Improvising with an Audience-Controlled Robot Performer

Claire Mikalauskas, Tiffany Wun, Kevin Ta, Joshua Horacsek, Lora Oehlberg

University of Calgary

Calgary, AB Canada

{crmikala, twwun, kta, joshua.horacsek, lora.oehlberg}@ucalgary.ca

ABSTRACT

In improvisational theatre (improv), actors perform unscripted scenes together, collectively creating a narrative. Audience suggestions introduce randomness and build audience engagement, but can be challenging to mediate at scale. We present Robot Improv Puppet Theatre (RIPT), which includes a performance robot (Pokey) who performs gestures and dialogue in short-form improv scenes based on audience input from a mobile interface. We evaluated RIPT in several initial informal performances, and in a rehearsal with seven professional improvisers. The improvisers noted how audience prompts can have a big impact on the scene — highlighting the delicate balance between ambiguity and constraints in improv. The open structure of RIPT performances allows for multiple interpretations of how to perform with Pokey, including one-on-one conversations or multi-performer scenes. While Pokey lacks key qualities of a good improviser, improvisers found his serendipitous dialogue and gestures particularly rewarding.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords

Improvised Theatre; Human-Robot Interaction; Crowdsourcing; Creativity-Support Tools
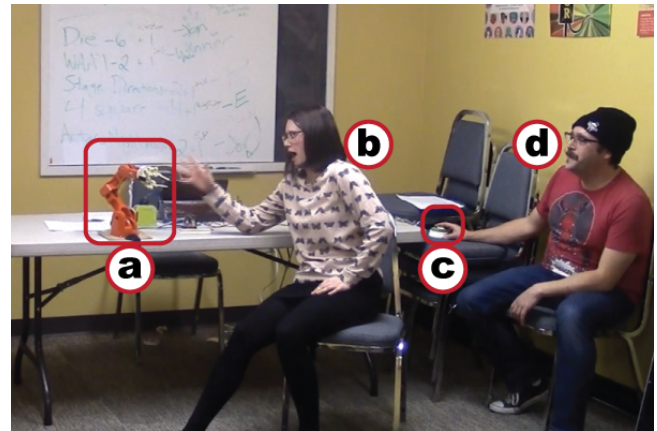
INTRODUCTION

Improvised theatre (improv) gives actors the freedom to improvise the well-known or explore something new without a script, pursuing plot details of particular interest. One common practice is to gather audience suggestions, which creates engaging interaction between the audience and the actors on-stage [13]. However, the actor must find a balance between going with an obvious choice based on the suggestion, or introducing a change or conflict that is inspired by the suggestion but not directly related [13]. Audience suggestions are used in many different ways in improv games. *Ask-Fors* allow the whole audience to shout suggestions or can be used to connect personally with one random audience member. In *Audience*

Figure 1: Improvising a scene using RIPT, with (a) Pokey, (b) an improviser, (c) the GO button, and (d) an off-stage button operator.

Speak [6], the audience must create one voice for a character, while *Moving Bodies* [6] brings audience volunteers on-stage to physically move the actors. However, there is a limit to the number of people who can meaningfully participate, or whose input becomes part of the show. For example, *Audience Speak* is most successful with a big audience, as the "funny" answers are drowned out by obvious answers. Improv is often not about being "funny" with every line, as the mundane suggestions are more relatable and playful on-stage [6].

Technology offers one possible means of mediating audience input, and improving the quality of audience participation in improvised performance. There are several open questions on how technology might address these issues:

- *How can technology increase audience engagement through more direct participation in an improvised performance?*

- *What are effective ways to solicit audience input for an improvised scene?*

- *How do performers integrate large amounts of audience input into their improvised performance?*

We introduce Robot Improv Puppet Theatre (RIPT) (Figure 1), an improv experience where one or more improvisers perform with an Arduino Braccio robot arm, Pokey. The robot arm performs a set of pre-recorded gestures and "mouth movement" actions with its gripper, while using text-to-speech to read aloud random audience input. RIPT solicits audience input

for both dialogue and gestures via a web-based application. The improviser creatively interprets the robot's dialogue and actions to construct new storylines. The robot's performance is directly based on audience participation, and offers improvisers unpredictable content to drive new, creative stories.

We offer the following research contributions: (a) the design and implementation of a system that allows an audience to control a robot actor in an improvised scene; (b) an evaluation of the aforementioned system with (i) informal performances with an amateur improviser, and (ii) a feedback session with seven professional improvisers as both performers and an "expert" audience; and (c) a discussion of how to manage audience input and how audience input impacts improvisational performance, when enacted by a performance robot.

BACKGROUND: IMPROVISED PERFORMANCE
Improv has many different styles and formats; for the purposes of this paper, we focus on *short-form improv*[1], which utilizes audience suggestions. Actors will often use a game or *ask-for* to begin their scene; the actors then respond to the rules or constraints of the game to create a new scene, on-the-fly. Key improv skills include carefully listening to fellow actors and introducing *callbacks* to previously discussed elements in the scene. The scene ends when one of the actors decides that the story has run its course or hits a strong punchline.

There are a range of ways to solicit the audience for suggestions. For example, *Bucket*[2] asks the audience to write words or phrases on slips of paper. Actors place the slips of paper in their pockets without looking at the contents; as needed throughout the scene, actors pull out a slip of paper to directly use as dialogue. Improvisers also draw "random" input from pre-existing bodies of text. In *Actor's Nightmare*[3], one actor has a script from which they must pull dialogue from one character's lines. The use of one character maintains the same perspective for the actor, while their partner contextualizes the seemingly random lines of dialogue into a cohesive scene. Games such as *Moving Bodies* rely on more active audience participation by bringing volunteers from the audience on-stage to directly manipulate the body positions of the actors during the scene [6]. For RIPT, we focus on direct forms of audience participation over thematic suggestions as the audience directly contributes dialogue and movement to the scene from one character's perspective.

Although most improv games do not currently use technology, the use of technical effects in improv is not unheard of. However, the use of advanced technology in improv is rare [3] and often limited to light and sound [14]. This type of control implies an improviser is in the tech booth of the theatre; in practice, improv can be performed anywhere. We focus on technologies that are lightweight and operate independently of existing on-set theatre tech; our goal is to focus on a mediated interaction between the performers and their audience.

---

[1] http://improvencyclopedia.org/glossary/Short_Form.html

[2] http://improvencyclopedia.org/games/Bucket.html

[3] http://improvencyclopedia.org/games/Actor%60s_Nightmare.html

RELATED WORK
In this section, we discuss past research on robot performance, and technology-mediated audience participation in performance.

Robots and Performance
Performing robots are often operated by a human performer or act out pre-programmed behavior. There are numerous examples of theatre incorporating robots in human-to-human performance. Puppeteering is one common approach, in which robots act as a surrogate for a human puppeteer. Some examples include using the robot as a marionette operator [7] or as a hand puppet through a VR headset [12]. We are interested in what happens when performers no longer have complete control over what the robot does.

Research has also explored how robots can effectively express emotion. Through human interpretation, robots can convey emotions and intent effectively. Movement alone can elicit emotional responses from an observer despite the lack of a face or other human characteristics [4]. By expressing forethought and reaction through movement, robots can make their character more appealing [18]. Furthermore, a human performer can elicit more lifelike interpretations from a robot by allowing the robot's actions to affect the narrative and build motives through its past actions [1, 8]. Our work focuses on the robot as a surrogate for crowdsourced audience input, whose motives and goals can be creatively interpreted by other actors to drive the story.

Performing with a robot provides different opportunities for the creation of an improvised show. Improvisational TheaterSpace [15] allows a human actor to perform with a virtual actor in the form of projected text, which understands simple words and phrases. The text is intended to be the inner thoughts of the actor. The system has since been used for improv, such as a person talking to their alter ego [16]. We focus on using both the robot's dialogue and physical action as openly interpretable material for improv.

Audience Participation in Performance
The HCI community has seen an increased interest in the role that online crowds can have in creativity practices — either as a source of ideas [20, 21] or constructive critique [2, 9]. Performance, however, is a fundamentally different setting for soliciting creative input from audiences. Every audience member brings their own individuality to their responses, but when put on the spot, they may feel pressured to self-edit and give a response they think the crowd will expect [6]. While there is no "right" answer for audience suggestions, improvisers want to elicit interesting, even personal, answers from the audience [13]. With any audience, there is a chance that an audience member is too embarrassed to provide suggestions [19]. Technology offers one way to mediate anonymous audience participation, protecting individual audience members from embarrassment and rewarding multiple audience members for their participation. Theatre Engine [11] gives the audience an opportunity to provide input to the performance through smart devices. The devices are used live throughout the performance, providing engagement by letting audience members

impact what occurs on-stage. However, Theatre Engine does not provide anonymity for the participants. With RIPT, we look at leveraging audience engagement, while maintaining anonymity in audience suggestions.

We are also interested in how performers adapt to technology on-stage. Honauer et al. [5] presented a scripted play to an audience, containing interactive special effects where the control is in the hands of the performers. The researchers studied how the performers adapt over the process, from rehearsals with the systems, fixes to the systems, and finally a performance. Although the audience is not in control of the interactive effects, understanding how they perceive technology is crucial to understanding its success on-stage. We extend actor-controlled technology to the stage of improvised theatre with RIPT. By gathering all input from the audience before the performance, rather than during, the input instead comes entirely from the audience.

POKEY, AN IMPROV ROBOT

Trisha arrives twenty minutes early to ensure she gets a good seat for viewing an improv show. When she walks into the lobby, she is greeted by one of the actors, and is instructed to go to a website on her phone to input dialogue and gestures for the upcoming show. Excited, Trisha pulls out her phone and visits the website. She first sees a prompt and asks the actor where the prompt is coming from. The actor explains that there is a set of prompts for each show, determined by someone who is not performing, in an attempt to get dialogue from the same perspective. Trisha asks who will be saying this dialogue, and learns that a robot arm will be performing on-stage as well. Trisha quickly enters some dialogue based on the prompt, and notices that she needs to also select a gesture. The names of the gestures are vague, but she is excited to see what it looks like on-stage. Trisha selects a gesture that seems the most interesting in combination with her dialogue. Once she hits submit, the screen confirms her submission and reloads with a new prompt and a new list of gestures. Trisha decides to submit as many entries as possible before she is asked to find her seat for the show. The show begins and Trisha is eager to hear her lines of dialogue delivered by the robot in the scene.

In the following section, we describe key RIPT system components: *Animating Pokey's Gestures and Movement* by defining and recording a library of possible gestures that the robot can perform with the "Gesture Animation Rig"; the *Audience Crowdsourcing Mobile Interface*; and the *Backstage Controller* that allows for minimal control by a behind-the-scenes "director".

Animating Pokey's Gestures and Movement
RIPT as a system facilitates the performance of dialogue and gesture by an Arduino Braccio Robot Arm ("Pokey"). First, we defined a series of gestures for Pokey to perform. To record gestures, we designed a custom animation rig that temporarily attaches to the exterior of the Braccio robot arm. The rig includes three lengths of 3D printed armatures and one base piece (Figure 2). Each joint of the animation rig has a potentiometer that adjusts its value according to the angle of the joint; as the animator moves the robot manually, the rig

Prompts
What you would say to...
• someone cooking in the kitchen
• a professional chef in a restaurant
• a friend in the kitchen with you
• a manicurist doing your nails
• a shop clerk selling you jewelry
• a friend while getting your nails done
• a sales clerk when buying business clothing
• your new boss at a part-time job
• your friend shopping with you for business clothing
• a lifeguard at the beach
• someone teaching you how to swim
• your spouse at the beach
• the police after witnessing a petty theft
• your sibling after witnessing a petty theft
• to show you are a law abiding citizen
• your child when trying to get them to go to sleep
• your dog when trying to calm them down
• a preteen you are babysitting
• a barista in a coffee shop who sees you everyday
• your sibling after not seeing them in years
• your best friend while quickly grabbing coffee
• a library volunteer
• your best friend while chatting on the phone
• someone in a whisper
• someone giving you a tattoo
• a doctor giving you a vaccination
• your favourite visual artist
• your mom while shopping for your prom dress
• a pushy sales associate who thinks you need help shopping
• a school chaperone for a high school event
• your sassy friend when you forgot something
• your sassy friend when you do something embarrassing
• your sassy friend about someone else when they do something

Table 1: List of prompt sets used in RIPT.

Gestures		
Idle*	Exaggerated Yes	Head Up
Confront	Defeated	Inspect
Nod Yes	Defensive	Jump
Laugh	Sulk	Dance
Raise Head Slowly	Shake No	Lean Back Quickly

Table 2: Table of Gestures. "Idle" gesture is not selectable in the mobile interface.

produces an analog value corresponding to the angle of each joint as a keyframe. Very few of our planned gestures used the yaw rotation about the base of the Braccio; thus, we did not design a jig to capture the yaw rotation of the base. Instead, gestures with yaw base rotation had each yaw angle manually defined for multiple keyframes.

To account for any misalignment of the potentiometers during the construction of the animation rig, we calibrated the translation from analog values to angles by first collecting the analog values while the arm is fully extended (as controlled by the provided Braccio library). We then reconstructed the affine map of analog values to angles using linear regression. This allowed us to disengage the servo motors and puppeteer the Braccio arm, recording the angles at each joint over time.

We recorded and implemented 18 gestures for Pokey (see Table 2). These include gestures such as: confronting the actor, inspecting something, or dancing. We recorded an idle animation for the robot to perform when there is no user-selected animation playing. This serves the two-fold purpose of giving the audience and performer a sense that the robot is "thinking", as well as informing the RIPT operator that all systems are working as expected.

In each recorded animation file, we manually placed cues to indicate both when the gripper should start moving for speech, and when the text-to-speech module should start producing sound. Mouth movements are not coordinated with the dialogue; the grippers simply open and close while talking and mouth movements stop once the text-to-speech process is terminated. Mouth movements operate independently from the arm gestures to accommodate different lengths of dialogue.

Audience Crowdsourcing Mobile Interface

Audience members access the crowdsourcing interface (Figure 3) by visiting a web URL on their mobile devices before the start of the show. The mobile interface prompts the audience for dialogue using one of three prompts from the selected prompt set. Dialogue entries need to be between 3-35

words and must be paired with a selected gesture. We created prompts that encourage a common character perspective with audience input for a show (see Table 1). The prompts come in sets of three, where each set is a show. The interface also asks the audience to select one of three gestures to go with their line of dialogue; the three options are selected at random from the full list of available gestures (see Table 2). The set of prompts and available gestures are chosen by a behind-the-scenes director using the backstage controller.

Backstage Controller

The backstage controller allows a director to initiate a show and collect submissions for Pokey. While the director can remove inappropriate submissions, they cannot use the back-
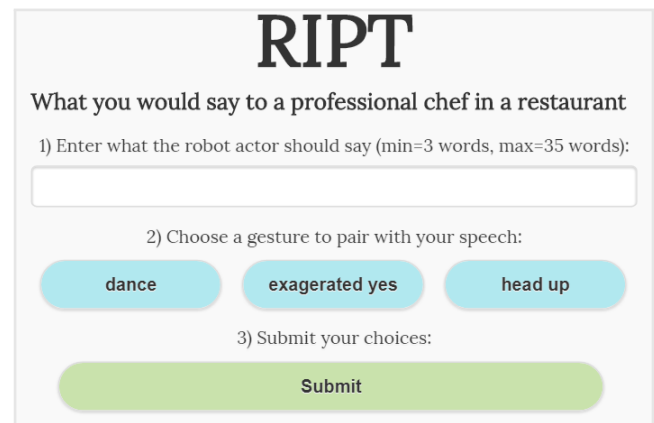
Figure 3: Audience Mobile Interface. One random prompt from a chosen set of prompts is presented per submission. The audience member supplies a line of dialogue paired with one of the three randomly presented gestures.

Figure 4: Backstage controller interface includes: (a) controls to initiate submission collection; (b) queue of next five submissions, to be used on upcoming GO button presses, chosen randomly from the collection; (c) the full list of submissions; and (d) manual controls for hardware buttons.

Figure 2: Animation rig. Armature is outlined in red and base piece outlined in blue. Each joint is fitted with a potentiometer to measure angles of the joint.

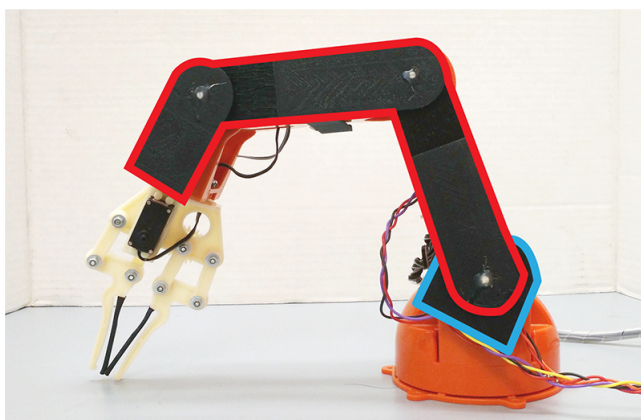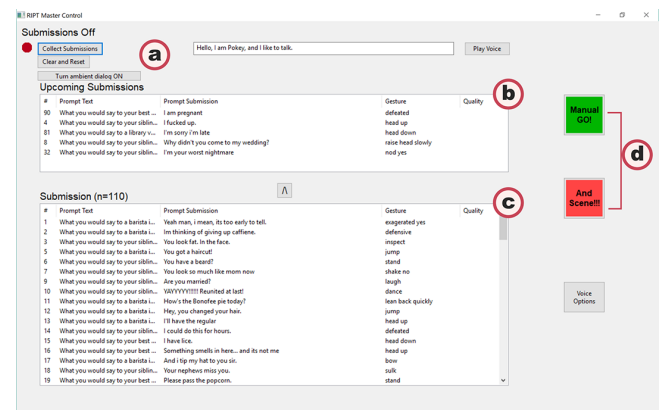stage controller to add submissions or control the ordering of those submissions. Once a submission is removed or played through Pokey, it can no longer be played again. To collect submissions, the director selects a set of prompts and gestures to be displayed on the mobile interface. A web-server then opens the mobile interface for submissions from the audience (Figure 4). Submissions are automatically saved to a file and can be reloaded for later. Once there is a satisfactory number of submissions, the director can close the submissions and begin the show.

The backstage controller and the mobile interface connect to each other via a proxy server. Each mobile interface connects to a server, which manages all mobile interface connections and relays submissions to and from the backstage controller. The backstage controller receives submissions and updates the server (therefore the mobile interface) with prompt sets, gestures, and submission collection status. The mobile interface was written in HTML5 and JavaScript, the server with Node.js and Socket.io, and the backstage controller with Python.

Initiating Pokey's Dialogue

Once the director signals that the show is ready to start, the show begins. An audience member, or one of the actors not currently in the scene, sits off-stage with the GO button. The button operator presses the GO button each time Pokey needs to deliver a line in the scene. A green LED provides feedback for when Pokey is ready to deliver another line; this prevents Pokey from playing another submission until the current one has finished executing. The performers interpret and respond to Pokey's gestures and dialogue to integrate the robot's performance into the scene. The scene ends whenever a performer, the director, or the button operator feels the scene has come to a natural end.

We rely on human intelligence to decide the timing of when Pokey should perform its lines. This not only simplifies the implementation of the system, but also keeps either an audience member or actor engaged in performing along with Pokey.

EVALUATION

To evaluate RIPT, we first conducted proof-of-concept performances both informally and at a demo session. However, for a more in-depth evaluation we brought RIPT to a rehearsal with professional improvisers.

Preliminary Evaluation

During the development of RIPT, we conducted several rounds of improvised puppet shows. One of our co-authors is an amateur improviser and performed scenes with Pokey using a sock puppet (see Figure 5). We initially used a puppet due to Pokey's size and movements resembling a sock puppet. Meanwhile, an audience member triggered Pokey's dialogue with the GO button. Our early performances led to usability improvements for the backstage and web interface. We made minor revisions to the prompts, robot voice, and gestures to improve the cohesion of the performance.

We also participated in a conference demo session at ACM UIST 2017, where a large number of people (appx. 200) passed by a booth featuring our co-author performer. We ran

four performances with Pokey during the demo session, and gathered audience contributions in the 20 minutes leading up to the show (avg. = 70). Each improvised show lasted approximately 5 minutes and used 15-20 submissions.

Naturally, performances with Pokey are unpredictable because of the randomness of the responses. While we intended for the prompts (Table 1) to guide the suggested dialogue, many audience members ignored the prompts by either inputting text copied from the internet or attempting to break the system with specific character strings. Although non-sequitur statements can be entertaining, submissions that followed the prompts provided more opportunities for a cohesive storyline. Ultimately, it is the performer's responsibility to make sense of what Pokey says and does.

In the demo session performances, the sock puppet limited the amount of emotion and movement the human performer could add to the show. Emotion can be crucial when bringing context and meaning to an improvised scene, yet Pokey's text-to-speech provided no emotion or tone. We decided to allow actors to use their bodies — instead of hand puppets — to allow them to fully emote during a scene. We brought our refined system and setup to a professional improv troupe to better understand how professional improvisers react to the audience-sourced experience. The improv troupe served as an "expert" audience, given their expertise with how audiences might behave in a performance, and as professional performers.

Professional Improvisers Evaluation

We brought RIPT to a rehearsal session with seven members from a local professional improv troupe. The actors (3 female, 4 male) had between 5 and 20 years of experience in improv, including experience in competitions, performing at corporate events, and working with multiple professional troupes.

We first demonstrated how RIPT uses the audience's input for dialogue and gestures. We stepped through how we selected the prompts and asked the actors to input dialogue as the audience for their own rehearsal performance. We demonstrated

Figure 5: Improvised puppet show performed by one of our co-authors using a sock puppet.

how the backstage controller could moderate input during a show; the actors agreed that it served a key purpose for public shows. In our study, we did not use a director to moderate submissions through the backstage controller (and informed our participants of this) as audience members themselves would take turns performing. Throughout the study, we asked the actors specifically about the usability and level of audience engagement with the prompts, gestures, and overall layout of the web interface. While they were adding dialogue, we asked them to speculate on how the audience experience of this type of performance might work out in practice. The semi-structured group interview format benefitted from the improvisers' retrospective and analytical tendencies stemming from their creative practice [10].

Once there was a sufficient amount of audience input, the actors took turns performing with Pokey, with one actor offstage to operate the button for each scene. The performers initially performed one-on-one scenes with Pokey, but eventually branched out to include multiple performers in the scene. In total, they performed seven scenes; each scene lasted approximately 3 minutes. Between scenes, the improvisers discussed how that scene went, and what worked (or did not work) with Pokey as a fellow actor; these discussions often informed how they experimented and played with Pokey in the following scene. At the end of the session, we gathered final feedback on what it is like to improvise with Pokey, and how Pokey's limitations affect the performer.

Our findings reflect actors' responses to our specific questions about the prompts, the backstage and web interface, and gestures. To understand the performance itself, we performed open coding [17] on videos of the rehearsal performances to identify emergent patterns.

FINDINGS

We discuss how RIPT solicits audience input and how a RIPT performance might be structured to coordinate audience input with the actors' performance. We then examine how improvisers crafted scenes using Pokey's random, audience-sourced dialogue and gestures. Throughout our findings we use pseudonyms to refer to our participants.

Soliciting Dialogue and Gesture from the Audience

Our improvisers were very familiar with improv performances that leverage audience input; finding ways to increase audience engagement through participation is a key challenge of their current practice.

Patty: When we do corporate shows, all of our games involve the audience, and that's to heighten the engagement of the audience. The audience loves nothing more than to hear themselves, center stage.

We asked the professional improvisers to provide feedback on how RIPT asked for audience input, given their extensive past experience with phrasing effective *ask-fors* more broadly in improv. The actors gravitated towards prompts that help the audience generate more interesting, personal answers.

Patty: When you're like "can we get a word?", you get 'dildo' or 'broccoli'. [...] When we're trying to get something a little

bit more sophisticated from an audience, we're often asking "Can I"... even just an adjective, an interesting word... "Can I get- What was the toy you played with as a child?"

Overall, the improvisers approved of the prompts' level of detail. They also appreciated the attempt to generate dialogue around a common perspective to identify a consistent role for Pokey in the scene: *"[The input is] like Actor's Nightmare, it's all coming from one perspective" (Alex).* They suggested alternate ways of prompting the audience to provide dialogue from a particular character or role, by including, *"a picture of a person or an actor or, like, a role or a character, and then, like, 'things that that person would say' " (Alex).*

When selecting a prompt for their own performance, the actors discussed prompts that stood out as good or bad for a show. Good prompts bring about scenes with *"the widest variety of intimacy and casualness" (Alex)*, character qualities that *"can be played in different ways" (Leah)*, and *"more emotional content and less specific content" (Alex).*

The actors were less attracted to scenes with plot *("We don't care about plot" (Elizabeth))*, and were particularly wary of *"transaction scenes" (Patty)* — where two actors go through an established series of exchanges to reach a clear goal. For example, *"'What would you say to a sales clerk while buying business clothes?' is very specific about where the scene is going to go. It doesn't leave you a lot of space [...] It gets boring" (Elizabeth).*

Instead, the actors gravitated towards questions that can lead to multiple, emotionally-charged interpretations.

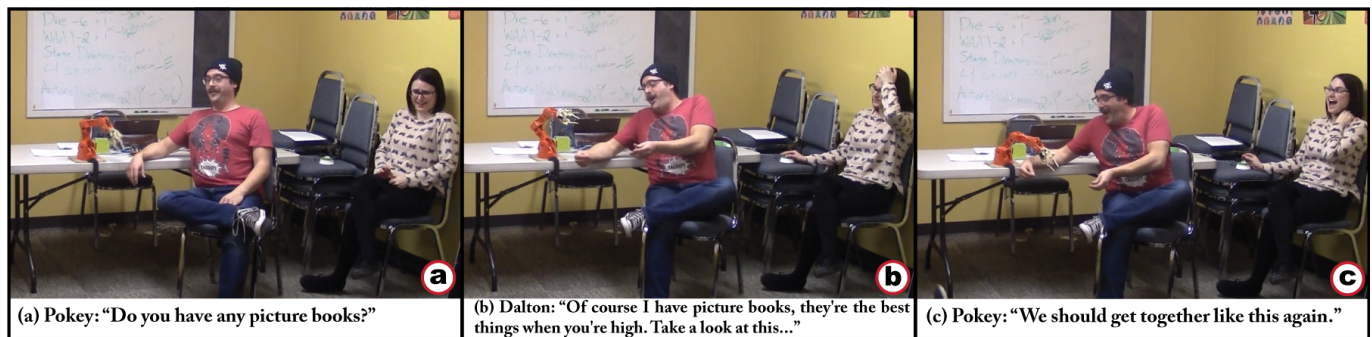*Dalton: I really like the question, "What is something you would whisper?"*

Leah: What I really like about the whisper thing is, like, why are you whispering? Like, you're embarrassed of it? Or is it creepy? [...] or, like, accidentally, like, racist or something...?

The actors chose the "barista in a coffee shop" set of prompts, swapping out the "best friend" prompt for the "whisper" prompt from another set (Table 1). They ultimately recommended formulating sets of prompts that (a) defines the context, (b) maintains the perspective, and (c) provides openness.

We also asked the actors to comment on how we asked the audience to select gestures. The actors approved of offering three choices of gestures for the audience. First, they reasoned, the limited set forces the audience to make a decision more quickly than if they had to pick from a broader set. Second, the random selection offered new possible gestures with each submission that could inform or inspire the scene. Finally, the actors felt the ambiguity of the gesture descriptions would entice the audience to look forward to discovering what their gesture looks like in the performance:

Elizabeth: It makes it more fun for the audience because then they don't know exactly what the gesture is, so then they're surprised when they see it.

With the dialogue prompts and gesture options set, the improvisers provided dialogue for their own scenes. They added

(a) Pokey: "Do you have any picture books?"

(b) Dalton: "Of course I have picture books, they're the best things when you're high. Take a look at this…"

(c) Pokey: "We should get together like this again."

Figure 6: In this scene: (a) Pokey suggests reading a picture book; (b) Dalton pulls out an imaginary picture book; and (c) Pokey, by chance, performs inquisitive gesture, which results in him moving in to look more closely at the book before delivering his line of dialogue.

input to RIPT both before they began performing, and between scenes, submitting 163 lines of dialogue in total.

While we expected the gestures to influence what the participants input into the system, they often ignored the gestures, and focused more on the dialogue prompts themselves. While RIPT accepted dialogue entries from 3-35 words in length, the experienced improvisers only provided short lines of dialogue (avg. = 6.11 words). They considered the impact of raising the minimum word count, or setting specific word count challenges as a way of getting the audience to provide more meaningful dialogue.

Patty: I got caught up in the "Oh, I want to do another one, I want to do another one". But if I have to get to fifteen words for this one, then I'll try a little harder.

Alex: Yeah, the system makes you want to write short sentences.

The performers also adjusted their input in anticipation of what they thought might be funny in a scene — without any sense of how the scene would actually play out: *"I found myself doing a lot of gag lines off of the top" (Leah).* Their excitement to create so much dialogue (albeit terse) was in anticipation of the performance itself:

Elizabeth: As an audience member, you want to hear what you write. You're so excited, so you're just, like, "I'm going to do as many as I can, so that I can hear my thing."

Structuring a RIPT Performance
We discussed several variations on how to run a RIPT show with the improvisers. They were hesitant to allow the audience to add dialogue to RIPT during the show, since they want the audience to focus on the performance, not their phones: *"It's just distraction-ville with all the phones out" (Patty).* They alternately suggested pausing the show between scenes or mid-scene: dimming the lights on-stage, bringing up the house lights, and inviting the audience to add new dialogue based on what just happened in the scene.

Elizabeth: You could be in a scene, like, you and me are in a scene and we're talking [moving hands like talking heads] "duh, duh, duhh…" and [the audience] is doing their thing for a

minute, and then we start the scene when the lights come up. But we're still doing the scene […] just no verbal.

This opened up the possibility for the audience to introduce *callbacks*, as a way of responding to what the improvisers had introduced in the scene. This could also heighten audience engagement, which is strongly impacted by how "live" the performance feels. In our earlier performances, we often encountered skepticism that the scenes were indeed improvised. The professional improvisers confirmed that this is a typical response to improv, and advised ways to use the prompts themselves to emphasize spontaneity:

Alex: [Better prompts are] anything that convinces the audience that it's more in the moment and that the robot's doing it live.

Having prepared prompts can communicate pre-planning to the audience — *"The audience is like, 'Oh, they've predetermined, like, five sets'" (Alex)* — even if the performers are unprepared and the show director (a non-performer) determined the prompts.

Performing Alongside Pokey
Once the improvisers added input to RIPT, they took turns performing a total of seven scenes with Pokey. When not performing in a scene, fellow actors observed as part of the audience. We discuss key elements that influenced a performance with Pokey, such as: how the improvisers created a cohesive scene, the comedic timing of Pokey's performance, the number of actors in a scene, and the physicality of Pokey's movements in relation to the performer.

Creating Cohesive Scenes
Pokey presented an unusual challenge for our improvisers. The key to making sense of Pokey's words and actions was for the actor to do the *"emotional weight lifting" (James)* of the scene. The improvisers were able to bring more emotional weight to their performance because they interacted with Pokey as a human, instead of with a hand puppet as in our initial performances:

Alex: What we can do while we are with Pokey is we can give all that nuance, [pointing to own face] all that stuff that, like,

the other sock puppet can't do. So [the sock puppet] really limits the range of expression [...] That's why Avenue Q is half-and-half, and, like, half-and-half is more interesting.

Most scenes were initiated by Pokey's first line of dialogue; the actors could then begin to contextualize what was said into a cohesive storyline. Only one scene did not begin with Pokey, where the actor struggled to incorporate what Pokey was saying; the actor often turned to clarifying the situation that Pokey's character obviously did not understand.

Patty: As an artist, what's your process? Just walk me through that.

Pokey: These elbows. A-A-A-A-R-P I will fight you.

Patty: Well...Well okay. You know, I know you're new to podcasts there [sigh] but can we get you some coffee?

Pokey: How was your day off?

Patty: Great, thank you. The interview is about you though.

In one of the more successful scenes, everything Pokey said was accepted as truth: Pokey's dialogue introduced a picture book, and later he gestured to "inspect" the book (Figure 6). The actor in this scene clearly used the "yes, and" principle[4] of improv to accommodate Pokey's unpredictable dialogue. While Pokey cannot "yes, and" in response, in this case the dialogue and gesture lined up such that Pokey appeared as though he had an understanding of the scene.

Timing
Pressing the GO button is another form of audience engagement — in both our early performances, and the rehearsal with professional improvisers, everyone enjoyed pressing the GO button. However, in our early performances, audience members needed a clear cue from the performer (e.g., eye contact, awkward silence) for when they should press the GO button to continue the scene. This gap between the cue and the execution of Pokey's line resulted in a more intermittent performance. In the rehearsal, the improvisers played the role of an audience member controlling the GO button — however, they advised against inviting a random audience member controlling the button, because of timing.

Patty: What's great about it is having an improviser on the GO button, because that timing is an important piece [...] As the person playing the button [in the 3 person scene] that was way more fun because I had to wait, and I was like, "where do I want it to be?" And then there were times where I was like, "talk faster, just talk over him, we don't even care what's happening anymore!"

The gestures also introduced their own performative timing — the "inquisitive" gesture stood out during the study, as it had a meaningful movement towards a specific, imaginary artifact on the table, and it delayed the dialogue by five seconds. We introduced this delay so that the gesture could carry more meaning. Due to the delay, the performer in the scene (Alex) could lean in and match Pokey's gesture, as though he were listening to a secret.

---

[4]http://improvencyclopedia.org/glossary/Yes_And.html

Patty: If we're starting to mimic the, uh, Pokey, then that's great. We'll do that all the time on-stage. We'll mimic each other, we'll copy each other.

Number of Actors
In our preliminary performances we only had one amateur improviser; in the rehearsal setting with professional improvisers, we were curious how RIPT might extend to a larger troupe. When we asked the actors about their expectations for the "barista in a coffee shop" scene, they clearly anticipated a two-person performance: *"It's going to be two people having a chat. Not a lot of action"* (Alex); *"Might be a lot of talking heads"* (Elizabeth).

While the actors mentioned the possibility of three-person scenes, no one started a scene with two human actors on-stage. However, it became clear that the scene could extend to more players, as needed. In one scene, Pokey's dialogue called for an additional character (Figure 7).

Pokey: Please pass the popcorn.

Leah: [passes mimed popcorn bowl to Pokey]

Pokey: Well, look who it is.

*Leah: [makes eye contact with Dalton who is off-stage]*

Dalton: [walks toward Pokey and Leah] Hey, hey guys.

Towards the end of the session, the improvisers became more playful with Pokey and decided to start a five-person scene from the start (Figure 8). This scene quickly devolved — all the performers played off of each other, while Pokey stayed in the middle of the scene interjecting non-sequitur lines that interrupted others. Pokey's character ended up being an unpleasant person at a bus stop, driving others out of the scene. In the end, Pokey was left alone to just say a few lines, creating a monologue, and finally drawing the audience in.

Leah: I was projecting so much emotion into everything he was saying. I was like, "poor guy."

Physicality of Pokey
When discussing alternative variations of RIPT, one actor suggested transforming a visual, animated character that could be projected anywhere on the stage. However, we observed that Pokey's physicality was an important element for the actors to play with during the scene. Pokey's physical presence was

(a) Pokey: "Well look who it is." (b) Dalton: "Hey, hey guys."
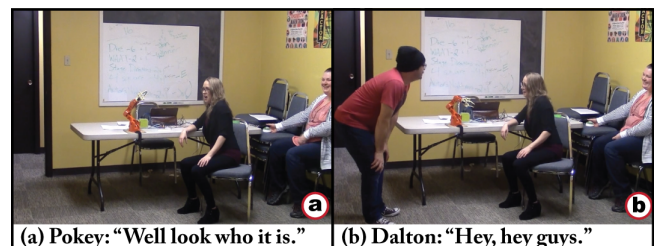
Figure 7: (a) Pokey delivers a line implying a new character and (b) Dalton takes advantage of this to invite himself into the scene.

not only mimicked by actors, but expressive in themselves. The actors valued the robot's ability to communicate beyond dialogue *("So much of communication is nonverbal" (Elliot))*, particularly those actions that clearly communicated an emotional state *("The dancing one was the clearest, like, the more expressive animal ones" (Alex))*. The participants saw potential in Pokey's gestural expression that was yet untapped.

Alex: It's almost like you want a puppeteer to spend an hour with it, there's some that aren't emotional enough. There's some [gestures] that are too subtle.

The most successful gesture — the "inspection" to look closely at a picture book — was only possible because there was such a clear indication of Pokey's "gaze". The actors wished they could play more with Pokey's gaze and eye contact — which was limited based on the location of his eyes.

Alex: You want to play with eye contact when he's looking at you and looking away, and you can't when [Pokey's eyes] are on either side [of his head], because he's a bird.

Pokey's physicality brings many of the same qualities as past work on robot performance [7, 12]. Because RIPT also brings in unpredictable audience input, Pokey could bring elements of a physical performer into an improvised scene while maintaining the spontaneous spirit of live improv.

DISCUSSION & FUTURE WORK

Our implementation and evaluation highlighted several current limitations of RIPT, as well as opportunities for future design iterations, research, and performances.

Pokey's Voices. We used a single male text-to-speech reader for Pokey's voice. This resulted in referring to Pokey as a "he", even though in practice, actors will adopt different voices to communicate a range of characters. In future iterations, Pokey's voice could be selected by the director to adapt to a particular role in a given scene: man, woman, child, or non-human (e.g., a gruff "wolf" voice).

Costumes and Props. Improvisers will often grab and use props or costume elements before the beginning of a scene

Figure 8: A multi-performer scene at a bus stop.

to help reinforce elements of their character. Our actors suggested dressing up Pokey with costumes or props that reflect a particular character while collecting audience submissions.

Physical Gaze Cues. Pokey currently has two googly eyes on the sides of his head — similar to a bird's eyes. However, the performers found it difficult to figure out where exactly Pokey was looking. Revisiting Pokey's design and gesture vocabulary — particularly from the perspective of a puppeteer — could improve the types of signals that Pokey gives to the actors and the audience.

Emphasize "Live" Performance. One way of emphasizing the "live" nature of RIPT would be to allow the prompts to immediately reflect audience input. For example, the director could ask the audience for a few initial elements of a scene (e.g., "What's a good Halloween costume?"; "Pirate"), and then create a set of three questions on-the-fly that reflect that prompt (e.g., three questions from the perspective of a pirate). Another way to emphasize the "live" nature of improv would be to pause the show so that the audience can provide the *callbacks* needed in Pokey's dialogue.

Audience Size. Audiences become engaged when they *"hear themselves, center stage"*; in an improv scene, an audience member's satisfaction heightens when their suggestion affects the show's outcome, such as: introducing a plot twist, making a joke, or creating a situational coincidence. When testing our system, we primarily used smaller audiences (6-20 people). However, as audiences become larger, an audience member may have increasingly lower odds of hearing Pokey perform one of their lines during a performance. While RIPT currently randomly performs input lines, greater awareness of where lines come from, or what their potential impact could be at different points in a scene, could increase the audience's satisfaction during a show.

Improv Format. RIPT was designed with short-form improv in mind; while the actors in our evaluation could imagine integrating a brief scene with Pokey in the context of a longer-format show, RIPT was designed for performances only lasting a few minutes. Due to its novelty, Pokey was generally the focus of a scene with one-on-one dialogue. However, improv troupes often include multiple performers on-stage, with new performers entering or exiting the scene as needed. While it is possible that our actors have been primed to use Pokey in one-on-one dialogue, our actors would have liked more opportunities outside of the suggested dialogue structure involving Pokey. Participant suggestions for future work include starting scenes with a second performer to break out of one-on-one conversations, or to include Pokey as a recurring character that appears periodically during a show (e.g., a landlord looking for rent money).

CONCLUSION

Is Pokey a good improviser? Clearly not. Even with the best prompts and audience input, Pokey is comedically terrible at many fundamentals of improv performance. Pokey is not aware of the scene at hand; he does not listen to his fellow actors; he does not create *callbacks*; he cannot "yes, and" in response to other actors; he does not know which actor on-

stage is addressing him; and he does not know where to look on-stage. But despite Pokey's lack of fundamental improv skills, RIPT offers a way for audience input to play a larger role in short-form scenes, leading to a fascinating, compelling improvised performance. And despite his naïvité, Pokey is capable of performing serendipitously good scenes. When Pokey recites the "right" random line of dialogue and gesture at the correct time, it's magical: inviting others into the scene as an improviser would, or playing into another performer's setup to deliver a scene-ending punchline.

ACKNOWLEDGEMENTS

REFERENCES

1. Allison Bruce, Jonathan Knight, Samuel Listopad, Brian Magerko, and Illah R Nourbakhsh. 2000. Robot improv: Using drama to create believable agents. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA'00)*, Vol. 4. IEEE, 4002–4008.

2. Steven Dow, Anand Kulkarni, Scott Klemmer, and Björn Hartmann. 2012. Shepherding the crowd yields better work. In *Proceedings of the ACM Conference on Computer-Supported Cooperative Work (CSCW '12)*. ACM, 1013–1022.

3. Dustin Freeman and Ravin Balakrishnan. 2016. Improv Remix: Mixed-Reality Video Manipulation Using Whole-Body Interaction to Extend Improvised Theatre. In *Proceedings of the ACM Conference on Designing Interactive Systems (DIS '16)*. ACM, 533–542.

4. John Harris and Ehud Sharlin. 2011. Exploring the affect of abstract motion in social human-robot interaction. In *RO-MAN, 2011 IEEE*. IEEE, 441–448.

5. Michaela Honauer, Patrick Tobias Fischer, Eva Hornecker, Julia Hahn, Bahar Akgün, Claire Dorweiler, Liese Endler, Yvonne Götzl, Muhammad Raisul Islam, Thomas Keßler, and others. 2017. Dusk: Adaption and Perception in Interactive Theatre. In *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems (CHI '17)*. ACM, 1037–1045.

6. Keith Johnstone. 2014. *Impro for Storytellers*. Routledge.

7. Woong-Ji Kim, Sun-Wook Choi, and Chong Ho Lee. 2012. MARIOBOT: Marionette robot that interact with an audience. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI '12)*. IEEE, 409–409.

8. Heather Knight. 2011. Eight lessons learned about non-verbal interactions through robot theater. In *International Conference on Social Robotics*. Springer, 42–51.

9. Kurt Luther, Jari-Lee Tolentino, Wei Wu, Amy Pavel, Brian P. Bailey, Maneesh Agrawala, Björn Hartmann, and Steven P. Dow. 2015. Structuring, Aggregating, and Evaluating Crowdsourced Design Critique. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work Social Computing (CSCW '15)*. ACM, New York, NY, USA, 473–485.

10. Brian Magerko, Waleed Manzoul, Mark Riedl, Allan Baumer, Daniel Fuller, Kurt Luther, and Celia Pearce. 2009. An empirical study of cognition and theatrical improvisation. In *Proceedings of the ACM Conference on Creativity and Cognition*. ACM, 117–126.

11. Charles B Owen, Alison Dobbins, and Lisa Rebenitsch. 2013. Theatre Engine: Integrating mobile devices with live theater. In *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*. ACM, 378.

12. Mose Sakashita, Tatsuya Minagawa, Amy Koike, Ippei Suzuki, Keisuke Kawahara, and Yoichi Ochiai. 2017. You As a Puppet: Evaluation of Telepresence User Interface for Puppetry. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, New York, NY, USA, 217–228.

13. Tom Salinsky and Deborah Frances-White. 2013. *The Improv handbook: The Ultimate Guide to Improvising in Comedy, Theatre, and Beyond*. Bloomsbury Publishing.

14. Sally Smallwood and Cameron Algie. 2011. Unsung Heroes Of Improv: The Tech. (2011). `https://peopleandchairs.com/2014/10/06/unsung-heroes-of-improv-the-tech/`

15. Flavia Sparacino. 1996. *DirectIVE–choreographing media for interactive virtual environments*. Ph.D. Dissertation. Massachusetts Institute of Technology.

16. Flavia Sparacino, Glorianna Davenport, and Alex Pentland. 2000. Media in performance: Interactive spaces for dance, theater, circus, and museum exhibits. *IBM Systems Journal* 39, 3.4 (2000), 479–510.

17. Anselm Strauss and Juliet M Corbin. 1990. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Sage Publications, Inc.

18. Leila Takayama, Doug Dooley, and Wendy Ju. 2011. Expressing thought: improving robot readability with animation principles. In *Proceedings of the International Conference on Human-Robot Interaction*. ACM, 69–76.

19. Gareth White. 2013. *Audience Participation in Theatre: Aesthetics of the Invitation*. Springer.

20. Lixiu Yu, Aniket Kittur, and Robert E Kraut. 2014a. Distributed analogical idea generation: inventing with crowds. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, 1245–1254.

21. Lixiu Yu, Aniket Kittur, and Robert E Kraut. 2014b. Searching for analogical ideas with crowds. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, 1225–1234.