

BRIAN SCHMIDT ON AUDIO FOR VR/AR

By Tom Kenny

Over the years, Brian Schmidt has become something of an informal resource to Mix on all things having to do with videogame sound and music, interactive storytelling, and now Virtual and Augmented Reality.

This year, in early October, Schmidt's GameSoundCon, one of the world's leading gatherings of technologists and artists involved in interactive audio, celebrates its 10th anniversary. We thought now would be a good time to check in with him on what more traditional film and TV audio post pros might need to know in entering the coming wave of VR experiences.

Mix: Congratulations on 10 years. From game sound to interactive VR experiences.

Schmidt: Thank you. VR and AR have definitely been exciting for interactive audio in general. From a technology delivery perspective, the way you need to do audio in VR kind of has its roots in games, which is why game engines like Unity and Unreal are so popular even among creators of non-game VR experiences. The whole HRTF and 3D sound and room modeling and sound propagation—the game folks have been working on these issues for a couple of decades. So to see it kind of reborn and be placed front and center in the experience is exciting.

And yet it's a whole new world for audio.

One of the things I love about interactive in general is that we're still sort of making it up as we go along. It's immersive, and it's simulation, but it's not. Let's say you're making a big VR experience and somebody walks into a cave. We can trace the reflections off that actual cave environment. But for a person viewing that VR experience, a recreation of an underground cave in Tennessee, they probably don't know what a real cave sounds like, but they know what they think it should sound like. So in a sense, the goal of the VR experience has these dueling sonic requirements. On one hand, you want it to be faithful to being in that cave. But at the same time, being too real could be detrimental. Much



like film and television.

And yet it's world's different from traditional film or television.

I kind of view VR audio as an audio magnifying glass: It will enlarge problems with audio in interactive experiences. In a typical game in a 5.1 system in a quiet room, it's easy to miss nuance. But in a VR experience, if there's a robot hovering in front of me with a lot of moving parts and I want to walk up to it and turn to the back and side of it, I want to feel like I'm up close. The robot becomes very complicated once I have the ability to walk up to it.

The tools that the game engines have—Unity, Unreal or custom—or interactive audio tools like Wwise or FMod or Fabric, they allow you to create these complex sounds like a car or a robot or an explosion and give you the illusion of sound spherically around you. And they do a pretty good job.

We tend to focus a lot on sound effects and music whenever a new surround format comes out, but dialog is still king.

Consider dialog in a VR experience. Suppose I have a character and suppose I have the ability to walk right up to them and they have some lines of dialog that are part of the story. If I'm right next to that character, he's probably going to talk in a relatively soft voice. But suppose I'm 30 feet away and they still have to deliver the line of dialog, well they're going to raise their voice a bit. Now all of a sudden I not only need HRTF technology and sound propagation and all that,

but I need to have a different vocal performance depending on how far away they are. It's literally another dimension. Not only do I need to place the sound there, but I need it to sound right from wherever they deliver it.

Forget about the VR aspect, just the interactive aspect. When that character delivers that line is not fixed in stone. There may or may not be other pieces of audio occurring. This whole notion of intelligent interactive mixing, where the mix occurs in real time, is new.

And music?

As I said, it's the Wild West and we're still figuring out what works and what is distracting. There is obviously a temptation to go whole hog with 3D music and VR, kind of like the emergence of Dolby Digital. But it's so application-specific. If the music is truly non-diegetic, meaning it is just there to serve the emotion and there is no kind of relationship of the music to the world that it is in, then you can be fairly traditional. It's there, stereo to the ears.

As soon as you start incorporating surround or 3D music, you are pulling the music from being non-diegetic to being more diegetic. It's almost on a continuum now. The more you make the user aware that the mix is changing or the location of the instruments is changing, you're changing the role of the music. We're still figuring that out—what is awesomely cool, and what is distracting from the story. If the whole purpose to immerse yourself in the music, as far as lending emotional support, plain old stereo seems to work pretty well. ■