Trouble Maker: Truth or Dare Prompt Generator Based on Markov Chain

Yibo Fu
Parsons School of Design, The New School
New York, NY, USA
fuy644@newschool.edu

Yuqing Liang
Parsons School of Design, The New School
New York, NY, USA
liany922@newschool.edu

Harpreet Sareen
Parsons School of Design, The New School
New York, NY, USA
sareen@newschool.edu

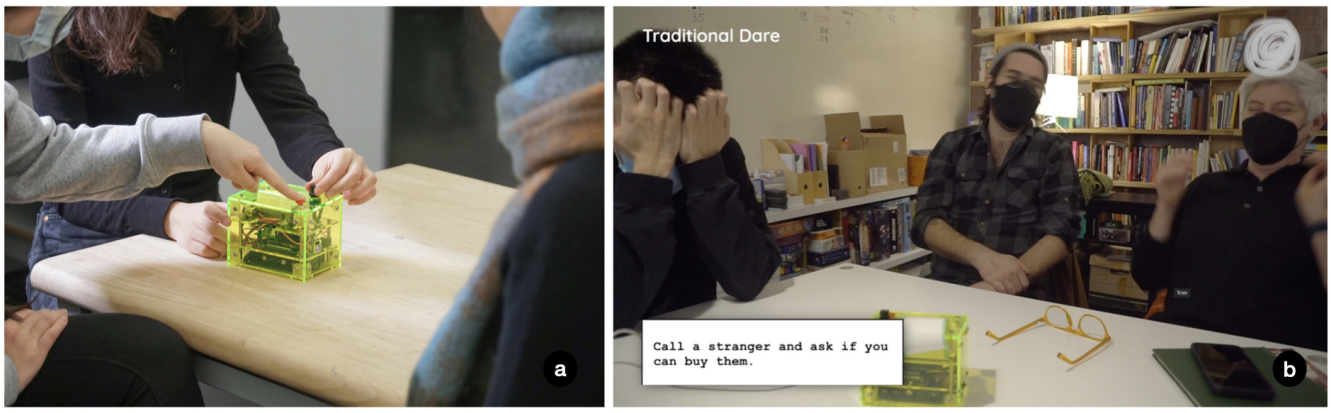

Figure 1: a) Players interacting with the physical device. b) Players' reaction after executing the machine-generated prompt

ABSTRACT

As a classic party game, Truth or Dare players always face the challenge of continuously devising clever prompts by themselves, where prompts can often get repetitive or hard to think up quickly. We present Trouble Maker, an open-source physical device that generates whimsical truth-or-dare prompts using Markov chain with a single button press. Our device autonomously generates surprising and unpredictable prompts using a lightweight machine learning algorithm that is printed instantly. ML-prompt generation and thermal printer are integrated altogether as a tangible artifact. Results from the preliminary user study show prompts based on our method amused and surprised players, and the device increased the overall playfulness and engagement of the game.

CCS CONCEPTS

• **Human-centered computing** → **Interaction design**; Systems and tools for interaction design;

KEYWORDS

Smart device, Party game, Machine learning, Prototyping, Language model, Markov chain

ACM Reference Format:
Yibo Fu, Yuqing Liang, and Harpreet Sareen. 2023. Trouble Maker: Truth or Dare Prompt Generator Based on Markov Chain . In *Companion Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY Companion '23), October 10–13, 2023, Stratford, ON, Canada.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3573382.3616040

CHI PLAY Companion '23, October 10–13, 2023, Stratford, ON, Canada

ACM ISBN 979-8-4007-0029-3/23/10.
https://doi.org/10.1145/3573382.3616040

1 INTRODUCTION

Living in the post-pandemic era where face-to-face parties are now returning, people are in strong need of in-person activities [19] and communication [22]. One classic and common party game, Truth or Dare, popular among friends, acquaintances or strangers, is commonly played during parties and many other gatherings. It places the players in a dilemma between answering a personal question or performing a challenging task. However, running out of truth-or-dare prompts is a common pain point in the game, players either search online with limited and repetitive prompts, or just get tired of thinking up new prompts.

In our work, we address this issue through the creation of a tangible interactive device that can autonomously generate prompts in real-time using a machine learning algorithm. Instead of using a large language model (LLM), we chose Markovify [20], a text version library that uses Markov chain algorithm. In many cases, Markovify is not an ideal algorithm especially with small amounts of input, as there are high chances of generating wrong,

nonsensical sentences. However, the fact that this algorithm is **relatively "primitive"** could **instead result in more surprising and whimsical** Truth or Dare prompts, making it **very unpredictable** from human's perspective, thus increasing the random nature of the game. By using a series of selected text inputs, including both funny and serious content, our algorithm can generate 'infinite' challenges for the players that are often beyond humans' creativity. We then integrated the algorithm into an open-source, hand-palm-size tangible device.

We argue that such design enables more surprising, playful and fun experiences for the Truth or Dare game and amplifies the dialogues between human players. In the preliminary user study conducted in different scenarios, participants found that our real-time generated prompts are highly readable, mostly viable and surprising.

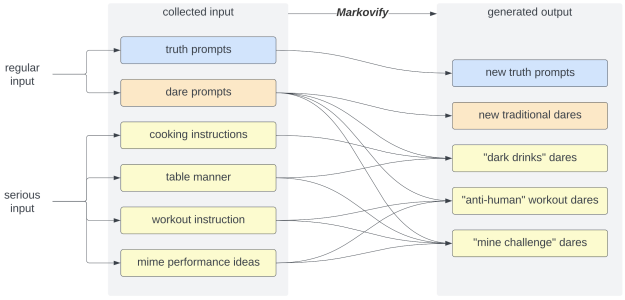

Figure 2: Collected input & generated output

2 RELATED WORK

2.1 Human-AI interaction in games

Integrating AI into games has a long history back to the 20th century [17]. Nowadays, AI has been used in many different ways in games, Treanor et al. [28] categorized a few genres, such as procedural content generation (PCG) [24][4][5], AI-agent playing along with or against human players (NPCs) [18][12], or AI algorithm that automatically balance in-game complexity [27], etc. One of the critical reasons that AI is implemented in many games is that it can react based on users' actions in real time, thus bringing highly customizable gaming experiences. Our work builds along with those works and aims to provide randomized and surprising Truth or Dare prompts. To be more specific, our work lies in the genre of text-based AI games, where AI is used for generating text that human players can interact with. In Crossword Puzzle [30], the algorithm can consistently generate crosswords on large word lists containing thousands of words, which humans cannot constantly generate. AI Dungeon [6] uses GPT-2 [7] to create any worlds, characters, and scenarios for your character to interact with based on text dialogue. Similar to Trouble Maker, those works intend to provide (almost) infinite text-based gaming sources for players to interact with through AI algorithms.

In Trouble Maker, the interaction between players and AI is indirect compared to the previous works, as the dialogue and actions emerged from the prompts happen among human players, and it is the human players who decide and improvise actions and dialogues on top of the generated prompts.

Specifically in our work, we used Markov chain, a mathematical algorithm as the base of our algorithm. Markov chain have been widely used in natural-language-generation applications [14]. White et al. created DrinkMaker [31], which is an algorithm that generates cocktail recipes based on an online cocktail database. Similarly in our work, we take advantage of Markov chain algorithm to enhance the random nature of Truth or Dare game.

2.2 Smart, tangible device for tabletop game

Tabletop games that aim to combine a physical game experience with electronic or digital elements have existed since the 1970s [21]. Kankainen et al. described those smart devices as a "hybrid" tabletop game experience with both digital and physical parts existing. Many consumer products [1][32] have also already explored such "hybrid" game experiences, with most of them focusing on enhancing visual game experiences, whereas our smart device specifically focuses on non-screen tangibility and text-based auto-generation enabled by digital technology.

Thermal printers are widely used in interactive devices [13, 15, 23] due to their convenience for rapid prototyping. In our case, we use a thermal printer to output each prompt in real-time. We hope that the tangibility of the printed prompts could help players focus on the game, enrich the playfulness and create a sense of ritual for the Truth or Dare game.

3 SOFTWARE FLOW & GENERATED PROMPTS

In this section, we describe our software flow (Fig 3) that generates prompts using a lightweight machine learning algorithm (Markovify), which could not only enhance the random nature of the game, but also ensure the readability of the output text. The software first ingests various types of text data, then parses and tunes the data, and finally assembles new sentences as the output prompts.

3.1 Prepare input data

To collect the input data for Markovify on a larger scale (thousands of sentences) from scattered platforms, we built a web-scraping tool using Node.js [3] and Puppeteer [2]. To better enable the unpredictability brought by Markovify, we experimented with a wide range of input types. We specifically looked for text materials that consist of more imperative sentences, as it fits the dare prompts the best. Example input sources include online truth or dare prompts, cookbooks, table manners, and workout instructions.

Based on this, we further developed the original truth or dare game from two modes (truth, dare) to five different modes (Fig 2) . We mixed different input types with one another and categorized the output into five playful categories for users to select:

(1) Remixed traditional truth
(2) Remixed traditional dare
(3) Dark drink dare
(4) Anti-human workout dare
(5) Mime challenge dare

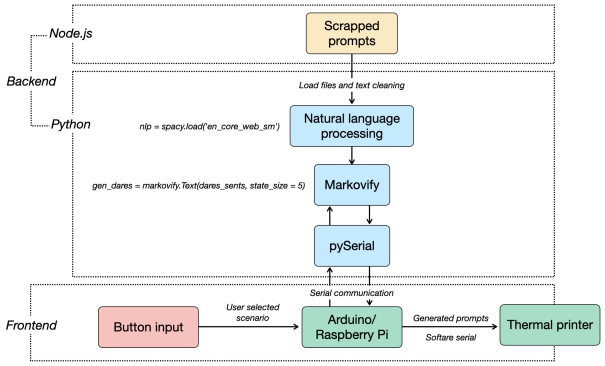

Figure 3: Software flow overview

3.2 Tuning scraped data

After data scraping, we conducted a series of steps on the raw data to make them more reliable without interfering with the "learning process":

(1) Use regular expressions to remove all the punctuation in the text.
(2) Removed sentences started with conjunctions like "then" or "after".
(3) Use Natural Language Processing to recognize sentences and rejoin them with a period.
(4) Convert the sentences into string type.

3.3 Generate prompts

In this final step, we feed text data into Markovify [20], an open-source text-generating Python library that is based on the Markov chain algorithm. The results are controlled by assigning input weights and then constraining them to the optimized length. This generates prompts that are suitable for the Truth or Dare game using the input types above individually or in combinations (Fig 2) .

As the core element in this game, prompts are hardly judged as correct or wrong. Instead, intimate, awkward truth prompts can invoke more interactions among players, potentially helping players learn more about one another. Wild dare prompts can excite the gaming and social atmosphere, often leading to stories everyone will be repeating for years. Surprising, novel prompts could potentially get everyone exicted for unforeseen questions and commands. We were able to observe many such kinds of prompts that are generated by our algorithm. A few examples are:

(1) *Let two people give you a new hairstyle.*
(2) *What is your favorite thing to do in front of your friends that you'd never do in front of your parents?*
(3) *Sing a praise song about a person in the group.*
(4) *What's a skill you wish you didn't have?*
(5) *What's the meanest thing you've ever done in a bathroom?*
(6) *Mix in the Dr. Pepper and beer serve in an old fashioned glass.*
(7) *Put your underwear into fridge for three minutes.*
(8) *Pretend it's cold and you are seeing something exciting or scary.*

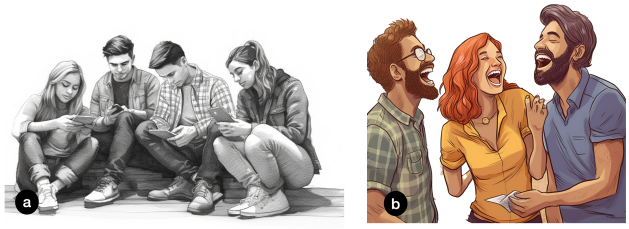

Figure 4: a) illustration of playing Truth or Dare with phones. b) illustration of playing with our device. Both image generated by midjourney, 2023 [10]

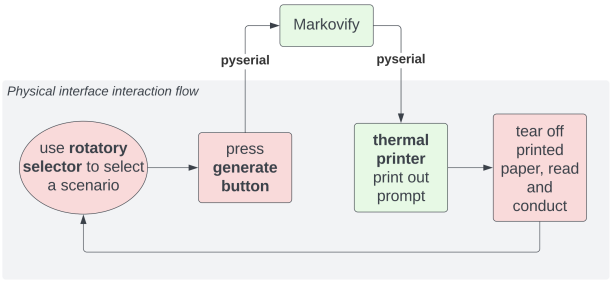

Figure 5: User interface flow

(9) *Go outside and try to sell a piece of trash to someone of the opposite sex and allow a picture to be taken of you.*

4 HARDWARE IMPLEMENTATION

In many party games, phones or laptops can be a distracting factor [16] to both the player themselves and other players during the game. In order to build a more engaging gaming environment among players and get rid of screens, we aim towards a non-screen tangible interface that allows players to focus on interactions in the real world (Fig 4) . Two version prototypes using different embedded system platforms were developed: a) Prototype 1: An Arduino Uno based device. b) Prototype 2: A Raspberry Pi Zero based device.

4.1 User interface flow

To create an easy-to-play device, we designed a simple and intuitive interface (Fig 5) integrated into a palm-sized device (Fig 6), with the following steps to use it.

(1) Select a game mode with the rotary selector.
(2) Press the red button next to the selector and the thermal printer will print a piece of prompt.
(3) Read it and complete the prompted action in front of other players.
(4) Rotate to the next player. If it ends, turn off the on/off switch on the side.

4.2 Embedded system & Open-source hardware

We designed two different prototypes. Prototype 1 (Fig 6.a, b) is a proof-of-concept prototype built using an Arduino Uno and requires

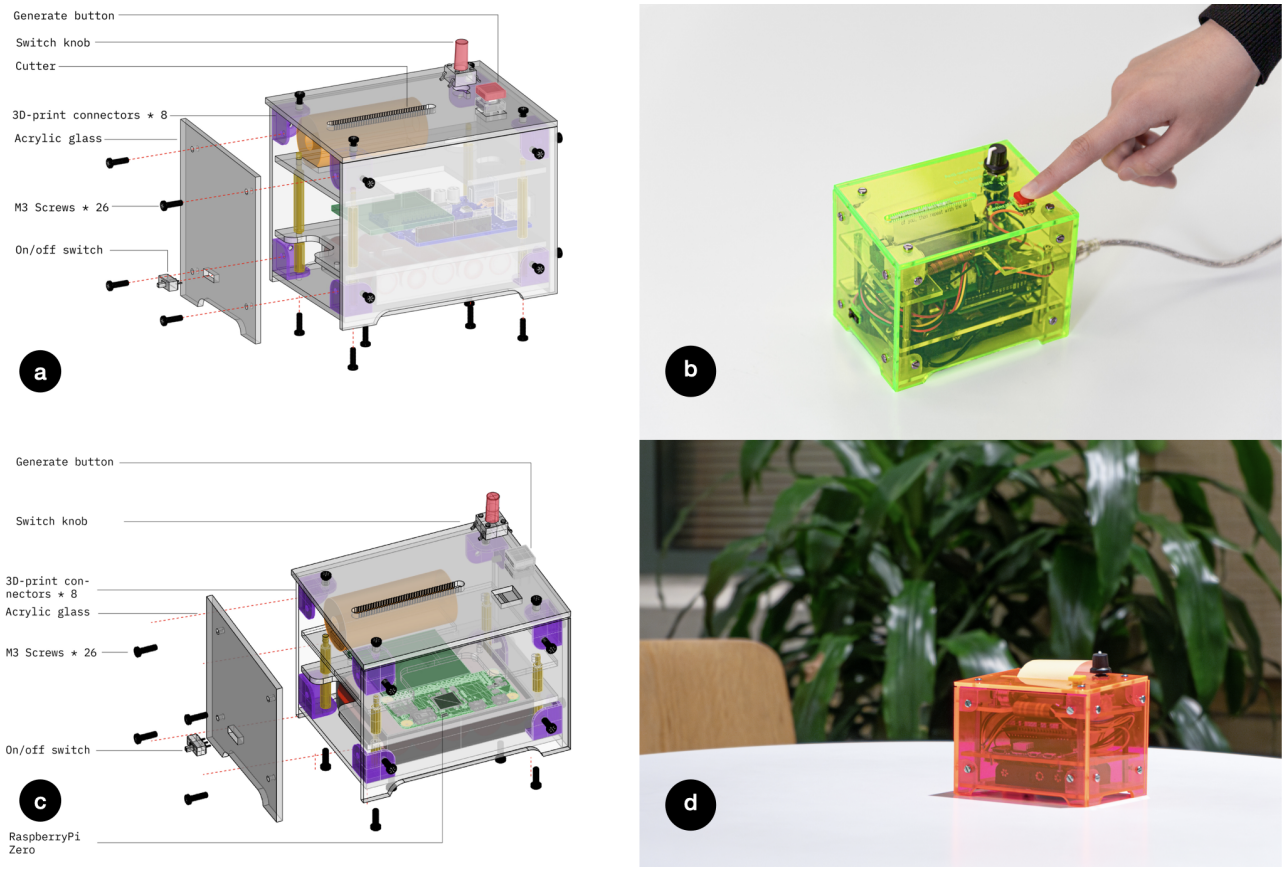

Figure 6: a) Explosion view of Prototype 1. b) Player presses the button on Prototype 1. c) explosion view of Prototype 2 d) Prototype 2

a laptop connected for the language model (back-end) to run on. To further increase the portability and usability of the device, we developed Prototype 2 (Fig 6.c), an untethered, smaller-sized model built with Raspberry Pi Zero that integrated both front-end and back-end (Fig 3) inside the device, allowing the whole generating process to happen locally. Both models provide the same interface and functionality.

With the aim of providing an open-source [11] system, all materials and components are off-the-shelf products. Both prototypes are enclosed with acrylic boxes assembled with M3 screws, making them easy to replicate with our open-source CAD files. Prototype 1 measures 12.4 × 8.6 × 9.8cm (L * W * H), and prototype 2 has the size of 11.4 × 8.6 × 7.9cm. The total cost, important to be kept down to make it easy to self-replicate or potentially mass-produce, was under 50 USD including the microcontroller. Arduino Uno or Raspberry Pi Zero could also be replaced with lower-cost microcontrollers or computers.

4.3 Game Design & Social Configurations

Truth or Dare, as a party game, offers social configuration among players. In our case, we designed this simple tangible interface with the aim of enhancing such social configurations by two main design adjustments.

First, due to the social configuration of the original Truth or Dare game, where players can either play with a few close friends, or with known people, or even a bunch of strangers playing Truth or Dare as a ice-breaking game, we provide multiple modes supporting different scenarios suitable for various player groups, party locations and scenarios. For example, dark drink dares are more suitable for restaurants, bars or home parties with drinks. Anti-human workout dares are interesting to pick in certain public spaces just to let the player to get more 'awkward' with more 'audiences'. On the other hand, more genres also means more selective and categorized generated prompts, thus allowing better targeted questions and executable prompts to be generated without wasting time finding a doable prompt in certain situation.

Second, In spite of this more and more digitalized world, the natural need for direct human and social interaction is underserved [29]. In game field, tabletop games with physical objects can well supported such need [29]. The continuous and never-extinct interests of young generations in board games [8] are testimony to this. Therefore, instead of offering a digital app, we provide a palm-size gadget as the tangible interface which can facilitate multiple players playing face to face closely. Players can gather around (either by a

table, in a bar or lying on sofa) naturally to play with the buttons, read and pass on the printed prompts. We believe the physicality of our design could mediate socializing needs and afford more interactions between players.

5 USER STUDY

With the main goal of validating the design assumptions during the design process, mainly to evaluate generated prompts from real players perspective, understand the tangible design's effect, we conducted a preliminary user study with 3 groups (3-6 players each) in different scenarios (Fig 7). Overall, participants' age ranged from 21 to 52 years old, with a mean of 27.8, a standard deviation of 10.33, all with prior experience playing Truth or Dare.

Since our priority is to test players' authentic, instantaneous reactions and feedback in a real gaming environment, we simply first let players play the game using Trouble Maker without disruption, with prior notice that prompts are generated by ML. Group A (3 players) played during a break between classes, with eighteen prompts generated. Group B (4 players) played at a dinner gathering among friends with eight prompts generated and group C (5 players) played at a house party among friends and newly met people with fifteen prompts generated.

After the games, we did a survey on all twelve participants to better understand prompts, we asked players to rate on each prompt based on different parameters (Fig 8). We first ask the players whether the generated prompt is readable or not. We found that only rating the readability is not sufficient enough for a smooth gaming experience. For instance, "Put your left hand on your back and extend your right leg to meet your left hand" or "Start lying on your back, legs above your head, and arms extended in front of you" are both readable but would be impossible for players to carry out. Therefore, a 'task viability' rating is needed after the readability layer, which would determine whether the prompt is able to be executed or not by actual players. Finally, We also let players rate the degree of surprise using a 5-point Likert scale as an optional third question.

Results (Fig 9) show an average of 94.5%, 93.75%, 86.7% (three groups respectively) in readability, showing highly readable sentences were generated in real-time. Viable task prompts average to of 88.8%, 62.5%, 66.7% in the total number of tasks. The results also show that most players were surprised by the prompts.

91 Finally, we had open discussions after each game. During this qualitative interview, A3 said, "I'm amazed to see (that a) machine created those sentences that we humans could never think of." A2 commented "I feel weird to do those... (Calling a stranger to buy them)", showing the controversial interpretations of the generated prompts. B3 described that "Everyone holding their breath waiting for the printing process actually makes the whole gaming experience more mysterious", which proved our expectation of bringing a sense of ritual to this game using real-time physical input/output. Additionally, C3 mentioned that compared to traditional truth or dare game, being able to select from the five different modes (originally two) put him in a more comfortable zone where he only needed to sip "dark drink" instead of possibly "kissing someone he doesn't even know well". "As an introverted person, it's really tough for me to do those awkward things," said C3. By adding more modes to the game, we are able to better target practical scenarios and give players more freedom in the game.

6 DISCUSSION & FUTURE WORK

Here, we draw from our preliminary user study with some findings, discuss the possible reasons behind the data and how we can improve our system.

Based on the results from the user study, readability of prompts is relatively stable (standard deviation 0, 0.43, 0 respectively), meaning we are able to deliver mostly readable prompts through a series of tuning processes. On the contrary, we noticed that the deviation of viability of prompts varies a lot (0.94, 1.60, 1.41 respectively). As the criteria for judging whether a prompt is viable is relatively subjective and dependent on scenarios. For instance, "Put your underwear into fridge for three minutes" might be considered either viable and not viable by different individuals and is also relevant to gaming environments (play with strangers or close friends). It is also hard to tell if the result from our preliminary user study is valuable due to the small groups of players participated. Therefore, a more comprehensive user study is needed with: 1) a much larger quantity of prompts, 2) a bigger group of participants in order to fairly categorize and analyze output prompts.

In addition, we noticed the ethical concerns regarding the algorithm as Trouble Maker sometimes creatively encourages players to do risky and controversial things, which could cause offense to others, endanger the player themselves or other players physically, or even create trauma to players psychologically, etc. However, there is a tradeoff between the key element of this game, surprise, and those ethical problems [26]. Such issues may not immediately be resolved [26] but can only be better balanced through algorithm control and constraints from the game design side, as ultimately players decide how and whether to play in the game [25]. However, we still encourage players to get consent from all involved players prior to each prompt to allow a comfortable gaming environment for all.

We are also interested in the game captivation mechanism and how ML can help encourage such a process in our case. For now, knowing the fact that those prompts are generated by ML could be an influential factor that attracts and captivates players, and such freshness could disappear once they played for a few rounds. Therefore, whether players can notice the difference between traditional human-generated prompts and our ML-generated prompts and whether generated prompts can better captivate players, encourage them to return to the game in the long term need to be studied.

Although the generated prompts have a high readability rate after a series of tuning processes, the sentences are still sometimes repetitive in structure, for example: 1) call a library and ask them if they have pizza. 2) call a stranger and ask if you can buy them. Here, even though the content themselves is irrelevant, it may make players lose interest as they keep "calling strangers". Therefore, we plan to further enlarge the input text data source to reduce the rate of getting similar prompts.

Another phenomenon we observed during the user study is that by autonomously generating prompts using ML, we potentially deprived some involvement of human players from part of the game,

Figure 7: a) Mime challenge mode. b) Dark drink mode. c) Anti-human mode.

Figure 8: User study flow to rate prompts

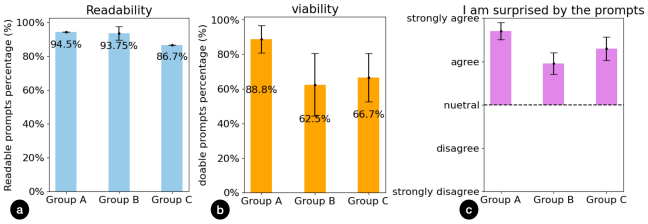

Figure 9: a) Evaluate readable prompts, b) Evaluate viable prompts, c) Rate the degree of surprise by the prompts

such as thinking up more tailored, intimate questions targeting a specific friend or follow-up (e.g. 'revenge') dares. This could be addressed by allowing customizable prompts. Players will be able to add real-time input and the algorithm will mix it with other text data while actual conversations happen, thus enabling human-AI co-creation with a subtle mixture of known and unknown elements in a prompt.

7 CONCLUSION

Recent trends of applications for ML are mainly adopting those large data (e.g. ChatGPT3 [9] was trained with 40Tb-80 Tb text data) based language models (LLM). Instead, in our work, we adopt a lightweight language model (89 Kb input text data) that requires light computing power to autonomously generate Truth or dare prompts for players. We enclosed the controller (with onboard generator algorithm, selection display, button), thermal printer in a tangible interactive device. According to our preliminary user study, prompts generated by Trouble Maker are relatively viable and mostly surprising, not only enhancing the experience of the Truth or Dare, but could also open up new meanings and extra playfulness to the original game.

We believe that our system is simple, elegant yet sufficient for the game Truth or Dare, and we hope our design can inspire new and playful experiences and interactions to human-AI interactions in games design.

ACKNOWLEDGMENTS

We want to thank Prof. Shirley Leung for guiding this work and providing technical support. We want to thank Prof. Colleen Macklin, Prof. Jessie Harding for participating in the user study and provide invaluable feedback.

REFERENCES

[1] [n. d.]. *"A New Dimension in Tabletop Gaming." The Last Gameboard.* Retrieved April 14, 2023 from https://lastgameboard.com
[2] 2005. *Puppeteer.* Retrieved April 17, 2023 from https://pptr.dev/
[3] 2009. *Node.js.* Retrieved April 17, 2023 from https://nodejs.org/en
[4] 2011. *Minecraft.* Retrieved April 17, 2023 from https://www.minecraft.net
[5] 2016. *no man's sky.* Retrieved June 22, 2023 from https://www.nomanssky.com/
[6] 2019. *Ai dungeon.* Retrieved June 22, 2023 from https://play.aidungeon.io
[7] 2019. *Better language models and their implications.* Retrieved June 22, 2023 from https://openai.com/research/better-language-models
[8] 2020. *The Rise of Board Games in Today's Tech-dominated Culture.* Retrieved Aug 24, 2023 from https://www.pitt.edu/pittwire/features-articles/rise-board-games-today-s-tech-dominated-culture
[9] 2022. *chatgpt3.* Retrieved June 22, 2023 from https://chat.openai.com/auth/login
[10] 2022. *midjourney.* Retrieved July 30, 2023 from https://www.midjourney.com/home/
[11] 2022. *troublemaker.* Retrieved June 22, 2023 from https://github.com/liangyq404/trouble_maker/tree/local
[12] AIBrain. [n. d.]. *Futurable: True Ai Game.* Retrieved June 22, 2023 from https://aibrain.com/futurable-true-ai-game/
[13] Berg. [n. d.]. *Designed in Hackney: Little Printer.* Retrieved April 17, 2023 from https://www.dezeen.com/2012/03/20/designed-in-hackney-little-printerby-berg
[14] Henri Berthiaux, Vadim Mizonov, and Vladimir Zhukov. 2005. Application of the theory of Markov chains to model different processes in particle technology. *Powder Technology* 157, 1-3 (sep 2005), 128–137. https://doi.org/10.1016/j.powtec.2005.05.019
[15] Olivier Brcknr. [n. d.]. *Designed in Hackney: Little Printer.* Retrieved April 17, 2023 from https://little-big-printer-messenger.vercel.app
[16] Jianxun Chu, Sara Qaisar, Zakir Shah, and Afsheen Jalil. 2021. Attention or Distraction? The Impact of Mobile Phone on Users' Psychological Well-Being. *Frontiers in Psychology* 12 (apr 2021). https://doi.org/10.3389/fpsyg.2021.612127
[17] Micael DaGraça. 2017. *How Artificial Intelligence has Shaped the History of Gaming.* Retrieved April 18, 2023 from https://www.gamedev.net/tutorials/programming/artificial-intelligence/how-artificial-intelligence-has-shaped-the-history-of-gaming-r4782/
[18] Google DeepMind. [n. d.]. *AlphaGo.* Retrieved June 22, 2023 from www.deepmind.com/research/highlighted-research/alphago.
[19] Goldman Jeremy. 2021. *A Return to in-Person Events? Here's What the Data Suggests.* Retrieved April 18, 2023 from http://ccrma.stanford.edu/~jos/bayes/bayes.html
[20] Jsvine. [n. d.]. *Markovify: A Simple, Extensible Markov Chain Generator.* Retrieved April 14, 2023 from https://github.com/jsvine/markovify

[21] Ville Kankainen and Heikki Tyni. 2014. Understanding Smart Device Tabletop Games. https://doi.org/10.1145/2676467.2676511
[22] Rachel Paul, Jesse Sharrard, and Song Xiong. 2016. The Importance of Face-to-Face Communication in the Digital World. *Journal of Nutrition Education and Behavior* 48, 10 (nov 2016), 681. https://doi.org/10.1016/j.jneb.2016.09.014
[23] Ticha Sethapakdi, Mackenzie Leake, Catalina Monsalve Rodriguez, Miranda J. Cai, and Stefanie Mueller. 2022. KineCAM. *Proceedings of the ACM on Computer Graphics and Interactive Techniques* 5, 4 (sep 2022), 1–9. https://doi.org/10.1145/3533613
[24] Noor Shaker, Julian Togelius, and Mark J. Nelson. 2016. *Procedural Content Generation in Games*. Springer International Publishing. https://doi.org/10.1007/978-3-319-42716-4
[25] Lucy Sparrow, Martin Gibbs, and Michael Arnold. 2019. Apathetic Villagers and the Trolls Who Love Them. In *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*. ACM. https://doi.org/10.1145/3369457.3369514
[26] Lucy A. Sparrow, Martin Gibbs, and Michael Arnold. 2021. The Ethics of Multiplayer Game Design and Community Management. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM. https://doi.org/10.1145/3411764.3445363
[27] EA sports. [n. d.]. *FIFA companion app - FIFA 23*. Retrieved June 22, 2023 from https://www.ea.com/en-gb/games/fifa/ultimate-team/fut-app
[28] Mike Treanor, Alexander Zook, Mirjam Palosaari Eladhari, Julian Togelius, Gillian Smith, Michael Cook, Tommy Thompson, Brian Magerko, John Levine, and Adam Smith. 2015. AI-Based Game Design Patterns.
[29] Evert van Loenen. 2007. Entertaible: A Solution for Social Gaming Experiences. (01 2007).
[30] Otis Peterson Wehar and Michael. [n. d.]. *My Crossword Generation Tools*. Retrieved June 22, 2023 from https://www.crosswordconstruction.com/generator
[31] Noah White. [n. d.]. *DrinkMaker: A Markov-Chain Based Script That Makes Random Mixed Drinks*. Retrieved April 17, 2023 from https://github.com/NoahWhite1115/DrinkMaker
[32] Jeff Wigh. [n. d.]. Chessup: Chess Smart Board for All Ages. Retrieved April 18, 2023 from https://playchessup.com/

Received 2023-06-22; accepted 2023-08-03; revised 2023-06-22; accepted 2023-08-03