Interaction with sound for participatory systems and data sonification

Liu, D.

CHAPTER 2

Interaction Models with Sounds for Real-time Participation

This chapter is partially based on the following publication:

Liu, D., & van der Heide, E. (2017). Interaction models for real-time participatory musical performance using mobile devices. In *Proceedings of the 2017 International Computer Music Conference, ICMC 2017, Shanghai, China, October 16-20, 2017*. Michigan Publishing.

2.1 Introduction

Sound interaction, as an interdisciplinary art form, it is relevant to the study of the arts, audience, behavior, techniques, interactions and so forth. Considering interaction with sound, participatory performance that involves music is an interesting area to gather examples on how such interaction can evolve. We consider the interaction consists of a system that communicates with participants, where a dialogue is possibly developed (cf. definition 2.1). If more people are involved this should be considered computer mediated interaction. In general, we start to consider systems in which the interaction is co-located and synchronous. In addition, remote but synchronous systems are considered. In terms of interaction this states whether or not the participants are in the same place as the system or not. For the further understanding we first introduce three major concepts that are important to the development of interaction models.

As indicated, crucial to the interaction is the dialogue:

Dialogue

Definition 2.1 *Dialogue has an interactive component, which is developed between the actions of the subject (i.e. audience) and the reactions of a system. It possibly also provides a means of learning.*

A dialogue can be applied in different context. As a paradigm we first investigate participation in musical performance defined as:

Participatory musical performance

Definition 2.2 *It is an immersive form of sound performance that directly invites the audience to be a part of, or influence the performance in real time.*

Participants are essential to dialogue and performances.

Participants

Definition 2.3 *In participatory musical performances, participants are the people who take part in a performance and partly contribute to the performance result.*

Besides the traditional western concert performance practice, the participation of audience in performances is becoming an emerging field. The current ubiquity of *mobile devices*, mobile phones or smartphones, makes it possible for large numbers of participants to interact with sounds in real time. The participants in such musical performance could be either just audience members or the audience in combination of professional performers. All considered audi-

ence takes different roles from performers, to composers, to editors, etc., varying from work to work. While some systems provide a relatively simple and passive participation form, other systems can develop a dialogue with the audience and achieve a complex participation form. Referring back to the ***dialogue model*** (cf. Figure 1.1), the verbal element enables audience input and actions, while the subject represents the entity or role that initiates the dialogue. Furthermore, the adjective provides descriptive information on how the system responds to the subject's actions, particularly in terms of sound production.

In this chapter, we first review documentation and publications of a series of real-time participatory musical performances, from which we deduce the relevant aspects as well as a journey map of participation. Then we discuss several forms of audience participation by mainly considering in three aspects: active / passive, direct / indirect, with / without limitations. From our analysis we derive a number of performance models that capture the interactions in different manners. Meanwhile, we assess usability issues for the interaction design of performances and follow the explanation of unidirectional and responsive interaction based on the approaches how sounds were designed and produced. Lastly, we propose possible directions of efficient sound interaction design for further research.

We provide video examples of some related work through QRcodes, which can be scanned with a mobile device for viewing. The QRcodes are numbered and given in the page margin.

2.2 Participants

In some existing participatory musical performances, participants could be divided into two groups, novices and masters (Lee & Freeman, 2013; Freeman et al., 2013). Miletto *et al.* considered a *novice* a music beginner, a person who lacks musical knowledge or who is just learning the rudiments of music (Miletto et al., 2011). We have excluded works that were designed for professional performers, as these required substantial practice or extensive prior musical knowledge. Thus the works we review in this chapter were not aimed at mastered musicians, but attempted to engage the audience in the performances. Furthermore for definition 2.3 we characterise participant as 1) audience and performers or 2) audience only. Performers may play a leading or a collaborative role in a performance.

2.2.1 Audience & Performers

In traditional western concert performances, there has been a clear separation between audience and performers. While the audience is watching and listening at the auditorium, performers are performing arts onstage in front of the audience. The development of mobile devices and wireless technology has opened doors for a completely different approach to engage the audience, creating the possibility for the audience to interact in a performance and become a part of the creation of the performance. In this respect, one of the pioneer works is *Dialtones*. This is premiered in 2001 (Levin, 2001). The audience was asked to register their phone numbers at web terminals and specific ringtones were installed on their phones. During the performance, the phones were dialed by the performers via a computer program that allowed 60 phones to ring simultaneously. In this way, the ringtones could be orchestrated as a musical performance (cf. QRcode 2.1). The audience's mobile phones became the performance medium, although their participation was *passive* (cf. definition 2.5) in the performance itself.

QRcode 2.1

Since then, performances have been developed to invite the audience to participate in a more *active way* (cf. definition 2.6). *massMobile* (Freeman et al., 2015; Weitzner et al., 2012) is an audience participation framework developed by Freeman *et al.* in 2012. It provides a possibility for the audience to shape an onstage live performance. It was used to develop an application for the live performance *Saxophone Etudes* (Freeman, 2012). The audience could vote for various musical factors through the application, including tempo, dynamics, note duration, articulations and measures of the music, which were displayed to the saxophonist for solo improvisation in real time. In this case, a dialogue can be initiated at lexical level via the interaction between the audience and the performer.

Both works combine audience participation with performers and construct different collaborations between the audience and the performers. While the performers directed the performance in *Dialtones*, the audience actually influenced the way the performance developed in *Saxophone Etudes*.

2.2.2 Audience Only

In some participatory performances, there are no professional performers and the audience is instructed to create or join a performance individually or form a group. In *Tactical Sound Garden (2006)*, participants can choose a sound from

a pre-designed sound library and plant it at a certain location by adding a GPS location (Shepard, 2006). Furthermore, they can modify the sounds planted by others and modify their volume and repetition time. All the sounds are mixed at the server-side and streamed to mobile devices that are used by the participants. As soon as someone walks into the garden, the sounds planted around him would be played in real time. When one plants or modifies a sound based on the mixture of former sounds in the garden, communication among the audience members emerges through the sounds. Each audience member contributes sounds independently, and one audience member can be seen as the contributor in others' view.

Swarmed is yet another case of audience only participation (Hindle, 2013). It applies a captive-wifi-portal allowing participants to generate audio via a webpage used on multiple mobile devices simultaneously. There are several pre-defined instruments on the webpage for the participants to choose from, which produced synthesised sounds based on gestures. The audience are able to hear the sounds they are playing on their phones independently, as well as the combination of the sounds produced by the other audience from onstage speakers.

SoundBounce shows the possibility for multiple audience members to be able to perform and interact with a sound in a group (Dahl & Wang, 2010). In *SoundBounce* a sound is regarded as a ball. Pre-defined gestures allow the players to throw and bounce the virtual ball to each other according to compass data using their mobile devices (cf. QRcode 2.2). The movement of the ball (or sound) is sonified through FM synthesis. Although the premiere of *SoundBounce* was initially performed by a group of professionals, we still included it in our audience category. *SoundBounce* utilises a physical metaphor to develop the mobile interaction and is designed in an intuitive way for a group of participants to perform and interact with each other. We see this as an equally meaningful approach in an audience only context.


QRcode 2.2

All aforementioned works are examples of audience only participation performances and the auditory results are non-deterministic. *Tactical Sound Garden* is not a performance with a beginning and an end and can therefore be seen as an installation format. *Tactical Sound Garden* and *Swarmed* have a similar approach in the sense that the participants make their own contributions and are not 'forced' to interact with each other. Interaction between the participants is an implicit requirement of *SoundBounce* since the ball has to be thrown and

caught. In all cases the participants can take a pause and just listen to what is going on.

2.3 Participation Journey Map

If we consider features of an interactive sound system, we take as a starting point that the feature is available to one or more participants. The system can be fully self-operational or can have interference with a performer. In the latter case the conditions, i.e. rules, of the system can change during the interaction. We investigate these systems in relation to the interaction that is provided. The interference of the performer/operator is, for now, of less interest.

In this section, we focus on the audience rather than professional performers. Participation journey map is a well-known tool to investigate participation across various states and how individuals transition between these states(Mast, de Vries, Broekens, & Verbeek, 2021). We analyze different forms of audience participation based on a user journey map (see Figure 2.1, cf. definition 2.4), which consists of three main stages: observation, learning, and interaction. Some participation forms are passive and may include little interaction, while other forms may in-

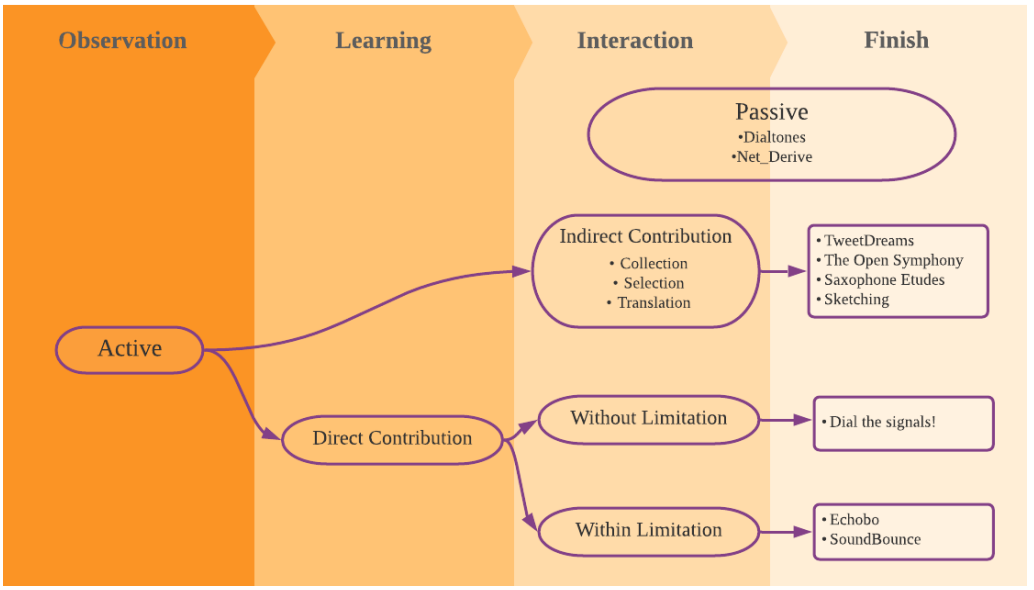

Figure 2.1: A journey map of audience participation.

volve a more complex progression that requires the audience to learn for effective interaction.

Participation Journey Map

Definition 2.4 *A tool to visualize the process that a participant experiences through when interacting with a system.*

2.3.1 Passive Participation

We first discuss the passive form of participation and define it as follows.

Passive participation

Definition 2.5 *Passive participation is a form in which the audience does not influence the result of a performance, but are nevertheless aware of participation.*

In *Dialtones* (Levin, 2001), the audience did not need to take any action but only brought their phones along. This innovative way of performing sounds tried to get participants involved in the performance, still the passive role might make participants feel surprised more than engaged.

In *Net_Dérive (2006)*, the participants were given a broader palette of actions to follow. There were several paths for participants to choose and follow around a specific location, i.e. a gallery (Tanaka, 2006). While the participants were walking, ambient sounds were automatically recorded by the participants' phones at regular intervals. These recordings were mixed at server-side and played in the gallery. Although the participants could choose different routes around the gallery, the process of recording was automatic.

Compared to *Dialtones*, in *Net_Dérive* the participants were activated more. But they did not have the opportunity to fully decide the recorded material. Both works require the audience to be present and participate but the audience has little or no influence on the sounding result itself. Thus in journey map, passive participation skips the stages of observation and learning (see Figure 2.1). Additionally, the way of interaction is quite limited.

2.3.2 Active Participation

The interaction of an active participant with a system requires from the participant to observe and pick up the clues that are provided by the system in order to understand the system (see Figure 2.1). The clues are given by the total dia-

logue (cf. definition 2.1) that is provided by the system. Visibility is an essential quality of allowing participants to easily recognise the clues (cf. definition 2.7).

Active participation

Definition 2.6 *Active participation signifies that the audience makes choices through interaction to influence and construct a performance via interaction.*

Visibility

Definition 2.7 *Visibility is a degree to measure whether the reaction from a system (feedback) is noticeable and understandable for the audience (Dix et al., 2003).*

In this section we will make a distinction between two forms of active participation: direct and indirect contribution.

Direct Contribution

Direct Contribution

Definition 2.8 *Direction contribution indicates the condition in which the audience is directly involved in the production of sound. The auditory results can be a clear feedback for the audience to perceive and help them learn the system (cf Figure 2.1).*

An example of direct contribution is that actions of participants are directly utilised as the input of sound generation or to trigger audible events. The audience is likely to become aware of how everything functions from aforementioned conditions, which brings possibility of learning to the audience (see Figure 2.1). In *Dial the signals!* (Ligna & Röhm, 2003), a matrix of mobiles phones was exhibited as an installation and the numbers of the mobile phones were passed to the audience to dial. Every tone the phones played was broadcasted by several radio stations and as a live stream on the internet. The audience had full control of deciding which phone to dial and their involvement was direct. Furthermore each of the phones corresponded to a specific sound, which provides a clear visibility of the auditory feedback for the audience to track. The audience is able to learn the mapping between the phone numbers and tones accordingly (see Figure 2.1). Meanwhile, the audience gets complete freedom and there are no specific choices made regarding the development over time. It is impossible for the designers of this work to predict which phone would be dialed first or in which order a series of phones would be dialed.

In some other works, specific interaction rules are used to limit the direct contribution from the audience. In *Echobo* (Lee & Freeman, 2013), an eight-key keyboard was displayed on the individual mobile phone as an instrument. The audience was instructed to play the keys of the chord selected by the composer or a master musician (cf. QRcode 2.3). The sequence of playing keys was, however, not fixed. The keys were marked with a black arrow and an electronic piano sound was generated directly from the phone as the harmonic backdrop of the performance. The harmonic structure was controlled by providing limited keys for the audience to play. Accordingly, the audience collaborated with the performer in harmony, while having partial freedom to play the instrument.

QRcode 2.3

In *SoundBounce* (Dahl & Wang, 2010), the audience performed with certain mapping rules pre-defined by composers and developers. The movement of the virtual ball was sonified through frequency modulation synthesis, which can be categorised in parameter mapping sonification approach (PMSon, cf. definition 4.3). The melodic pitch got higher and the sound became louder as the ball rose. Additionally, the sound cross-faded from the thrower's phone to the receiver's phone. *SoundBounce* is an example showing how motion data has been sonified in an intuitive way for participants to interact with. Limitations were set up to control the randomness and unpredictability in *Echobo* and *SoundBounce*, to improve the structure and visibility of auditory results.

Indirect Contribution

A participatory performance using interpretation or manipulation of the audience's contribution might have a more indirect nature. Therefore, we define indirect contribution as follows:

Indirect Contribution

Definition 2.9 *The audience's active participation is passed through a selective and translation process, whereas the audience members themselves do not control or produce sound directly or instantaneously. Indirect contribution reduces the visibility of feedback.*

So if the contribution is indirect, it is harder for the audience to observe or track their contribution due to the lack of feedback. While being indirect, the audience plays a crucial role in the performance. For example in *TweetDreams*, the audience was asked to tweet during the performance (Dahl et al., 2011).

QRcode 2.4

Tweets with a certain hashtag were able to be searched and collected. The specific hashtag and the rate of appearance of each tweet were determined by the performer, furthermore the tweets were used as an input for an algorithmic composition (cf. QRcode 2.4). The combination was indirect in this example, thus it was harder or impossible to learn how the input data (tweets) had been transformed into sounds (see Figure 2.1).

The Open Symphony is another example of indirect contribution. Four professional musicians improvised in accordance to a score resulting from the audience's votes (Zhang et al., 2016). The audience could vote different playing modes for different players through a user interface, including single notes, melodies, free improvisation, silence, etc. The result was displayed as graphic notations on a big screen from left to right. With a limited amount of features to vote, the audience contributed compositional resources to the performers to improvise.

2.4 Performance Model and Sound Production

The interactive experience is constructed through a dialogue between the audience and the system. Such dialogue is possibly developed in a performance (cf. definition 2.1). By participating in an interactive sound based system, an audience member can control or influence the auditory outcome of the performance stepwise (cf. definition 2.10). Therefore, we first introduce the concepts of *performance* and *performance model*. In regard to the sound production in each performance, we analyze it with respect to the concepts of *deterministic* and *stochastic*.

Performance

Definition 2.10 *A performance consists of the join of actions of the audience and the reactions of an interactive system, which involves how a sound is produced through interaction.*

Performance model

Definition 2.11 *A performance model describes the connections among audience members and between the audience and the performance system.*

Deterministic System

Definition 2.12 *There is a fixed set of mapping rules between input data and output sounds.*

Stochastic System

Definition 2.13 *The output sounds are generated from the input data via a model with randomness and mutations.*

Schraffenberger and van der Heide considered mutual influence between audience and interactive artwork as an important underlying principle of interaction (Schraffenberger & van der Heide, 2015). Some performances might have an evolving interactive dialogue, whereas the interaction model could be static in other performances. Accordingly, we will categorise and discuss two models in this section, the inherent performance model (cf. Figure 2.2) and the responsive performance model (cf. Figure 2.3). In order to align with the dialogue model proposed in Chapter 1 (see Figure 1.1), same color coding is used in Figure 2.2 & 2.3 to analyze the elements from a dialogue in both performance model.

2.4.1 Inherent Performance Model

In *Dial the signals!* (Ligna & Röhm, 2003), participants were free to dial the mobile phones exhibited in the installation. The corresponding phone responded immediately with a specific ringtone. Although there is possibilities for participants to learn the mapping between ringtones and phone numbers and choreograph the combination of ringtones to create a larger composition. There is no varying interactive dialogue taking place between the actions of the audience and the reactions of the system. The interaction was unidirectional and ended when a phone was rang (see Figure 2.2a). In similar fashion, in *Moori* (Kim, 2011), the audience responded to a list of guided questions sent by the performer to form a story. The text messages then were spoken with text to speech software. Despite the questions that were designed like a narrative script to keep the audience answering the questions, the interaction ended when one answer was spoken (see Figure 2.2b). Both aforementioned works are examples for a deterministic system (cf. definition 2.12), where there are simple fixed relations between the actions of the audience and the sounds being played. The performance model is linear and unidirectional. Therefore, there was no real opportunity for an interactive dialogue.

In *TweetDreams*, tweets were used as input for an algorithmic composition (Dahl et al., 2011). The algorithm brings a stochastic system creating a more diverse auditory outcome (cf. definition 2.13). A dedicated computer program analyzed the selected tweets and checked whether one was new as a root or was

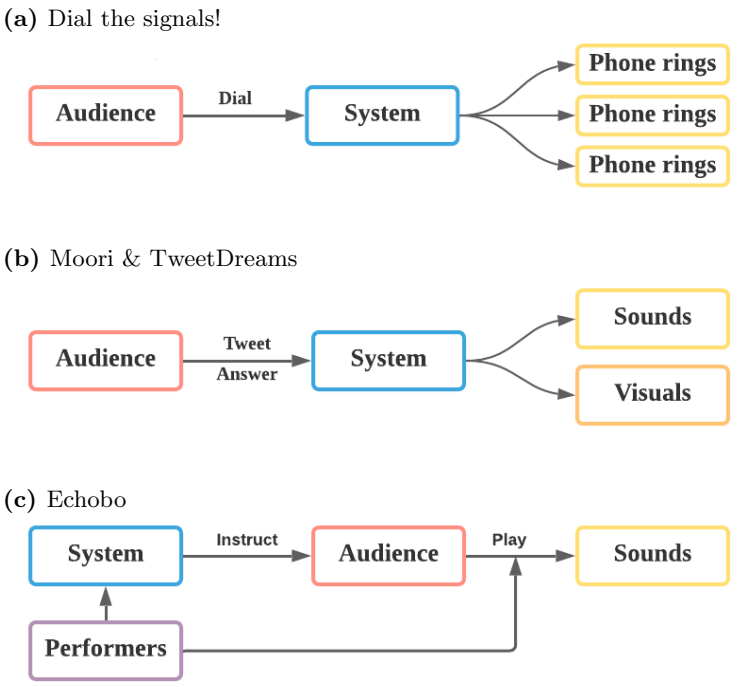

Figure 2.2: Diagram of Three examples of Inherent Performance Model.

related to an existing root. When a tweet starts a new root, a new melody will be chosen from pre-composed melodies, otherwise the melody will be mutated from its parent melody. This is another example how model based data sonification (MBS, cf. definition 4.4) has been used in real-time audience participation performance. Although the visualisation of tweets gave the audience feedback as clues to track and locate their own contribution, the musical transformation may not be clear enough for a true interactive dialogue.

In *Echobo* (Lee & Freeman, 2013), there was one musician controlling the chord progression, and an acoustic instrument player performing melodic material. The audience can play only with the harmony defined by the musician. Still, the audience was free to decide the order of keys they played. The performer would probably affect the decision of the audience about the order. The unidirectional interaction started with the chord selection and ended when the audience finished playing the notes from the chord without any further dialogue.

In order to characterise this kind of performance model, we introduce the

concept of ***inherent performance model***. From the diagram (cf. Figure 2.2), it can be deduced that the interaction in the inherent performance model is often unidirectional.

Inherent performance model

Definition 2.14 *The inherent performance model is a form of interaction that does not provide an actual interactive dialogue between the audience and the interactive system. The audience is not able to fully understand the system due to the lack of perceivable feedback.*

2.4.2 Responsive Performance Model

On opposed to the inherent performance model, we introduce the concept of ***responsive performance model***.

Responsive performance model

Definition 2.15 *The responsive performance model provides a constant interaction between the audience and the performance, in which the audience is motivated by various forms of feedback from the system.*

As an example considered, the improvisation of the performers could give a clear feedback to the audience and create a continuous interactive dialogue (cf. definition 2.1). It may also make the audience feel that they were engaged and influencing the performance. Because the audience might be curious about, or anticipate on how their contribution is being translated or performed.

In *Sketching* (Freeman et al., 2013), the audience could draw different shapes through a web page developed within *massMobile* (Weitzner et al., 2012). Each shape was assigned to an instrument. Other features such as colour, size, opacity, border were mapped to different musical parameters. Performers from a jazz band would pick and play a jazz standard and improvise based on the drawings from the audience, or entirely improvise with the drawings. The audience might draw a new shape or adjust the features according to the previous performance, and try to develop the performance further. This is a typical form of a responsive performance model.

In *The Open Symphony* (Zhang et al., 2016), the graphic notation presented the individual and collective feedback for both the audience and the performers. The decision of a single member of the audience may be influenced by the other

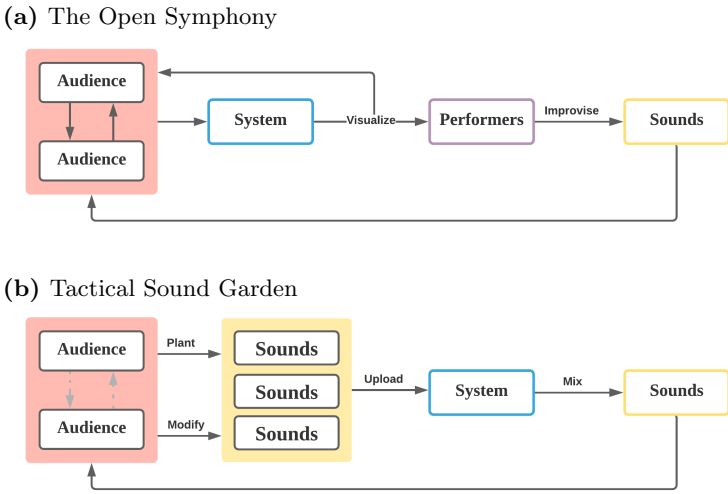

Figure 2.3: Diagram of Two Examples of Responsive Performance Model.

audience members as well as by the performers (see Figure 2.3a). Differently, in *Saxophone Etudes* (Freeman, 2012), although the points selected by the other audience members were presented on mobile of each audience member, the audience did not know the results of voting for each musical factors, such as speed, dynamics, etc.

In *Tactical Sound Garden* (Shepard, 2006), the participant was not only able to plant a sound, but also modify the sounds planted by others. Influenced by the mixture of the sounds in the garden, one participant may come up with various ideas about planting and modifying sounds. Such interaction remained among the participants and the sound mixture. In addition, it extended the experience of the audience beyond the time that the audience is interacting themselves. Therefore, a responsive performance model not only provides better visibility to stimulate constant interaction with the system, but also provides the clear interaction in a group of participants.

2.5 Discussion

In contrast with visual feedback, the auditory feedback might be less noticeable and understandable when the contribution of the audience is indirect. This depends on whether the audience is capable of associating the musical outcome with

their actions or contributions. Sound design or the adjective of a dialogue can possibly improve visibility and affordance (see Figure 1.1). Of the performances that we have reviewed in this chapter visual feedback is used to create insight in the performance and make the interaction more engaging. On the other hand, the visual feedback might also distract the audience from the actual auditory result.

Our interest focuses on the interaction with sounds and therefore initially we concentrate more on the sound and music than the visuals. In this context, we prefer a form of interaction in which the contribution of the audience and the auditory results can be clearly perceived, which possibly provide feedback for the audience to understand the interaction (cf. Figure 2.4).

The inherent performance model (cf. definition 2.14) has the characteristic that lacks an evolving interactive dialogue, therefore we argue that it would be harder to keep the audience participating in a performance with low visibility. The responsive performance model (cf. definition 2.15), however, could lead to a more dynamic and ongoing form of interaction.

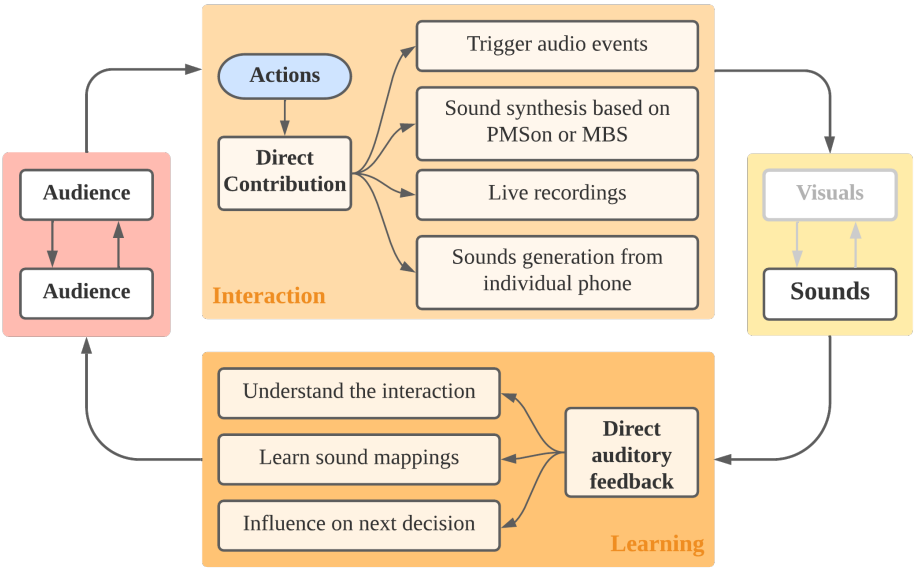

Figure 2.4: Proposed Framework for Future Research. A responsive performance model combining with direct contribution and direct auditory feedback, aiming to achieve a constant loop between interaction and learning. The color code refers to the elements of the dialogue model: subject-audience, verbal-actions, adjective-sounds.

Although efforts have been made to maintain the audience participating longer by elicitation, for instance, questioning (Kim, 2011) or continuous instruction (Lee & Freeman, 2013). The feeling of participating could be broken at any time during the performance since the audience does not get enough hints of participation from the result. Furthermore, a static form of interaction could be in the way of developing a true dialogue between the audience and the performance system. In most of the discussed works in section 2.4.2 the participation form has an indirect nature.

It is worth a try to combine the approaches used for direct contribution in the context of a responsive performance model (cf. Figure 2.4). The proposed framework attempts an explicit description of a dialogue between the audience and the system. It includes two main components from the participation journey map (cf. Figure 2.1), interaction and learning. In this case, sound will play an important role to help the audience to learn the interaction rules and understand the interactive system. While learning can motivate the audience to better interact with the system, the auditory feedback from the interaction can boost the learning process.

Additionally, a stochastic system can bring along an element of surprise (cf definition 3.2). In that case the dialogue might develop in an unpredicted way as the feedback has a stochastic element. However, it may result in low visibility of the feedback. This is also seen in case if the system is operated through a performer that changes the state of the system and thereby changes the expectations that were acquired by the participants. In this case, a clear feedback of sound design is required to achieve obvious affordance and improve the visibility.

2.6 Conclusion

The mobile device is a widely available medium for the audience to participate in a performance in real time, but it is not necessarily used. We have chosen not to focus on the device or technology itself but rather on the possible forms of interaction. Participatory musical performance is a great start point for our journey of exploring and learning about sound interaction.

In the course of our review, we have presented the descriptions of *participants* (cf. definition 2.3). The audience turns out to be the participants when they take part in an interactive system. Performers could be included to lead or collaborate

with the audience in a performance but are not a requirement. The journey map has been an essential tool for us to visualize the participation experience from observation, learning to interaction (see Figure 2.1). It is also the foundation for the proposed framework (cf. Figure 2.4).

We used the audience as an example to investigate the behaviors of subjects in a dialogue. We have categorised audience participation forms from their behaviors via the participation journey map (cf. Figure 2.1). There is little interaction existing in passive participation. Differently, active participation supports a wider range of interactions. In some works, certain rules were set up to limit the possible results and create more control over the performance. Still, participants can be aware that their decisions are forming and influencing the performance if their auditory contribution is direct. Indirect contribution employs a system to collect choices of participants, and translate those materials into compositional resources to create a performance. Participants experience less clear direct feedback of interaction from this form.

We have distinguished two performance models. We have analyzed the interactions among audience members, and between the audience and the performance. Subsequently, we have labelled existing works with inherent or responsive performance models. Aspects such as interaction, feedback, randomness, control have been discussed. We are particularly interested in exploring the experience of a responsive performance in combination with direct contribution in order to create a more engaging and interactive dialogue experience (cf. Figure 2.4). The system's responses can vary based on the actions of the audience and the context of the interaction. By considering the three elements of the dialogue model, we can create more engaging and tailored dialogue systems that enhance the audience's experience. Accordingly, we need to investigate interaction design (verbal) and sound design (adjective) further. While interaction design can influence a dialogue at semantic level, including how the participants comprehend the system. Various sound design approaches need to be examined what works for better auditory feedback and how to realize a dialogue at syntactical level for learning.