using a mini program to train ASD children's empathy and social skills will provide a better training platform, and promote the development of ASD education fields.

In the future, researchers will continue to collect and design more questions and materials in our mini program, and researchers hope our mini program can be introduced to and help more and more ASD children or their families.

6. Discussion

In the era of artificial intelligence, the rational use of artificial intelligence can help us conduct various research more conveniently and efficiently and apply it to more scenarios. First of all, researcherscan use artificial intelligence to predict more scenes and create communication situations that are closer to life. At the same time, researcherscan replace scenes more conveniently. For example, researcherscan replace places or people in life scenes to allow ASD patients to summarize the scene. and summary. researcherscan also use artificial intelligence to identify patients' needs, extract keywords and key information, conduct semantic text analysis, process patients' answers, provide tips and suggestions, guide the patients, and generate more appropriate questions to make training more targeted.

Limitations of the current study are as follows:

1. The availability of potential subjects makes the results obtained from our participants unable to generalize to more ASD children population.

2. Negotiation with the participants' parents is not adequate to obtain enough number of data groups. Thus, data analysis part lacks substantial numerical data.

3. One type of quantitative data is about the choice of computer-generated male teenager sound or middle-aged female sound, and this is designed to test whether the participant is more willing to hear instructions given by a familiar figure such as his mother. In the future, researchersplan to train AI tools to mimic the sounds of real people instead of using automatically generated sounds in the database.

4. When a participant is doing the tasks, guardians and teachers are needed to ensure he is concentrated on the task and comprehends all the instructions. Full automation in the process of "supervision" has not been achieved yet.

Reference

*Clements, C. C., Zoltowski, A. R., Yankowitz, L. D., Yerys, B. E., Schultz, R. T., & Herrington, J. D. (2018). Evaluation of the social motivation hypothesis of autism: a systematic review and meta-analysis. *JAMA psychiatry*, *75*(8), 797-808.

*Davis, M. H. (1980). Interpersonal reactivity index.

*Gweon, H., & Saxe, R. (2013). Developmental cognitive neuroscience of theory of mind. *Neural circuit development and function in the brain*, *3*, 367-377.

*Pervushina, O.N., & Gileva, D.I. (2023). Possibilities of Forming Empathic Reactions in Children with ASD. *Reflexio*.

*Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and brain sciences*, *1*(4), 515-526.

*Schwenck, C., Mergenthaler, J., Keller, K., Zech, J., Salehi, S., Taurines, R., ... & Freitag, C. M. (2012). Empathy in children with autism and conduct disorder: Group-specific profiles and developmental aspects. *Journal of Child Psychology and Psychiatry*, *53*(6), 651-659.