Talk Like Somebody is Watching: Understanding and Supporting Novice Live Streamers

Terrance Mok
terrance.mok@ucalgary.ca
University of Calgary
Calgary, AB, Canada

Colin Au Yeung
colin.auyeung@ucalgary.ca
University of Calgary
Calgary, AB, Canada

Anthony Tang
tonytang@utoronto.ca
University of Toronto
Toronto, ON, Canada

Lora Oehlberg
lora.oehlberg@ucalgary.ca
University of Calgary
Calgary, AB, Canada

ABSTRACT

We built a chatbot system–Audience Bot–that simulates an audience for novice live streamers to engage with while streaming. New live streamers on platforms like Twitch are expected to perform and talk to themselves, even while no one is watching. We ran an observational lab study on how Audience Bot assists novice live streamers as they acclimate to multitasking–simultaneously playing a video game while performing for a (simulated) audience.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

Live streams; Streamers; Chatbot; Audience; Virtual audience

1 INTRODUCTION

Online video streaming is an increasingly popular form of entertainment, education, and community building. It offers audiences a place to participate collectively around a variety
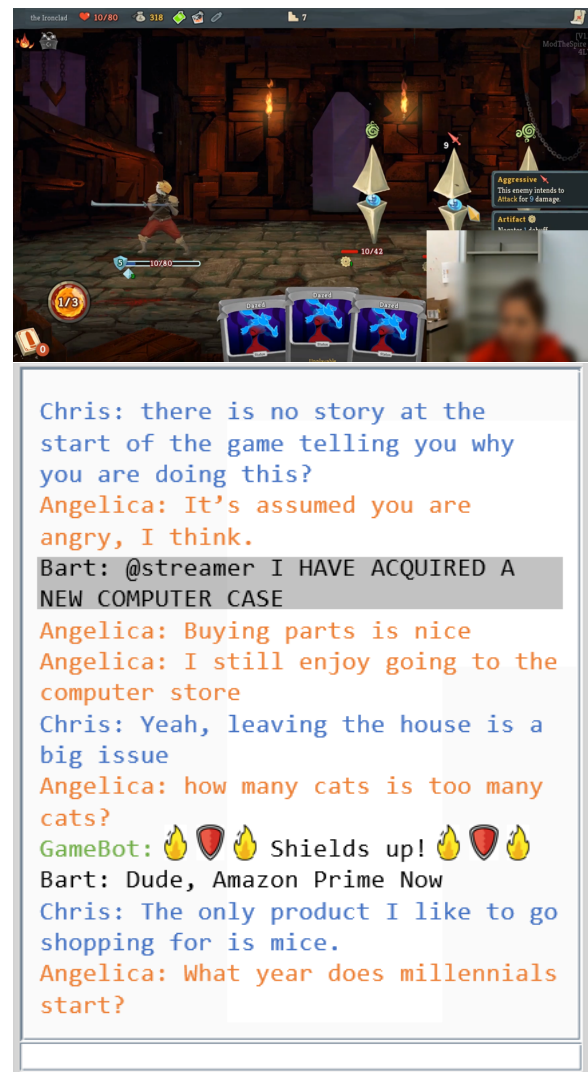
Figure 1: Participant on webcam performing a mock live streaming task while playing Slay the Spire with the help of Audience Bot, a simulated audience text chat.

of content [1, 7, 10]. Popular Twitch streamers have hundreds or thousands of viewers, a portion of which actively participate in text chat. However, most first-time streamers begin with an empty room [8]. To develop an audience, new streamers must learn to simultaneously be engaging, respond to messages in chat, and play their game.

Common advice from online guides for new streamers include suggested discussion topics (e.g., hobbies), ways to talk (e.g., being energetic), and how to engage (e.g., asking the audience questions) during a live streaming performance. Beginners are encouraged to be entertaining and talkative while alone in their stream, behaving as if they had viewers even if nobody is watching.

We developed Audience Bot (see Figure 1), a prototype chatbot design probe to help us understand how novice streamers could develop live streaming talk-aloud skills while playing a video game. Audience Bot uses log data from the game along with a curated selection of archived Twitch chat logs to interject prompts into the chat, encouraging the streamer to talk.

We conducted a small observational lab study with five participants with no streaming experience. We asked our novice live streamers to play a game (Slay the Spire shown top of Figure 1) while talking aloud, both with and without Audience Bot. Compared to having no audience, participants made more non-game related remarks while talking to Audience Bot and enjoyed their live streaming experience with Audience Bot. However, they still found it difficult to simultaneously manage reading and responding to chat while playing the game.

We offer two main contributions. First, we contribute Audience Bot, a system that leverages video game data and archived Twitch chats to provide a simulated audience. Second, we discuss initial observations of the novice streamer experiences and challenges, a topic that has thus far received little attention from the HCI research community.

2 RELATED WORK

Live Stream Experience

Live streaming research has focused on the experiences of established live stream channels with an interactive and engaged audience. There are many motivations for producing a live stream, such as learning to program [4], sharing crafts [6], teaching [3], community engagement [9, 14], or general social activity [6, 9]. All of these motivations rely on the existence of an interactive audience. While nearly all live streamers begin with no audience, very little research thus far has examined the experience or challenges of live streaming to an empty audience.

Woodcock and Johnson [13] as well as Taylor [12] highlight the intensive labor of maintaining a continuous performance during live streaming. Woodcock and Johnson elaborate, "Streamers must be constantly 'on,' friendly, or witty and are required to remain in character for long periods" and "that live streamers feel the pressures of such labor" [13]. Our work asks how novice live streamers undertake this labor, particularly during live streams without an audience to perform for.

Virtual Audiences

Our approach is to introduce a virtual audience into live streaming as a training tool for effective live stream behaviours. Simulated virtual audiences have been used to support public speaking, providing feedback through non-verbal means such as expressions [2] and gestures [11]. These systems create an environment for training public speaking skills "offline". While public speaking is one component of a live stream, streaming also requires the performer to acclimate to being on camera, performing a simultaneous task (such as playing a video game), and reading and reacting to text-based communication from their audience.

3 AUDIENCE BOT DESIGN AND IMPLEMENTATION

Audience Bot is a chat bot that simulates a small group (1-3 people) text chat from a Twitch streaming channel. We designed Audience Bot as an offline training tool to practise live streaming performance. Audience Bot generates two types of simulated audience chatter. First, Audience Bot responds to in-game actions and events, such as the player choosing a card to play. Second, it provides a scripted chat message dialogue based on real-world Twitch chat logs.

The Audience Bot implementation consists of several components (Figure 2). We modified Chatty[1] as a front-end for displaying simulated audience messages. We also chose to build Audience Bot around a specific video game that we could modify to generate chat content: Slay the Spire.

Messages consist of two components: a username and message content. Usernames are text-only characters; message content contains text, emoji, and Twitch emote images. The Chatty window shows the 20 most recent lines of dialogue.

Video Game Choice: Slay the Spire

Live streaming platforms support a wide range of games ranging from slow turn-based strategy games to fast paced real-time first-person shooters. We developed Audience Bot around Slay the Spire (top of Figure 1), a popular single-player turn-based strategy card game for several reasons. First, the game is turn-based which relieves new streamers

---

[1]https://chatty.github.io
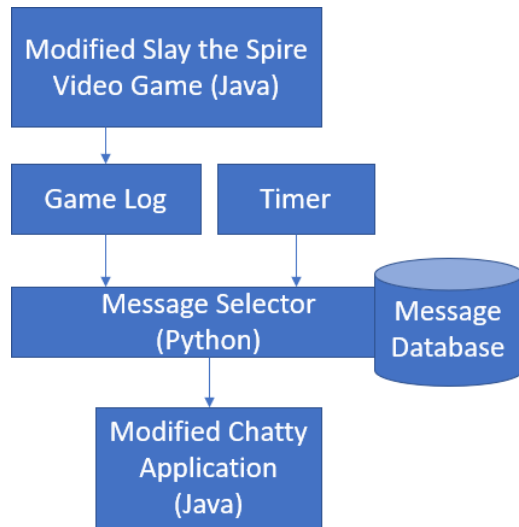
Figure 2: Audience Bot generates messages both by 1) responding to a log of Slay the Spire and 2) a timer. Messages are selected from a database of messages and shown to the streamer in a Chatty window.

Figure 3: Study condition order and timing for all participants. Participants performed the task conditions alone without researchers present.

from time pressure while playing. Second, Slay the Spire also allows for custom modifications to its source. We wrote a Java *mod* to track events within the game and output them to a simple text file. Our mod let Audience Bot "listen" to in-game events and player interactions, and use this content to generate chat message in response to game actions.
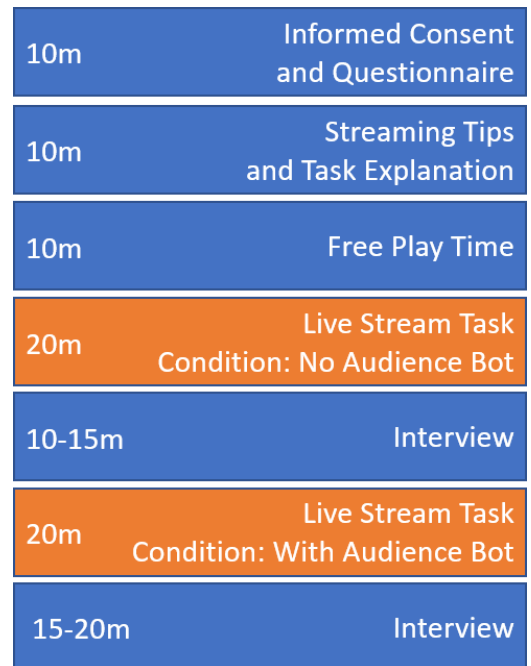
Design Choices

We developed Audience Bot's parameters–message rate, message content, and number of bots–based on feedback from iterative pilot testing.

Number of bots. We chose a combination of single bot and multiple bot interactions. Audience Bot starts with a single bot and slowly introduces a second and third bot. This gives the streamer time to adjust to having an audience. Through pilot testing we found that three bots were just enough to simulate conversation threads, while keeping the audience size limited for a novice streamer.

Message Rate. The rate of chat messages on Twitch varies based on number of viewers in the audience. Ford et al. [5] found massive online audiences (over 10,000 concurrent viewers) produced chat text as quickly as 1.74 lines/second. For our novice live streamers, we simulated a small audience, and thus limited our message rate to 1 message every 10 to 20 seconds.

Message Content. We modeled our messages on real Twitch chats, by collecting chat logs from Twitch Videos On Demand (VOD)–archives of live content that contain both video and corresponding chat. We retrieved the top 10 VODs (ranked by

view count) for Slay the Spire, resulting in a sample of 38,901 messages. We randomly selected sequences of messages from our sample, and looked for messages that could be appropriately adapted to Audience Bot. For instance, we excluded messages that were specific to something the streamer said or did in-game. We created a curated set of 200 messages (example messages in Table 1).

Two co-authors manually created "Game Action Response" messages, generated from in-game actions. Audience Bot responds to in-game events including: cards being picked for the deck, cards being drawn by the players, important cards being played, monster encounters and other room encounters. When one of these events happens, Audience Bot makes a short interjection, regardless of other chat, using game event related comments from our database. The bot is a simple system and can repeat messages over the task period. However, the slow nature of the game meant this only happened occasionally.

4 STUDY

We ran a small observational lab study to understand some of the challenges of streaming with no audience, and to observe whether a simulated text-chat audience with Audience Bot could help new streamers develop stream performance skills. We summarize our study design in Figure 3.

Type	Message
Game Action Response	Clash might be hard to use sometimes
Game Action Response	That's a lot of cards
Game Action Response	There's no way that Hemokinesis is a real word right?
Game Action Response	GIVE ME STRENGTH AHHHHHHH
Dialogue	When you pour milk into cereal, would you consider it a beverage, a sauce or a broth?
Dialogue	tomatoes are the best
Dialogue	Koalas are horrible animals. They have one of the smallest brain to body ratios of any mammal, additionally - their brains are smooth. A brain is folded to increase the surface area for neurons. If you present a koala with leaves plucked from a branch, laid on a flat surface, the koala will not recognise it as food. They are too thick to adapt their feeding behaviour to cope with change. In a room full of potential food, they can literally starve to death.
Dialogue Addressed to Streamer	@streamer remember to drink plenty of water to stay hydrated
Dialogue Addressed to Streamer	P.S. Remember not to die @streamer
Dialogue Addressed to Streamer	oooh hey @streamer any games you looking forward to?

Table 1: Example Audience Bot messages

Participants

We recruited participants through university campus posters and snowball sampling, asking potential participants to complete an online screening questionnaire. Out of 15 responses to our eligibility survey, we recruited five participants, age 21-56, (3 male, 2 female) who had some video game experience, as well as little or no streaming experience. We provided $20 remuneration for participating in our study.

Study Setup

Participants completed a brief questionnaire to gather demographic information and self-assess their experience with video games, live streaming, and public speaking.

We then introduced the idea of performing as a live streamer to participants, and discussed seven tips for talking while streaming, based on common online advice:

(1) Keep Talking: About Anything
(2) Verbalize Your Thoughts
(3) Commentate Your Gameplay
(4) Give Opinions About the Game
(5) Talk About Topics You're Interested In
(6) Share Stories
(7) Ask Questions

Our seven tips summarize live streaming advice from online sources; we reviewed the top 5 YouTube video tutorial results for "new streamer tips", and the top 10 online guides for new streamers from a Google search for "new streamer tips". For each piece of advice, we provided examples. For "Share Stories", we suggested that streamers might talk about their activities on the prior weekend. We answered any participant questions about the tips or streaming.

We then gave participants 10 minutes to become familiar with the Slay the Spire game. During this time we answered participant questions about the game, live streaming, or the upcoming task.

Mock Live Streaming Task

To introduced the study task, we described how new live streamers need to perform to any potential viewers that might drop into their audience. We asked participants to imagine they were live broadcasting their game play publicly, and asked them to attempt a live streamer performance–simultaneously playing the game while talking, entertaining, and engaging with an audience. Participants' work station was set up with a view of Slay the Spire, a webcam, and a chat room (see Figure 4).

Participants performed a 20-minute Mock live streaming task under two conditions–*No Audience Bot* and *With Audience Bot*–with an 10-20 minute interview after each condition. For *No Audience Bot*, participants performed the task without Audience Bot, and the Chatty application window stayed empty. For *With Audience Bot*, participants were informed that they would not be streamed to a live audience and would instead perform with Chatty display messages generated by Audience Bot.

Figure 4: Study participant streamer task setup. The left monitor contains only gameplay. The right monitor displays Chatty, Open Broadcaster Software (OBS) showing a live feed of gameplay and the participant on webcam, and our seven streaming tips.

Data Collection and Analysis

Participants completed short questionnaires before the live streaming task. We screen recorded the mock live stream setup, which included a synchronized webcam view of the streamer's face, a capture of the game, as well as the current state of the chat window. We used an additional microphone to record participants' dialogue during the task and interviews.

For our analysis, two co-authors watched and took notes on participants' mock live streaming performances, and transcribed all participant interviews. We also measured the duration of "dead-air" (i.e., silence) of participants during the streaming task.

5 FINDINGS AND OBSERVATIONS

Participants' verbal performance was noticeably different when Audience Bot was present. When performing without Audience Bot, all participants focused on the game as they played, rarely heeding the latter half of our list of streaming advice related to topics outside the game. Alternatively, when Audience Bot was active, all participants used Audience Bot as a source of ideas and topics for their performance.

Verbal Shadowing

In both conditions, all participants defaulted to describing the actions and visuals readily apparent in the game, reading game text aloud or simply verbalizing an event that just occurred. For example, P1 verbalized the count of damage being done by a monster (*"5 damage, 3 damage, okay"*) as the corresponding numbers showed in the game. In another case, P5 noted, "*Again, I got my cards*" after receiving their cards for the turn; P5 did not add any additional discussion of what cards they received or which cards they were considering

Table 2: P4 Responds to Audience Bot

Source	Content
Audience Bot Chris [*Text*]	people who stay at the same job for 40 years are crazy
Audience Bot Angelica [*Text*]	both my parents did that, but doesn't work in this day and age
P5 [*Verbal*]	"No, not at all. I love my job. I love my job a lot. I can't even think of a day without my job. I hope to work as long as I am physically fit."

playing. Participants' flat toned, sans-affect speech filled the silence, but this style of basic description–repeating what can easily be seen on-screen–added little to the live stream performance that was not otherwise visually apparent to an audience member.

We note that simple descriptions of mundane game occurrences was expected; one of the streaming tips provided to participants encouraged them to talk as much as possible (*"about anything"*), but did not specify what to talk about. Playing a new video game for the first time and performing for a live streaming for the first time leaves little cognitive energy to form and verbalize deeper thoughts. Thus, simply describing the game seemed to be the easiest way for participants to heed our advice to continuously verbalize.

Talking with Audience Bot

In the *With Audience Bot* condition, participants took time to read and react to Audience Bot messages, and respond to topics brought up by Audience Bot. The simulated audience introduced opportunities for participants to discuss topics unrelated to the game. In the *No Audience Bot* condition, topics beyond immediate game play were mentioned in only 18 instances across all participants' performances; P1 never discussed any topics beyond the game in the *No Audience Bot* condition. This is in stark contrast from the *With Audience Bot* performances, where participants discussed topics outside the game in 72 instances. With Audience Bot, P4 made 14 non-game related remarks and discussed topics ranging from their sleep schedule, chicken burgers, plans to go shopping, and love for their job (see Table 2).

The structure, formatting, and content of Audience Bot messages influenced participants' responses.

Expressive Commentary. Audience messages using emojis to show an emotional response to gameplay were largely ignored by participants. For instance, Audience Bot sent the

Table 3: Participant silence and non-game remarks in each task condition

	Silence Duration *(s)*		Non-game Remarks	
Participant	No Bot	Bot	No Bot	Bot
P1	871	1000	0	15
P2	468	537	3	12
P3	932	935	5	15
P4	702	717	4	14
P5	278	222	6	16

message "🔥 🌶️ 🔥 Shields up! 🔥 🌶️ 🔥" each time the card *Flame Shield* was played. However, none of our participants responded to these messages. While this type of message is based off of actual Twitch chat behaviour and lends authenticity, they did not necessarily prompt participants to engage with the chat.

Message Size. Audience Bot messages are 11 words long on average. However, our samples longest message about koalas (see Table 1) grabbed all our participants' attention. Participants read some or all of the message aloud and reacted verbally or with a laugh. After reading the entire message P3 responded: "*Woah, that's something I didn't know.*" The message stands out due to its size compared to other messages–it grabs attention because it introduces a large block of text to the chat display. While the odd content may also play a factor, we saw similarly strange content in shorter messages go unnoticed.

Dead-Air

We had hoped Audience Bot would reduce the duration of dead-air (i.e., silence) during a novice live stream performance. However, Audience Bot demanded more attention, as participants not only needed to play and talk at the same time, but also read, understand, and develop responses to the Audience Bot text. This was evident from both participant's performances and from post-task interviews. Participants exhibited more silence in the *With Audience Bot* condition, with the exception of P5 (see Table 3). For instance, when asked whether Audience Bot changed how they interacted with the game, P1 reflected, "*I found I was pausing a lot more so I could actually read the chat.*"

Enjoying the Virtual Audience

In post-task interviews four participants (P2-P5) said they liked having Audience Bot present (P4: "*I enjoyed it a lot*"). However, P1 did not necessarily enjoy Audience Bot; while Audience Bot was "*accurate of the Twitch experience*", they "*didn't really enjoy it but it's basically an educational tool. You're not supposed to enjoy it.*"

6 DISCUSSION

Experience and Multitasking

As participants read and responded to Audience Bot messages in chat, they suspended play. While this slowed game progress, we intentionally selected a turn-taking game so that pauses in game play would not be an immediate detriment to game performance. Other game genres would have been more challenging: Games with real-time elements demand more constant attention and action, while multi-player experiences require players to be aware of teammates' in-game actions and to listen for verbal communication.

Our participants were novices to both live streaming and to the game, which likely compounded the challenges of multitasking. We expect that as a live streamer gains experience with either playing the game or reading and responding to their audience, multitasking across both activities simultaneously would become easier.

Message Rate

We used a slower message rate to simulate a small audience, which is more typical for first-time streamers. However, popular streamers have on occasion sent their viewers to watch a smaller streamer, causing a sudden flood of viewers [4, 12]. Future work could expose novice live streamers to faster chats to simulate larger audiences.

The Value of Streaming to Nobody

Is there there value to the experience of streaming to nobody that is lost when using Audience Bot? Audience Bot is premised on the value of having an audience; however it is possible that performing for nobody has benefits for a streamer's growth. Further interviews with streamers who perform without an audience could provide implications for support tools beyond simulating an audience.

7 CONCLUSIONS

We highlight several difficulties facing novice live streamers through a small observational lab study of our prototype simulated audience system–Audience Bot. For our participants, it was challenging to discuss non-game related topics without an audience prompting them with topics. However, Audience Bot increased multitasking demands by the need to read, respond, and play simultaneously. Our virtual audience is a promising support tool for novice streamers; future work is needed to explore how adding bots impacts play, workload, and learning of streaming skills.

8 ACKNOWLEDGEMENTS

REFERENCES

[1] Laurensia Anjani, Terrance Mok, Anthony Tang, Lora Oehlberg, and Goh Wooi Boon. 2020. Why do people watch others eat? An empirical study on the motivations and practices of mukbang viewers. In *CHI 2020: Proceedings of the 2020 SIGCHI Conference on Human Factors in Computing Systems*. 10 pages.

[2] Mathieu Chollet, Pranav Ghate, Catherine Neubauer, and Stefan Scherer. 2018. Influence of Individual Differences when Training Public Speaking with Virtual Audiences. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*. ACM, 1–7. https://doi.org/10.1145/3267851.3267874

[3] Travis Faas, Lynn Dombrowski, Erin Brady, and Andrew Miller. 2019. Looking for Group: Live Streaming Programming for Small Audiences. In *International Conference on Information*. Springer, 117–123. https://doi.org/10.1007/978-3-030-15742-5_10

[4] Travis Faas, Lynn Dombrowski, Alyson Young, and Andrew D. Miller. 2018. Watch Me Code: Programming Mentorship Communities on Twitch.Tv. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 50 (Nov. 2018), 18 pages. https://doi.org/10.1145/3274319

[5] Colin Ford, Dan Gardner, Leah Elaine Horgan, Calvin Liu, A. M. Tsaasan, Bonnie Nardi, and Jordan Rickman. 2017. Chat Speed OP PogChamp: Practices of Coherence in Massive Twitch Chat. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI EA '17)*. ACM, New York, NY, USA, 858–871. https://doi.org/10.1145/3027063.3052765

[6] C. Ailie Fraser, Joy O. Kim, Alison Thornsberry, Scott Klemmer, and Mira Dontcheva. 2019. Sharing the Studio: How Creative Livestreaming Can Inspire, Educate, and Engage. In *Proceedings of the 2019 on Creativity and Cognition* (San Diego, CA, USA) *(C&C '19)*. ACM, New York, NY, USA, 144–155. https://doi.org/10.1145/3325480.3325485

[7] William A. Hamilton, Oliver Garretson, and Andruid Kerne. 2014. Streaming on Twitch: Fostering Participatory Communities of Play within Live Mixed Media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1315–1324. https://doi.org/10.1145/2556288.2557048

[8] Mehdi Kaytoue, Arlei Silva, Loïc Cerf, Wagner Meira, and Chedy Raïssi. 2012. Watch Me Playing, i Am a Professional: A First Study on Video Game Live Streaming. In *Proceedings of the 21st International Conference on World Wide Web* (Lyon, France) *(WWW '12 Companion)*. Association for Computing Machinery, New York, NY, USA, 1181–1188. https://doi.org/10.1145/2187980.2188259

[9] Max Sjöblom, Maria Törhönen, Juho Hamari, and Joseph Macey. 2019. The ingredients of Twitch streaming: Affordances of game streams. *Computers in Human Behavior* 92 (2019), 20–28. https://doi.org/10.1016/j.chb.2018.10.012

[10] David Struzek, Martin Dickel, Dave Randall, and Claudia Müller. 2019. How Live Streaming Church Services Promotes Social Participation in Rural Areas. *Interactions* 27, 1 (Dec. 2019), 64–69. https://doi.org/10.1145/3373263

[11] Hiroki Tanaka, Sakriani Sakti, Graham Neubig, Tomoki Toda, Hideki Negoro, Hidemi Iwasaka, and Satoshi Nakamura. 2015. Automated Social Skills Trainer. In *Proceedings of the 20th International Conference on Intelligent User Interfaces* (Atlanta, Georgia, USA) *(IUI '15)*. ACM, New York, NY, USA, 17–27. https://doi.org/10.1145/2678025.2701368

[12] TL Taylor. 2018. *Watch me play: Twitch and the rise of game live streaming*. Vol. 24. Princeton University Press.

[13] Jamie Woodcock and Mark R Johnson. 2019. The Affective Labor and Performance of Live Streaming on Twitch.tv. *Television & New Media* (2019), 1527476419851077. https://doi.org/10.1177/1527476419851077

[14] Tim Wulf, Frank M Schneider, and Stefan Beckert. 2018. Watching players: An exploration of media enjoyment on Twitch. *Games and Culture* (2018). https://doi.org/10.1177/1555412018788161