directing is proper

dope

directing attention in virtual reality

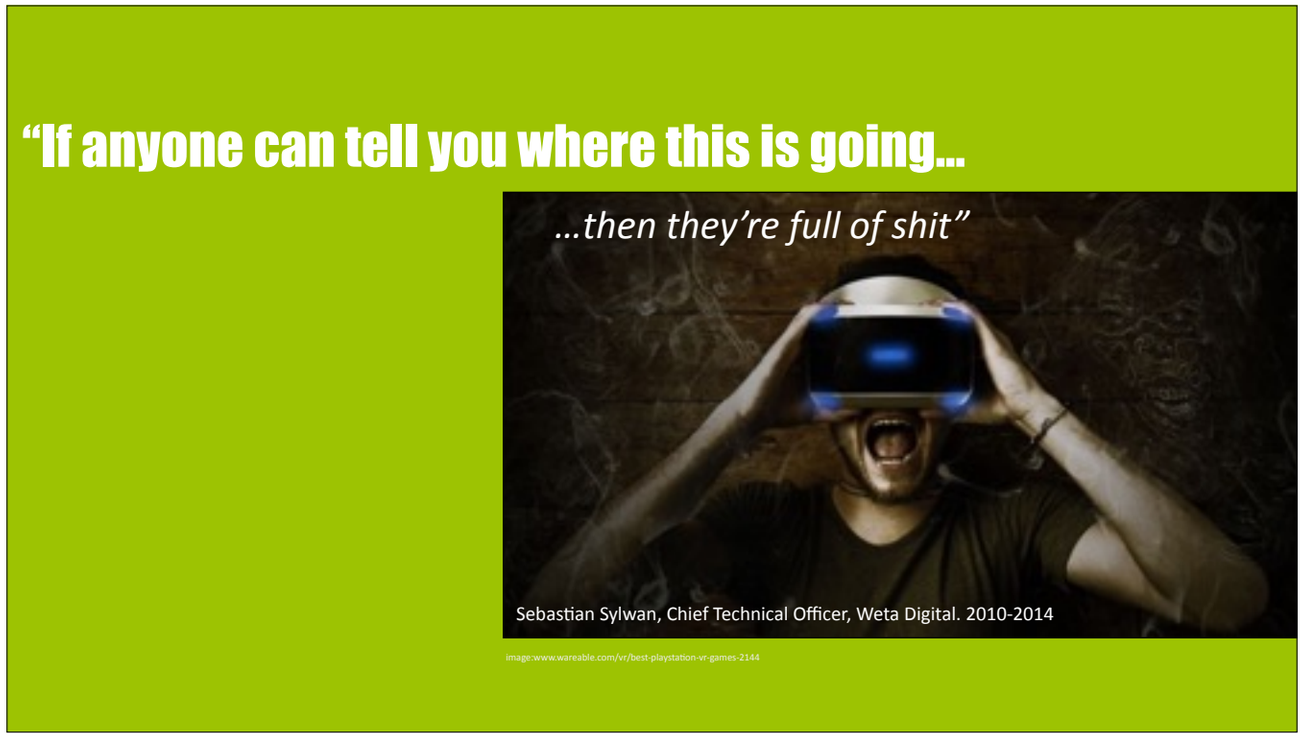
andy joule. falmouth university. cornwall. uk

“If anyone can tell you where this is going…

image:www.wareable.com/vr/best-playstation-vr-games-2144

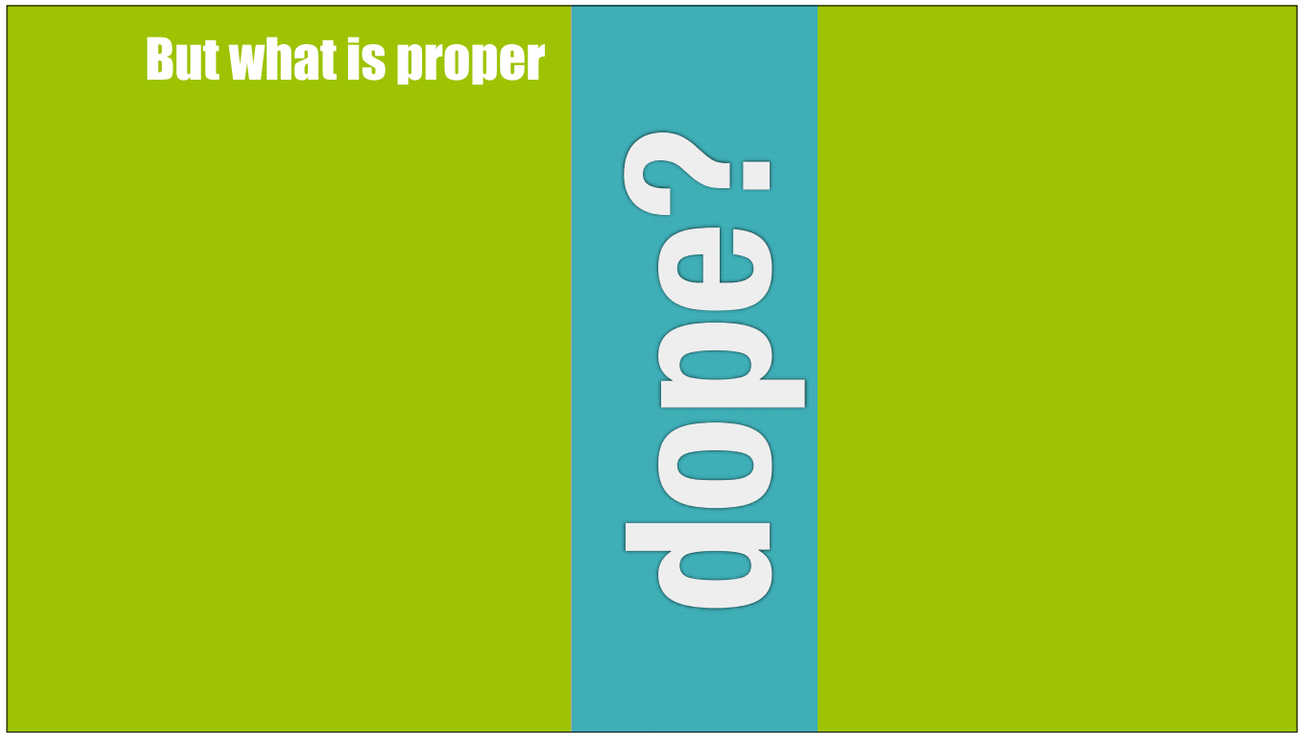
I probably need to introduce my presentation with something of a caveat. I am by no means an expert on virtual reality, so apologies if you were expecting me to enlighten you as I certainly don’t have all the answers. Why should I?

As Sebastian Sylwan, one time Chief Tech Officer of Weta (allegedly) said of VR: “If anyone can tell you where this is going…

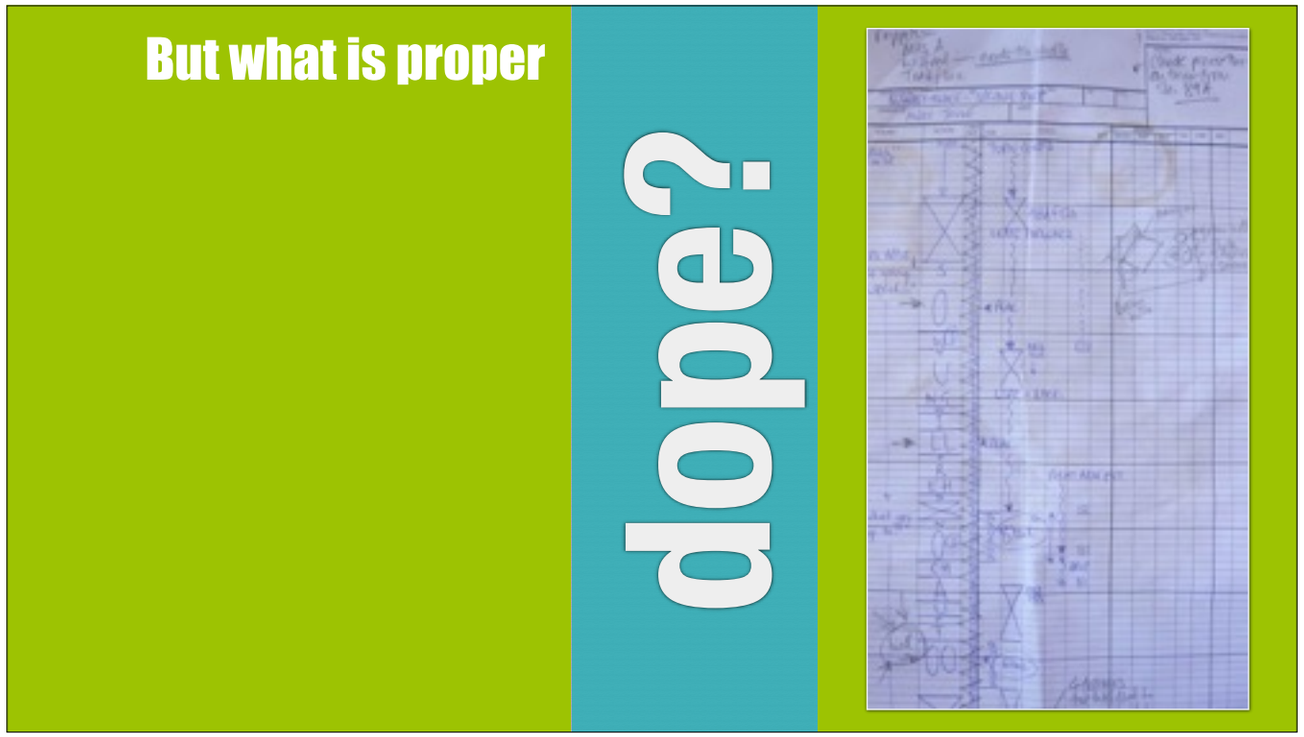


…that”.

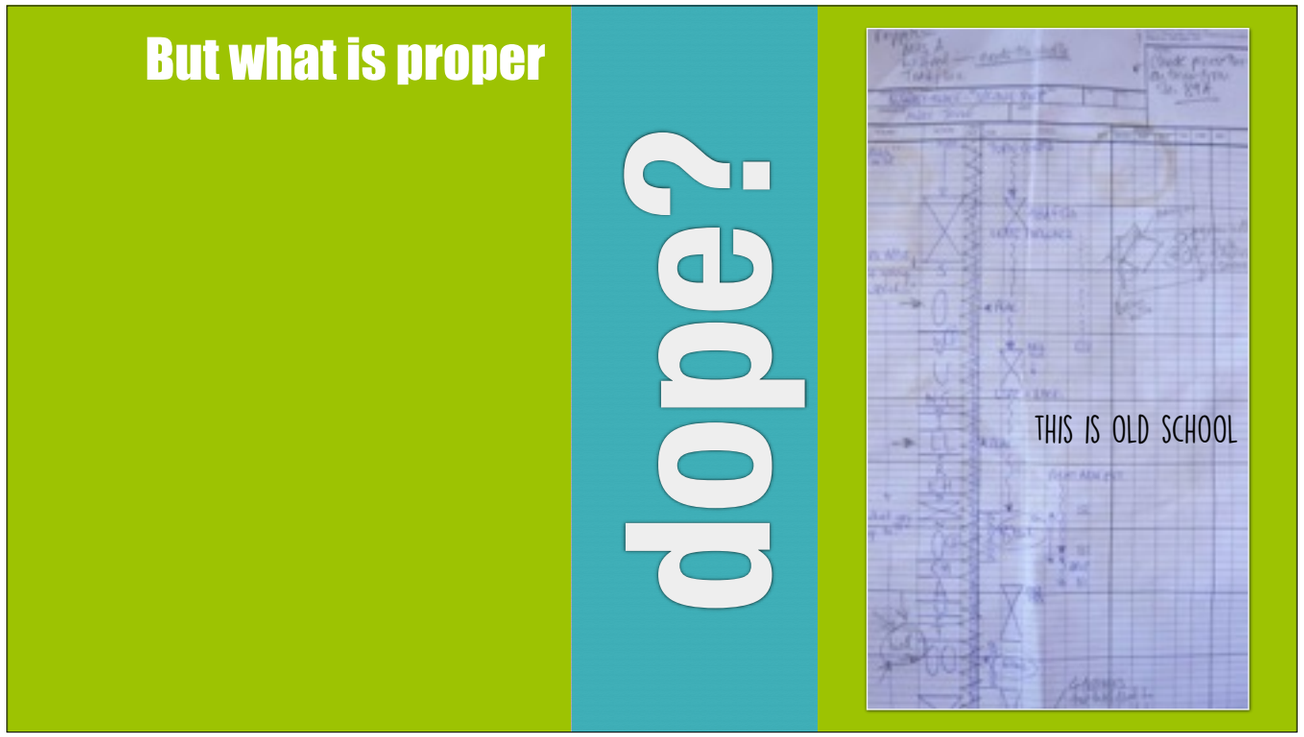
Why is directing it proper dope though? I’ve been researching how directors shift attention and alter points of view working in this emerging medium. Do old techniques apply? And what, if any principles should be at the centre of a directors thoughts?



What is dope? In animation it’s the voice that shapes the performance created by the animator. As a director I often ended up working with a script written by someone else. In those instances it was really important that I had figured out what actions each of the characters were going to perform before I went to the voice record.

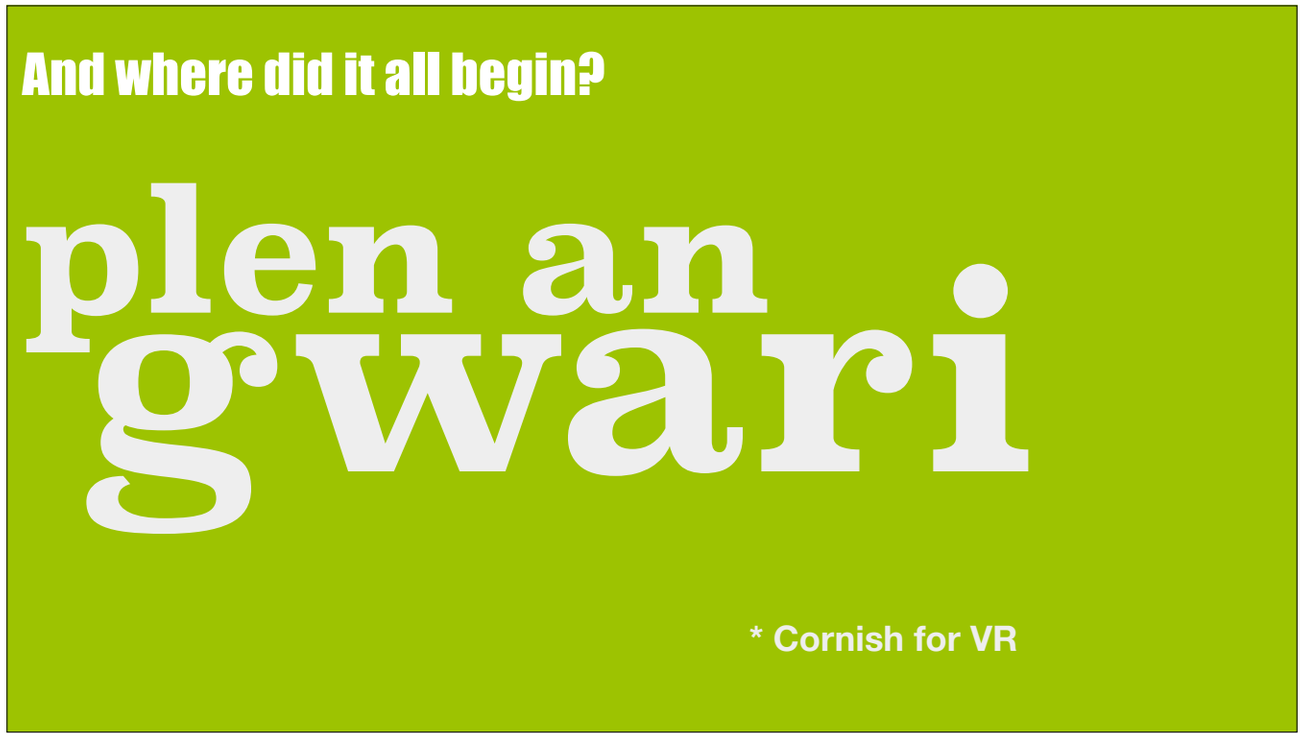


As a director it was vital that I knew the script inside out so that I could guide the voice artists - to describe to them what actions the characters were to perform when animated. Once the script was recorded and the voice track assembled from the various takes the audio was then broken down phonetically a frame at a time onto a ‘dope sheet’. Hence animation is proper dope. This necessity to treat the vocal performance in this way is unique to animation.



The dope sheet allows the animator to plan out their animation, to be frame accurate with lip synch and for them to plan out the body movements and gestures pegged to the vocal performance in line with its rise and fall. The planning and visualisation of the vocal performance are paramount decisions made by a director and directly inform the animated performance. Therefore the performance is imagined right at the start, yet the performance itself only becomes evident towards the end of the process. But how an animation director directs the audiences attention is the same as it is for live action. The point of view is controlled by what the camera shows us. Audio cues also works in the same way.

But what if the point of view is not fixed, giving an audience freedom to look all around. How can their attention be pointed to specific locations, particularly if the experience is to be an immersive narrative one. This is one particular challenge for VR but perhaps there are solutions to be found in the world of Plen an Gwary.



Their VR started pre 1400 AD. Ok so it was a VR kit built from stone and earth, theatre in the round. But that’s where it all began. In the UK at the very least!

And where did it all begin?
the playing places of

cornwall

In his book ‘Plen An Gwari - the Playing Places of Cornwall’ the theatre-maker and academic, Will Coleman argues that “rather than sat passively around the sides, the audience thronged through the whole plen an gwari space ‘on the hoof’. With artfully constructed scenery, costumes and props, massed chorus, live animals, guns etc a Gwari Meur would have been an epic immersive active experience”.

And where did it all begin?
the playing places of

image: Heidi Ball/Will Coleman. goldentree.org.uk/plen-an-gwarri-book/

image: www.bewnanskernow.org/cornish-culture-blog/invitation-to-the-launch-of-plen-an-gwarri-the-playing-places-of-cornwall-project

The top image shows what Coleman imagines the original Plen an Gwari would have looked like, with scenery and performing stages set around the stone and earth bank and the audience thronged in the middle. Below are the remains of Perran round plen near Perranporth, Cornwall.

Coleman supports Dr Richard Southern's research in his 1975 book 'The Medieval Theatre' where the thinking was that the audience's attention was directed to segments of the ring on small 'stages' so that their sight lines were ensured. Action then moved from one 'stage' to another using either loud sounds or dramatic actions to switch focus. Further research into stage directions by historian Sydney Higgins picked up by Coleman identified that "each section of the play would occur in a certain segment of the arena with only a few pavilions (stages) being active at any one time. For each section of the play a different segment of the arena is cleared of audience. Each section is concluded with a big… event that would draw the audiences attention and reorient the audience to a new segment".

So perhaps we can use these devices, used successfully over 500 years ago, in VR. Indeed we probably already are. Only the segments of the plen an gwari now become cones of focus. But more on cones of focus later.

Directing attention in virtual reality

sound or image

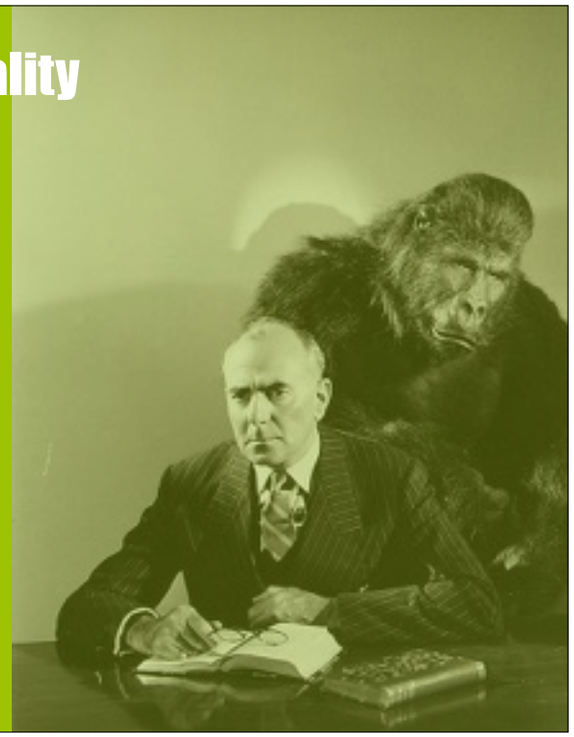
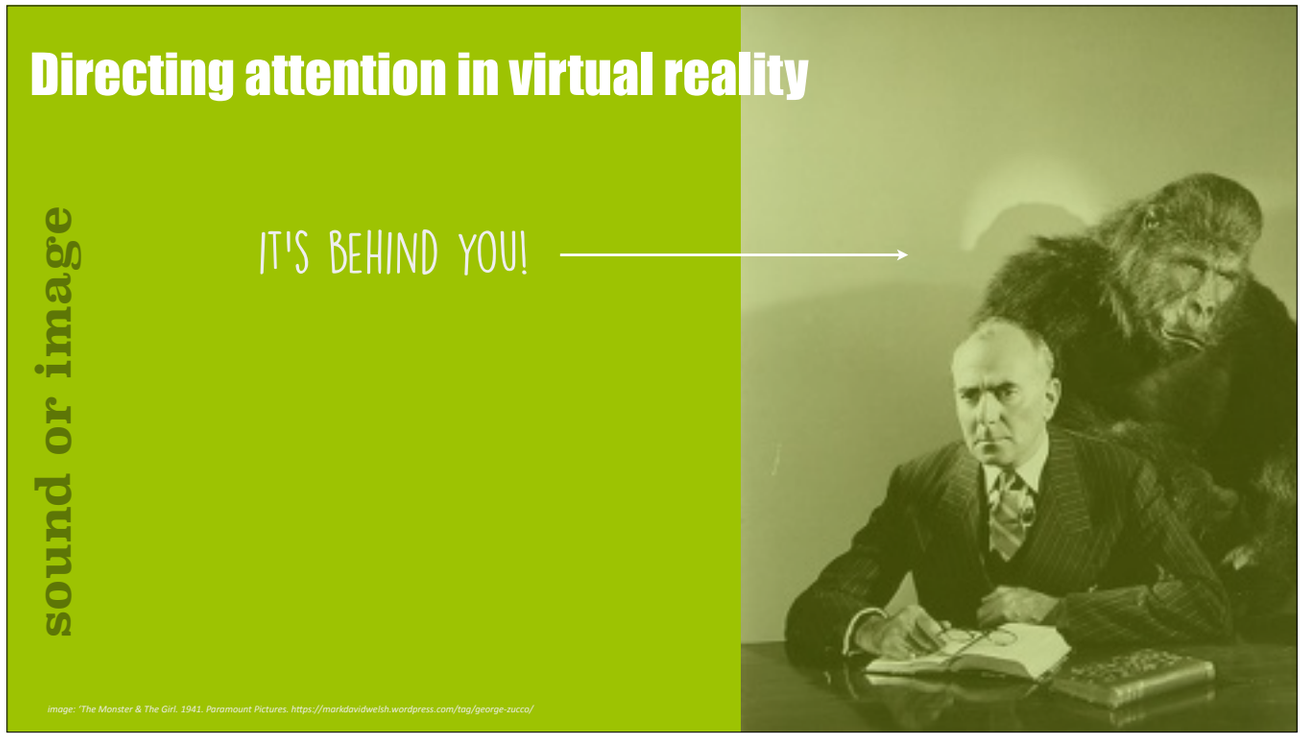


image: 'The Monster & The Girl. 1941. Paramount Pictures. https://markdavidwelsh.wordpress.com/tag/george-zucco/

John Mateer in the Journal of Media Practice, endeavoured to identify unifying factors within film and VR and explored the idea that “existing film production methods are considered in a manner adapted to establishing ‘presence’ in a CVR (cinematic VR) space”. One such consideration was the way in which the voice is used as a means to direct the audience to specific locations.

Mateer’s thinking is underpinned by work from the BBC Research and Development, UK. In their 2016 paper Sheikh, Brown, Watson and Evans explored the medium in their paper ‘Directing Attention in 360-Degree Video’. Although using 360-Degree film as opposed to immersive VR they were keen to test “unobtrusive techniques for directing the viewers attention within a 360 degree panorama”. This research was crucially supported by qualitative and quantitative data drawn from two key questions:



1. What attracts attention, what refocuses attention and what techniques can a filmmaker use to direct the attention of a viewer?
2. How does the distance at which action occurs impact the experience of the viewer?

The first of these questions is obviously of interest in the context of my research, although the second is also something to consider when directing within VR – where do you place your action for maximum effect depending on the level of comfort/discomfort you wish to induce based on the tolerances of broad interpersonal distances.

Directing attention in virtual reality

sound or image

BYSTANDER

TARGET

BENCH FOLK

image: 'Directing Attention in 360degree Video', Sheikh, A. Brown, A. Watson, Z & Evans, M. BBC Research & Development, UK. 2016

Clip	Summary	Cue 1	Cue 2	Cue 3
A_1	Motion across main characters	Bystander walks to target		
A_2	Motion across main characters with gestural cue	Bystander walks to target, waving		
A_3	Motion across main characters with audio and gestural cues	Target shouts "Alia"	Bystander responds with wave and "Hi"	Bystander walks to target
A_4	Motion of a main character following gestural and audio cues	Main character looks at target	Main characters talk about target	Main character walks to target

data: 'Directing Attention in 360degree Video', Sheikh, A. Brown, A. Watson, Z & Evans, M. BBC Research & Development, UK. 2016

Sheikh, Brown, Watson & Evans' research showed participants four clips with two characters having a conversation on a bench. The research team tracked the participants head movements to ascertain which method best directed attention - movement alone - where the bystander walks to the target, movement and gestures through to movement, gesture and audio cue.

Their results identified that where there was an audio cue that no matter where the audience were looking their attention shifted quickly, within 7 seconds of the cue. The researchers also found that using an audio cue where it referenced the target followed by the protagonists moving towards the target over 50% of the participants responded to the second cue (the mention of the target). In their research the results showed that where there was movement alone (A1) roughly 60% of the viewers to spotted the target after 64 seconds.

Where there was motion of the main character following gestural and audio cues it took only 47 second for a similar number to locate the target.

Of additional interest was the answer to one of their subsequent questions:

Are presence, immersion and enjoyment affected by characters in the content addressing the camera directly?

The participants revealed that they "felt more immersed in the content, and
enjoyed it more when they were acknowledged by characters in the scene. The technique of having the actors directly reference the camera as another character was effective"

Directing attention in virtual reality

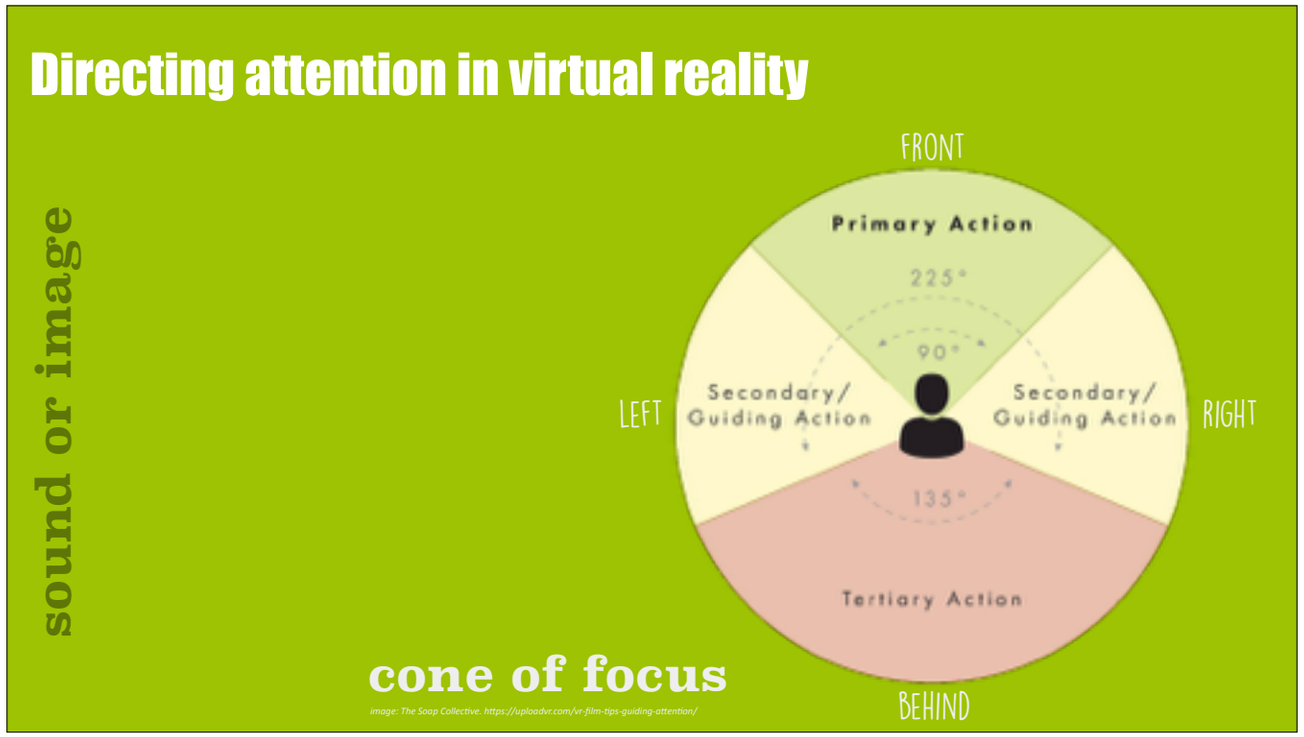
sound or image

image: www.pinterest.com/pin/200550989641855551

As a director of animation I have been keen to understand whether you need to approach the art of directing differently when working in VR, and more specifically when directing the voice or use of sound. Are there the same considerations, more things to consider? After all there is no one fixed point of view. One thing that is clear, is that the animated performance in the world of VR is as much defined by the voice as it is in traditional animation.

Just as it is with traditional animation so it is with animation in VR. Both Rich Evans from BAFTA winning studio Syncbox and Benedict Green at LA's Ecco VR - another specialist VR design studio are clear that when it comes to the animation pipeline, visuals come first, and audio follows. Evans and Green also state that 100% of their skill set from the traditional pipeline has been utilised in the VR environment, working with ProTools. According to Green "a great deal of our existing skillet translates perfectly". However to facilitate the spacialisation of sound - that is the placement of a sound that remains fixed in a space, but that 'moves' as the headset explores the 360 world giving a realisation of depth, height and distance with the attenuation of sound complicates matters.

The one thing that both Evans and Green were clear on - the differences are in the pre and post processes. Green states that: "Traditional treatments and techniques apply (eq, compression etc) but we add spatialising plug-ins to the chain. Spatialised VR has headroom challenges as limiting has to be used very sparingly if we're to avoid deforming our sound sphere. This can result in lower final output levels. Both agree that audio in VR also demands more in time. According to Evans the post process alone can be expected to take six times longer. So be warned!



So to cones of focus a term coined by the team at Soap Collective in LA. They looked at what attracts attention, what refocuses attention and what techniques a filmmaker can employ to direct the attention of a viewer.

Creative Director at Soap Collective Logan Dwight recognised that directing the viewers attention in VR is all about two things: placement of the viewer in the first place, and secondly the focus of action.

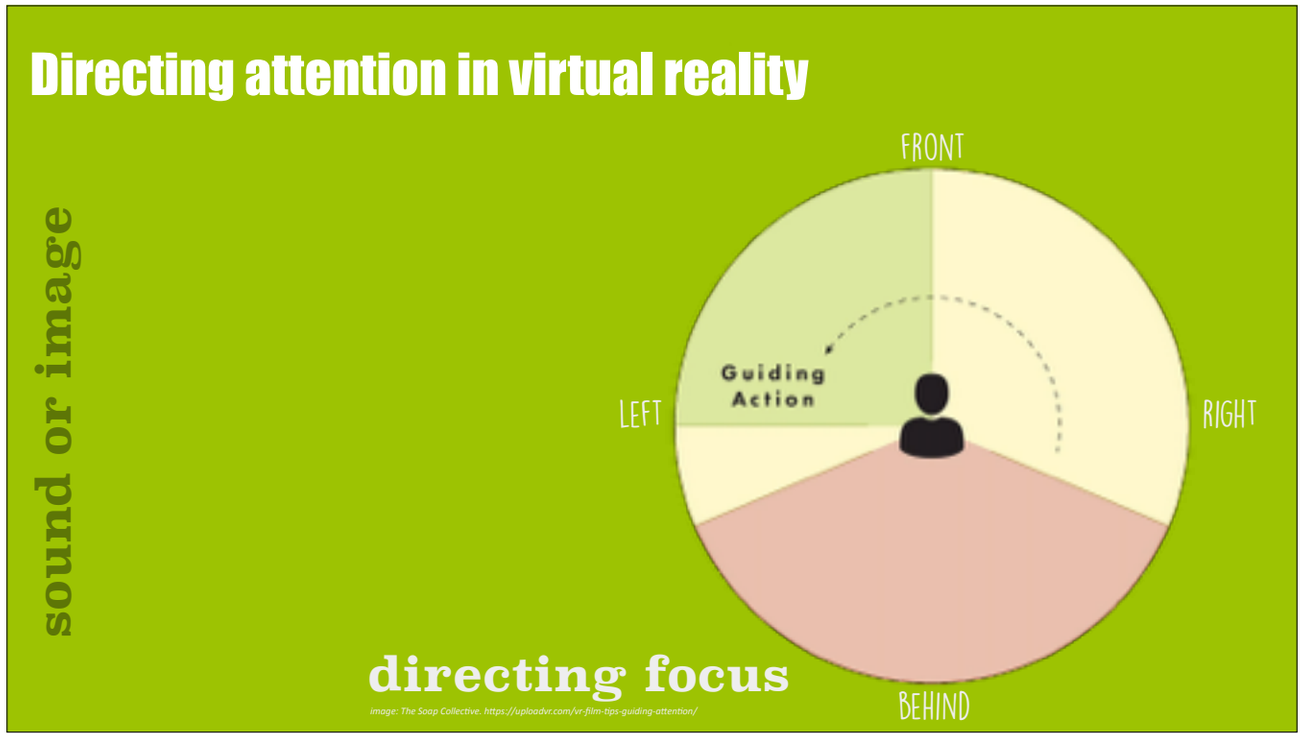
The assumption made by Dwight is that most viewers will be seated. If you have tried standing whilst in a VR environment you'll know why that's the case. With your viewer seated the sound and action can be designed to control their focus. Dwight terms this as "setting boundaries for the action". Soap Collective defined a cone of focus, dividing the 360 environment into sectors or cones with Primary, Secondary and Tertiary Action.

Dwight states that "When thinking about your story, think about choreographing the action so that it fits within the user's comfort zone. We like to break our action down into three categories: Primary Action, Secondary Action,and Tertiary Action".

Primary Action is where you want the viewer to focus on, and Dwight is clear that at the start of your story, this action should always be staged right in front of the viewer.

Secondary Action is the experiences "supporting interest". In film terms this is the off screen space. For Dwight this is "basically window dressing for the main story".

Tertiary Action is the experiences "minor supporting details", again still off screen space, with "low interest elements that do not draw attention". The ability to have this tertiary/off screen space is unique in 360 storytelling. It's existence keeps the viewer immersed. However in the main most viewers wont choose to turn fully around and look behind them (unless specifically directed, or the primary action is not engaging).



But what if that Primary Action needs to move? Placing the action within a 90degree cone in front of the viewer is limited and doesn't utilise the capabilities realised through 360 film-making. Here the use of Guiding Action, a term coined by Dwight, directs the viewers attention, and utilises the same techniques already applied in theatre and games. A change in lighting, through considered sound design or simply an action can all be used to shift this focus.

A consideration then for directing attention in VR is to design opportunities for characters to address the audience directly, to break the fourth wall. For a more engaging experience stories would do well to engage the audience as agencies for change, as active participants. Jesse Damiani terms this as the Builder-Participator Storytelling Paradigm.

But what considerations need to be made – stepping from the traditional world of animation in to a animated virtual world where the rules are still being defined.

Maybe we need to look at how everyone wandering into this virtual domain considers the use and direction of sound – from screenwriters, directors from theatre in the round, film and animation through to games developers and artists.

As Green puts it: "Audio is probably the most powerful, effective and subtlest of attention directing tools: we are very used to the principle of turning to face an identified sound source that's not in our field of vision. Thanks to HRTF (head-related transfer function), we can create very convincing attention directs that draw the users attention without being horribly unsubtle."

Since conducting my initial research I have realised that there is a gap between the research from Sheikh, Brown, Watson & Evans and that of the Soap Collective. What their researches have identified are where to effectively place sound and action to direct a point of view, and the time taken for this point of view to shift. However I believe this warrants further investigation as there appear to be gaps that need filling. For instance will a test participant react quicker to speech coming from the right hand side

Directing attention in virtual reality

working with sound

image: Dispatch. Dir. Ed Robels. Here Be Dragons

I wanted a greater insight into directing attention in VR, particularly in animation. But also wanted more to explore a few narrative experiences to see whether there were some common approaches being utilised before focussing solely on Aardman's 'We Wait'. As part of this research I analysed four different animated VR experiences. These helped me establish a base line from which to build.
Dispatch: Is an impressionistic narrative mini-series visualization of police call handler Ted's world. Handlers are trained to see with sound, and 'Dispatch' goes some way to interpret what Ted sees. Director Ed Robels overarching thought here was that:

"The minimalist reductive style comes down to a single thesis. If you're a dispatcher and you're speaking to somebody you have no idea what they look like and the sound is filling in your world view."

From the moment you enter the world of 'Dispatch' sound is used to focus attention. In fact in most instances sound comes first. From the footsteps, as the camera slowly tracks through a forest, steps crunching on dried leaves focussing the viewers attention to the front. A 'monster's' first thumping footstep to the rear forces a change in point of view with a self induced 180degree whip-pan. All the time the action slowly tracks forward, as if on a dolly. At other times Robels uses a simple tableau framing device with pulsating neon rectangles, action centered within these. Here dialogue is spacialized – with the voice over pushed to the front and an inner monologue overlaid and pushed to the rear. In a later scene of domestic abuse the action is again presented in a tableau, but then 'camera' takes control and the viewer is moved uncomfortably into the scene. The action moves on to a tracking shot through the house, with a neat wipe as we pass through a wall in to the bathroom. Here the viewer's point of view is fixed on a bathtub, while the action occurs to the left. Our attention is drawn to the left as the bathroom door is repeatedly kicked, sound forcing the viewer to look left within the left hand 67 degree sector of the cone of focus (more on that later), with a point of view taking up position behind the victim. Robels cleverly uses distinct punctuating sounds as his means to direct attention and establish desired points of view.
In the 360degree world of Dispatch the clever use of minimal visuals allows sound to take centre stage. As an animated piece the pipeline follows a traditional route – where sound, particularly dialogue is created first and the action matched to this. Though rather than using the intonation of dialogue to determine action the cadence is

Directing attention in virtual reality

working with sound

image: Dispatch. Dir. Ed Robels. Here Be Dragons

image: Allumette. Dir. Eugene Chung. Penrose Studios

Allumette: This VR experience leans heavily on traditional animation techniques, creating a charming hybrid narrative film where the CG animation is styled in such a way as to produce something that superficially appears to be stop motion. Serendipity according to its creators, who were shown an early iteration of rough ‘blocked’ animation and realised that this gave the whole production a completely different feel. Penrose Studio said of this: “it placed the characters in a miniature, hand-crafted, stop-motion world. We’ve dreamed about this sort of thing as children, with little toys, but they’ve never come alive and moved the way our characters move in VR”.

The vocal performance (emotives rather specific dialogue) has been doped (broken down frame by frame) to allow the animated performance to match the intonation just as you would expect with any animated film.

With Allumette the viewers point of view is primarily fixed straight ahead, in continuous wide shots, though of course the viewer has some ability to investigate the world albeit limited to a narrow cone of movement of approximately 30 degrees, giving an impression of some freedom. Cutting from scene to scene is managed via fades to white – using the flare from matches struck by the little girl central to the story. Similarly sound is focused to the front, where the action occurs, in stereo, though it is of course spacialised so that it moves from left to right as the viewer moves their head when looking around. However there are two occasions where the viewer’s point of view is directed to the left as the flying ship enters the scene. Effectively Allumette utilizes the 360 world simply to place the viewer at the heart of the action, in the same way that Plen an Gwari does. It doesn’t try to break to mould in any way, it is in essence an animated film in a VR world.

Directing attention in virtual reality

working with sound

image: Dispatch. Dir. Ed Robels. Here Be Dragons

image: Allumette. Dir. Eugene Chung. Penrose Studios

image: Alteration. Dir. Jerome Blanquet. Okio-Studio – Arte – Saint-George – Metronomic

Jerome Blanquet's 'Alteration' portrays the descent of protagonist Alexandro Burroughs into a world of AI. Both sound and image are manipulated with dexterity, in this vfx heavy live action experience that crucially utilises AI algorithms produced in conjunction with FAIR - Facebook Artificial Intelligence Research team in Paris. The experience uses these where 'real life' blends with the paintings of Burrough's partner, Nadia. The result is trip like and adds a surreal dimension to the experience.

Action is focussed in a primary cone (cones of focus a concept developed by Soap Collective), and a primary cone being a point of view within 90 degrees of the viewer, facing forward. The viewers attention is fixed on Burroughs who occupies centre screen, but sound cues from other characters including Elsa (AI in the guise of a woman) and Dr March come in from either right or left and from the very peripheries of the secondary cones of focus - each some 67.5 degrees. These two cones sit either side of the Primary cone. Otherwise sound, designed by Tim Dornbusch is treated in the same way that it is for a regular film only with spatialisation of stereo dialogue, and music fixed and non-focussed - existing in full 360. Syncbox's Rich Evans identified that music in VR must be fixed in order to keep at least one element grounded.

Blanquet isnt afraid to utilise some conventions carried over from traditional filmmaking either. Cameras track - and the viewer is carried along with this. Action cuts from scene to scene & he uses wipes. In one scene we also witness a straightforward cut from one shot to another - from over the top to effectively an eye level 3/4 shot. It works, but jars as the frame size doesnt alter, but it shows that a dramatic shift in point of view in VR can be tolerated, though perhaps with a little more thought it could be made to appear seamless.

Blanquet quoted in Variety says: The hardest part of creating a VR experience is the writing. Coming up with the mise en scene must necessarily embrace several worlds, especially theatre and gaming. As with all successful narratives, in any format, it boils down to the foundation — the script".

If anyone is interested FAIR may well open-source the algorithms used.

Directing attention in virtual reality

working with sound

image: Dispatch. Dir. Ed Robels. Here Be Dragons

image: Dear Angelica. Dir. Saschka Unseld. Oculus Story Studio

image: Allumette. Dir. Eugene Chung. Penrose Studios

image: Alteration. Dir. Jerome Blanquet. Okio-Studio – Arte – Saint-George – Metronomic

Dear Angelica: Created with Facebook's Quill, Dear Angelica gently guides the viewer through the thoughts of the Jessica as she reflects on the passing of her mother Angelica. In this experience diagetic sound is spacialized, with dialogue in the main simply in stereo. Attention is primarily directed by visuals first, augmented by sound, so that the swirling thoughts, seen in Jessica's scrawling handwriting, circle the viewer. Following these a viewer's attention is taken from one point of view to another. It's clear that Unseld wants you to look from place to place in a structured filmic way. Each frame of the animation is hand painted, though character movement is extremely limited, almost non-existent. Is it animation, probably not, but as the characters are hand painted they have the ability to form or disintegrate as desired. This is the device used by Unseld to overcome limitations of conventional edits. Scale becomes the means to alter shot types – from long shots with scenes placed at a distance from the viewer, held in virtual space, to close ups created by characters forming seemingly inches from the headset. What is clear with Dear Angelica is that it is image that leads direction and establishes point of view, and sound augments this.

Directing attention in virtual reality

CASE STUDY

'WE WAIT' DIR. DARREN DUBICKI

AARDMAN ANIMATIONS

image: 'We Wait', Aardman Animations

I wanted to go into more detail with the process for directing attention within VR, and went to Aardman Animation to interview the director of their VR experience ‘We Wait’, Darren Dubicki.

Interview with Darren

Directing attention in virtual reality

sound or image

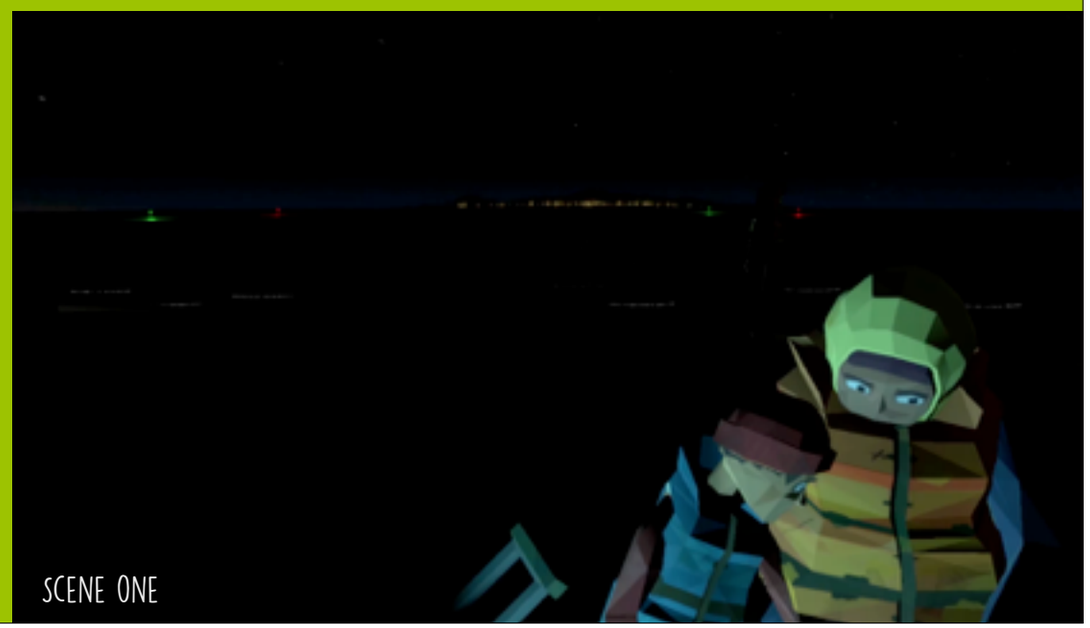


image: 'We Wait', Aardman Animations

Scene1.

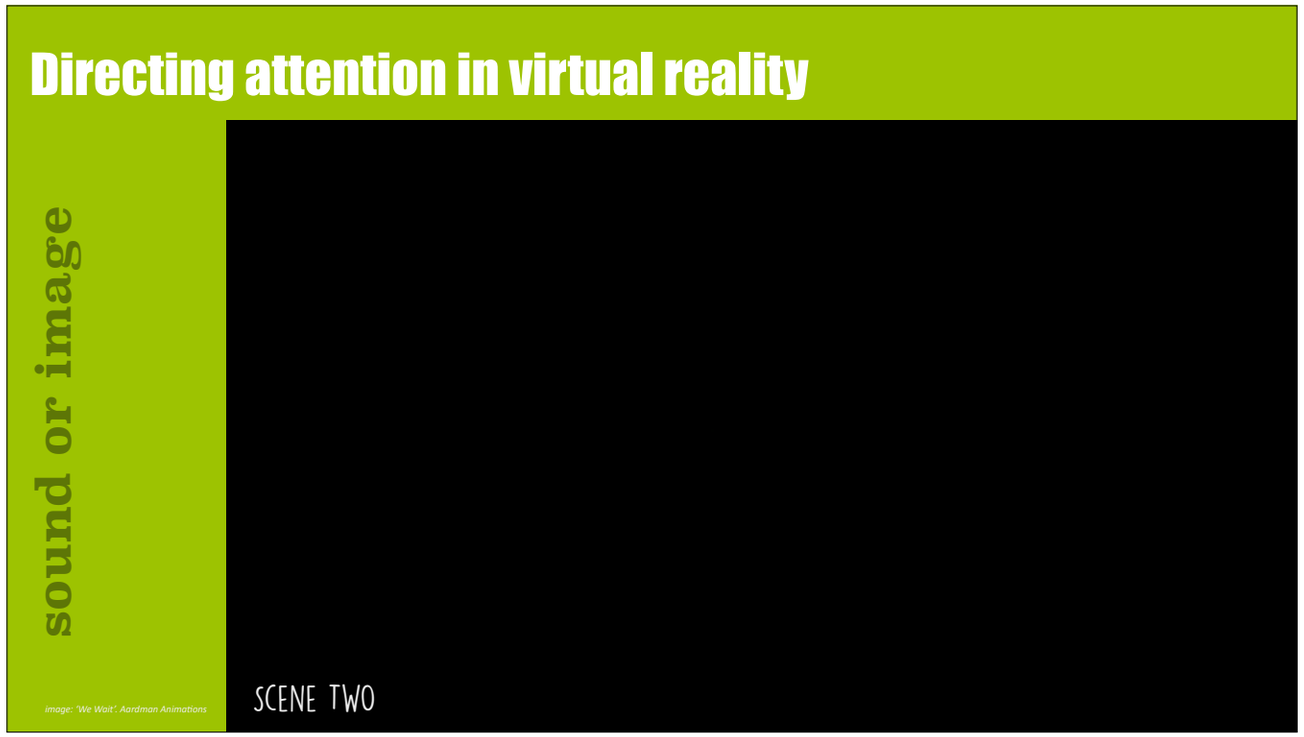
The film opens on a wide shot of a beach. Lights glow on the horizon - an island or mainland somewhere not too far away, but centrally placed in the Primary Cone. This centrality helps identifying it as the objective. On the beach all the action is staged within the Primary Cone, though non-active groups of characters inhabit the Secondary Cones to both right and left. To the rear the Tertiary Cone is a revealed as a low rock face with no particular detail. This featureless design encourages viewer to concentrate on engaging with the Primary Cone. A high altitude aircraft comes in from behind, out of the Tertiary Cone, inviting the viewer to look up, to explore the 360 environment. A motorbike passes right to left in Tertiary Cone of focus.

On the beach each group is lit by a single gas lamp, except one far to the right that is lit by embers of a fire. There is the low mumble voices, almost meditative or as if in prayer. Otherwise crickets and the sound of waves lapping break the silence.

The viewer is placed in the central group on the beach, with two women and a child. One man is pacing on the shoreline, on his mobile phone.

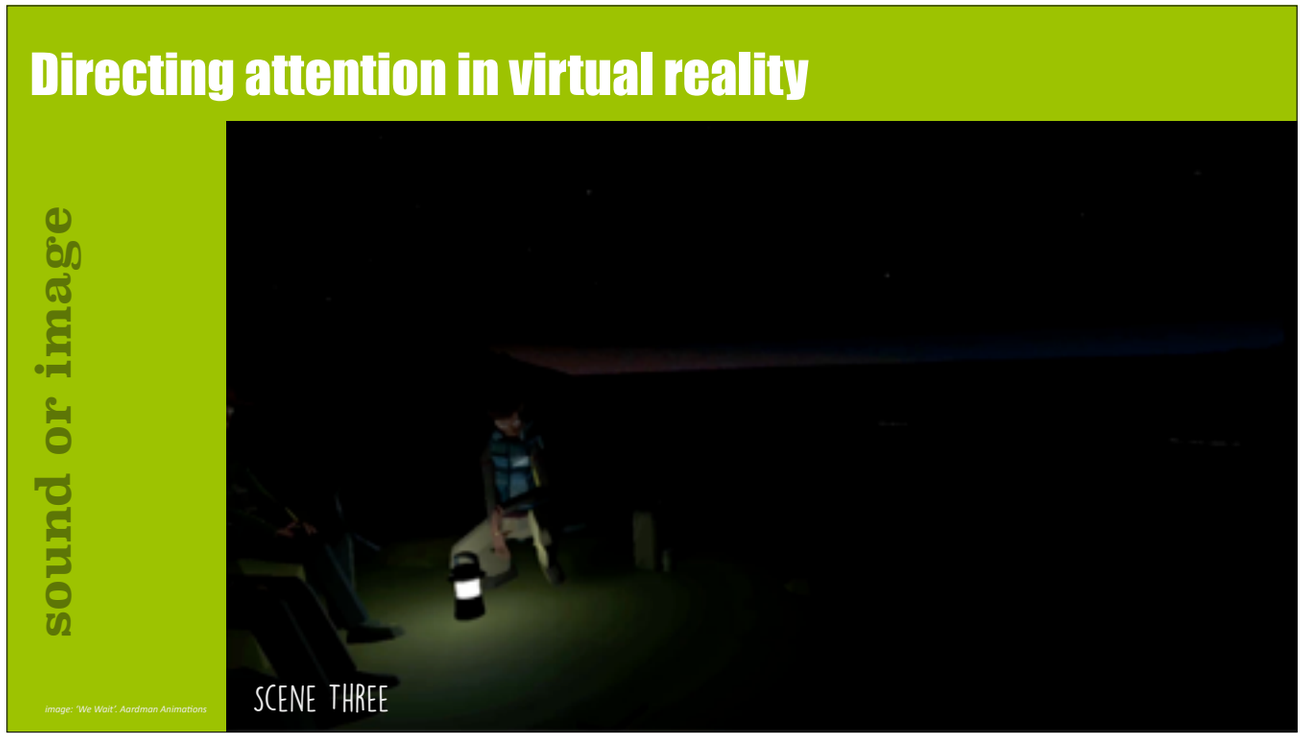
Until the viewer decides to interact or make a movement in this opening scene we are simply there, waiting with these refugees. Looking out to sea - into the Primary Cone. To trigger her engagement the viewer needs to lean towards her, make eye contact.
This woman, a mother with son in arms in front starts talking. The mother is hopeful. optimistic. She reveals her hopes for the future, that their crossing will be successful, their lives safer and more secure. This sets the tone for the beach scenes. They are scenes of hope. In the background her other son or son in law perhaps, walks towards our central group, mobile still in hand, though we are yet to hear from him. her daughter.

Hurriedly the reverb increases on the mother's dialogue as she begins to recount the story of their first crossing. The background and other characters in scene one fade



Scene 2.

We are on a small inflatable boat. It is night again. The scene is less saturated than previous one, adding to the sense that this is a recent but perhaps by necessity a fading memory. To begin with all we can see is the mother and young son, but gradually the other refugees become clearer, though they are more visibly muted that the mother and son. This slow reveal again invites the viewer to explore the surroundings. Count seventeen people on our boat, and looking around the viewer can see lights visible on either horizon - to left and right. Our mother and child are sat in front of viewer in Primary Cone. Here they are top lit, attention is focussed on them again if there was any doubt of where we should be looking. However being unsure of where we are to begin with is unsettling, that uncertainty adding to the sense of empathy. All this is new, unclear. The viewer left trying to make sense of the environment with few clues - just as the refugees are trying to make sense of their journey.



Scene 3.

We are returned to beach as in Scene One. The young man with mobile phone is sat with our central group, we are still with our mother and young son, while another man walks along the shoreline (in silence). Now the film shifts, this becomes in part his story, his future and his fears that will be shared in the next scene. However our focus remains on the mother and young son as if to indicate that it is material who is optimistic or who is afraid, these are common to everyone here. And we are invited to empathise with them. The witnessing of the drowning of the other refugees sets a sombre melancholic tone. Physically our refugees look more broken and you sense a rise in frustration as they wait for the smugglers to return with another boat. The young man remains positive that they will soon be at sea, after all, they will be paying the smugglers more money. He gets up, looking ahead and out to sea and slowly walks towards the shoreline. As before everything else fades away, leaving him as the focus of our attention as we transition into the boat. And just as before there is an increase in reverb, this is his inner monologue. The sound of gently lapping waves increase in volume as one scene transitions into the next

Directing attention in virtual reality

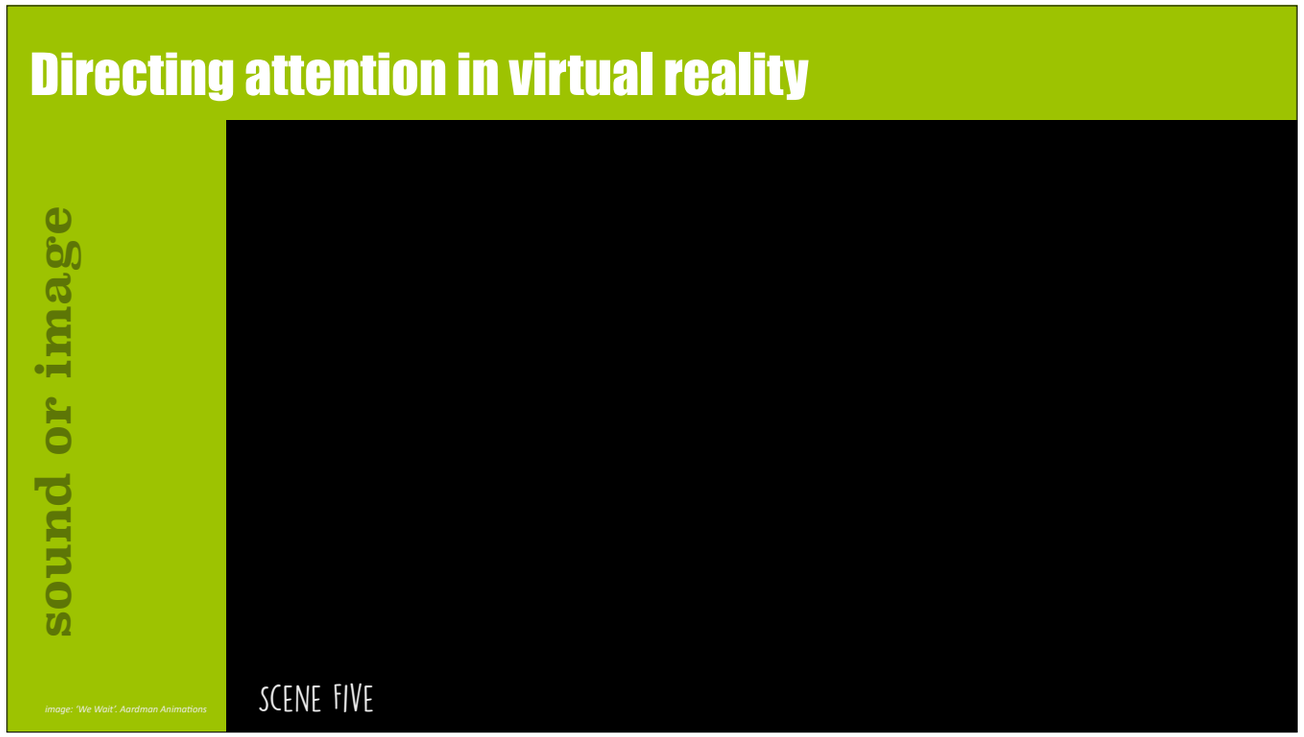
sound or image

image: 'We Wait', Aardman Animations

We are back on a small inflatable boat again. And it is night again. Our focus is soon shifted from the young man to the mother and child with them again sat in front of viewer in the Primary Cone. They are top lit, and therefore more prominent than the others who recede into the background gloom but we are invited to understand the fears of the young man as he continues with his thoughts. He explains the difference between being rescued by the Greeks as opposed to the Turks, the implications. A Greek rescue means hope and freedom. A Turkish rescue means being taken back to the beach. He tries to be brave, ‘what good is it to cry’, claiming that no-one won an award for bravery. His monologue focusses on his inner fears and this is intensified as other men on the boat raise their voices, compelling people to sit down, to keep the boat still, avoid the rocks!

There is a rising tide of anxiety. From behind a ships fog horn sounds in the Tertiary Cone. The viewer has to look to backwards to see an anonymous ominous black warship looming up from behind. Suddenly there is an order to ‘STOP’ from the ship. It is repeated and a searchlight from the Tertiary Cone floods the group with light. Panic ensures and this action working in combination with the sound again encourages the viewer to look all around as if to make sense of what is going on.

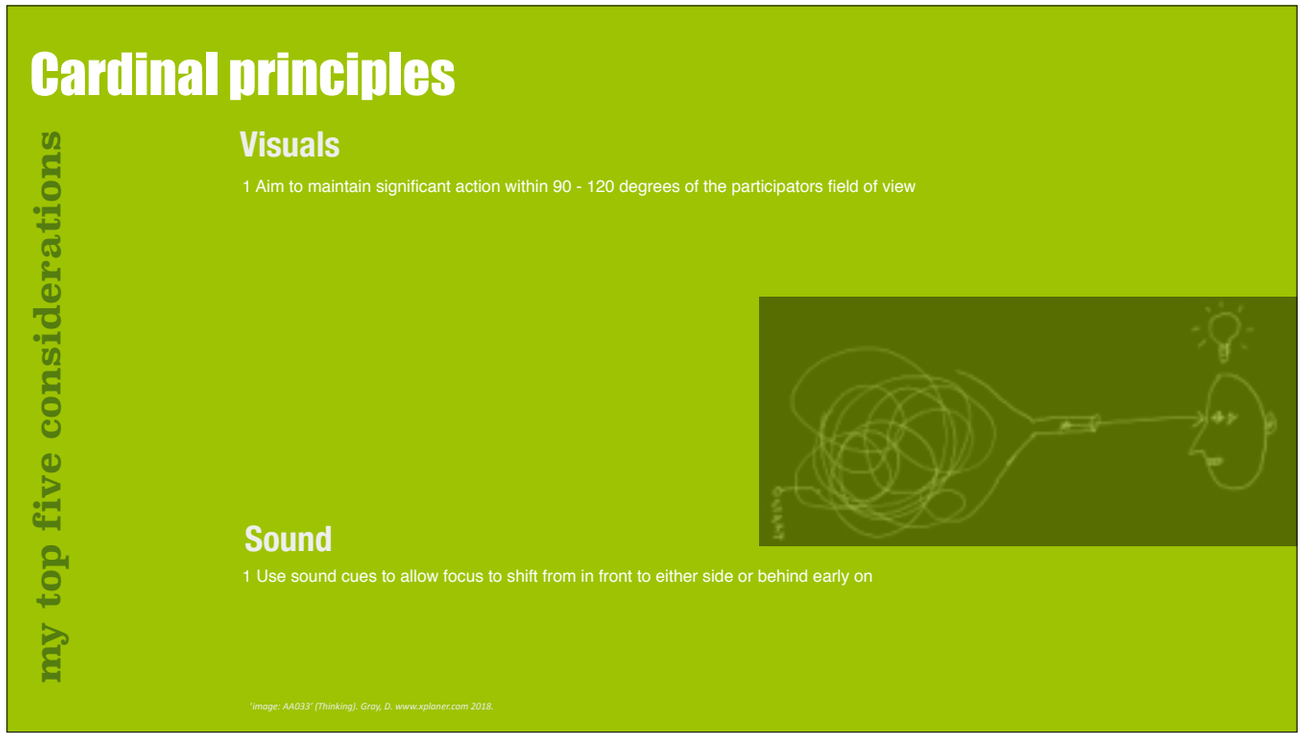
The intensity of light from the search light builds as the young man delivers the devastating line ‘It was the Turkish Coastguard’. This Intensity continues to build and serves as a dissolve/fade into next scene.



Just like our refugees we are returned to the beach. Again our mother and young son are right in front of us. The young man with mobile phone is sat with us, to the right of the periphery of the Primary Cone of Focus
The sound design returns to the low mumble voices, of crickets and waves lapping. Another motorbike passes right to left in the Tertiary Cone of focus.
These events are testing the optimism of the refugees. Our young man tries his best to stay positive, but there soon unfolds an argument with Yana, either his wife/ girlfriend or sister (their relationship is uncertain). Yana does not want to risk it again, two failed trips are enough, emotionally and financially. She will walk to Athens. Her fears are that they will be 'in the hands of the sea' to which he seeks to reassure her by saying that they will be 'in the hands of God'. He reassures her it will be ok, that even if she falls in the water, and despite not being able to swim himself, he will make sure she is safe. Again we see the switch from the internal emotional conflict and doubt, to external confidence and optimism.

Through all the set backs, failed crossings, drownings, exploitation and hardship we ultimately we see the pathetic situation that they are in when the young man takes Yana's hand and says 'It's ok. I will protect you'.

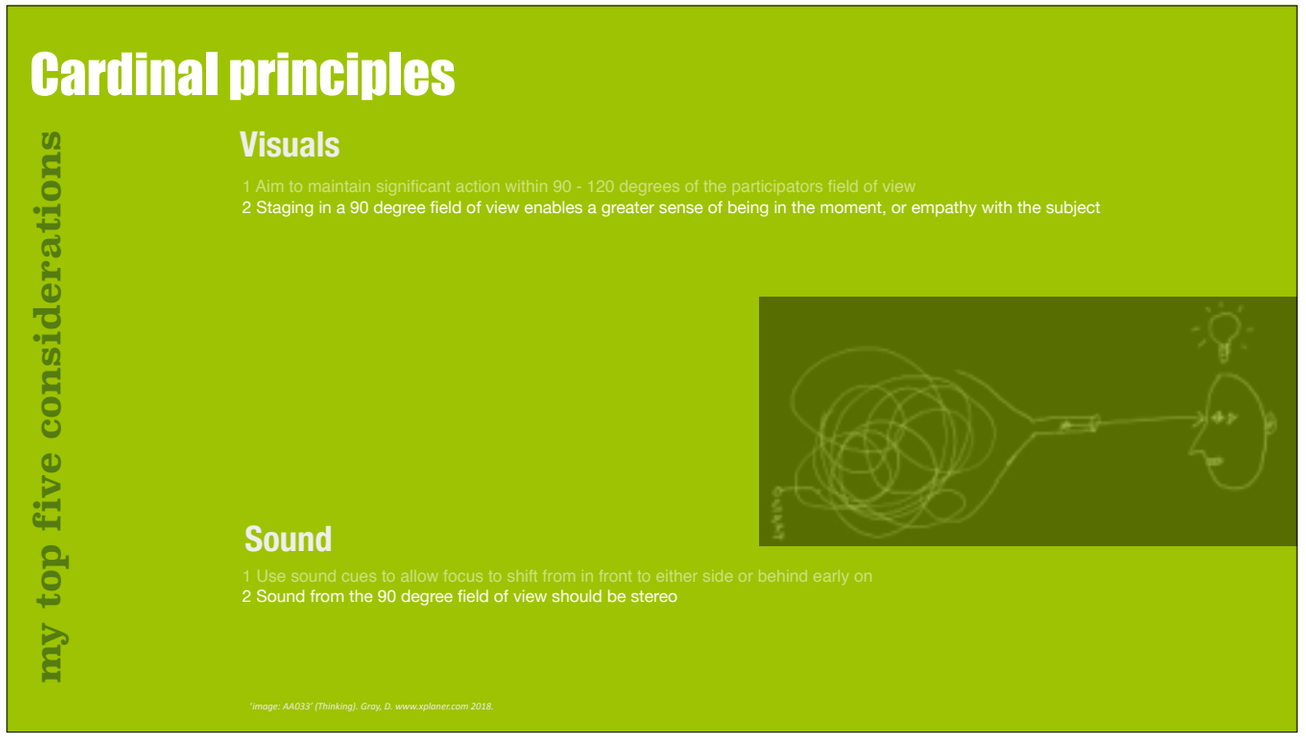
But we know, as do they that no-one can give such assurances as another boat arrives, quietly, discretely from the left.



So, from research and observations thus far I have arrived at what I am calling my cardinal principles. These are considerations that I would steer directors towards when developing a narrative VR project.

Visual considerations:
1 Aim to maintain significant action within 90 - 120 degree field of view in front of the user.
VR should immerse, but should not turn the user into a 'Wimbledon spectator', endlessly turning around to follow the action!
This doesn't mean there should be no 180 or similar turns, just that these should be used sparingly and definitely not gratuitously.

Sound considerations:
1 Use sound cues to allow focus to shift to either side or behind early on - this enables orientation (this includes up and down)
Onboarding is still important. People need time to take it all in and feel a sense of ownership and comfort, as well as to understand the VR basics of 'wow, I can look anywhere'. Sound really helps here. Very comparable to theatre - almost all plays deploy onboarding. We seldom here the first line as the curtain rises! It's usually a few minutes with one actor moving around the set as the audience takes it all in. Users should never be made to feel like they've missed an important detail or are falling behind.



Visual considerations:

2 Use of the 90 - 120 degree point of view enables a greater sense of being in the moment, or empathy with the subject
So much of the empathy effect of VR comes down to a) simply being in the same actual space as the VR characters and b) that fact that being looked at by characters in VR seems to have much more impact than the same trick in standard 2d footage. Ecco recently did post-production for a piece called 'State Power' (Dir. Jeff Stanzler). This tough bit of viewing drops you into a kind of Gilead America (Hand Maid's Tale) where you have already been 'black bagged' and caged. The user experiences it from the point of view of a prisoner, so I feel it's the sum total of all the VR methods that immerse here: being tortured with loud music, being 'made' to walk in shackles round a cage, whispered ideas and dissent from your fellow captives.

Sound considerations:

2 Sound from the 90 - 120 degree point of view should be stereo
However the distinction or importance of stereo files vs mono is reduced in VR audio. Traditional cinematic technique, for example, states that dialog should be mono and ambiences are stereo, but this only because of the core '2.1' nature of cinemas. You can't really use broad or hard panning, especially for dialog (with minor exceptions) so we need something other than music to explore or leverage the stereo field. In VR, if you dont strip a stereo file to mono (rare) then it'll be an ambience which will be deployed in a stereo panner so one can move the L and R channels independently and create a more encompassing feel. It is possible to do the same thing however with a few mono ambiences for an equal if not superior effect.

Cardinal principles

my top five considerations

Visuals

1 Aim to maintain significant action within 90 - 120 degrees of the participators field of view
2 Staging in a 90 degree field of view enables a greater sense of being in the moment, or empathy with the subject
3 Further action can emanate from either side or even behind after audio signposting

Sound

1 Use sound cues to allow focus to shift from in front to either side or behind early on
2 Sound from the 90 degree field of view should be stereo
3 Ensure that sounds come from one side or the other, but not at the same time

'image: AA033' (Thinking). Gray, D. www.xplaner.com 2018.

Visual considerations:

3 Further action can emanate from either side or even the rear after audio signposting.
Audio attention directs are the more powerful and effective way to turn a head. A good audio 'pull' can easily turn a secondary cone into the new primary. Without the audio attention pull, the director is running the risk that important action will go unnoticed or partially missed

Sound considerations:

3 Try to ensure that sounds come from either side, but not at the same time. You want to enable clarity and not confusion - your viewer may be stood not seated. Using HRTF (Head Related Transfer Function) convincing 3D audio can direct viewers attention subtly.
There are two concepts at play here: 1 - amount of sound sources vs effectiveness, and 2 - all directions are equal, but some are less equal than others! Ecco VR confirm that they definitely found that there is a diminished return curve where quantity of sound directions are concerned. 1 - Anything over 5 or 6 sounds and the viewer seems to give up trying to perceive the directions. Less, as ever, is more. 2 - Azymuth differences are very powerful and should be favoured. Even small amounts of pan in the primary cone are still very audible. To Ecco, the most powerful pull is from 'behind and to the side'. We are so used to the human concept of 'stimulus is not in cone of vision, turn and acquire' (friend taps you on the shoulder, car horn in blind spot etc.) that it translates, wonderfully intact, to VR. The least useful azimuth pull is the 6, or perfectly behind. Although HRTF treatment will try and 'sign' the sound as being on your 180, this can only be through attenuation and eq and not from direction as, effectively, there is none. Green always recommends the avoidance of placing a source 'on the six' and will advise directors against it. He sometimes 'cheats' an object on the six and actually place it at the 5:30 as it were, just to duck this issue.

Cardinal principles

my top five considerations

Visuals

1 Aim to maintain significant action within 90 - 120 degrees of the participators field of view
2 Staging in a 90 degree field of view enables a greater sense of being in the moment, or empathy with the subject
3 Further action can emanate from either side or even behind after audio signposting
4 Action alone will not guarantee a shift in viewers POV

Sound

1 Use sound cues to allow focus to shift from in front to either side or behind early on
2 Sound from the 90 degree field of view should be stereo
3 Ensure that sounds come from one side or the other, but not at the same time
4 Sound cues in combination with action to promote a faster shift in POV from the viewer

'image: AA033' (Thinking). Gray, D. www.xplaner.com 2018.

Visual considerations:

4 Action alone will not guarantee a shift in viewers POV.
Even with attention directs in the action and in the sound, people will blithely look the wrong way. Or not at all! One interesting exception or extreme case with attention directing is VR games. In a game, it's not unusual for important items, targets, directions etc. to be physically signposted with actual pointing arrows. You can't get more literal than that. These are usually in the HUD tradition and help users to not miss vital aspects of the experience.

Sound considerations:

4 Use sound cues in combination with action to promote a faster shift in POV from the viewer
The 'teamed-up' effect is practically infallible. Always a good idea where possible to bring the sound cue in ahead of the visuals - this gives the user an extra beat to spot the action before it's evolved too much. Ecco used that technique in the recent piece for Zoom's new VR mic. They visually overlapped to help the viewer, but still brought in audio cues ahead of the visual cards to assist the attention direct.

Cardinal principles

my top five considerations

Visuals

1 Aim to maintain significant action within 90 - 120 degrees of the participators field of view
2 Staging in a 90 degree field of view enables a greater sense of being in the moment, or empathy with the subject
3 Further action can emanate from either side or even behind after audio signposting
4 Action alone will not guarantee a shift in viewers POV
5 Cinematic conventions can be applied, though certain cuts don't!

Sound

1 Use sound cues to allow focus to shift from in front to either side or behind early on
2 Sound from the 90 degree field of view should be stereo
3 Ensure that sounds come from one side or the other, but not at the same time
4 Sound cues in combination with action promote a faster change in POV from the viewer
5 Ensure music or atmos are omnipresent and not spacialised

'image: AA033' (Thinking). Gray, D. www.xplaner.com 2018.

Visual considerations:

5 Cinematic conventions can be applied. Camera moves, cuts, wipes all work. Be mindful of changing 'lenses' to avoid jarring cuts.
Earlier VR tended to hold tight to the 'single view, long take' approach, but in the past few years good and creative directors have expanded this style to include a variety of shots. VR is easier on the viewer when you use fades and dissolves (jump cuts are too confusing) and when you avoid 'angle/counter-angle' techniques. Users don't like to feel 'teleported' or flung around the experience. Ecco encountered some very interesting cuts and POV changes in a recent piece for Diageo/Jaunt. These demonstrated a great use of multiple camera angles and view points but not a jarring experience

Sound considerations:

5 Ensure music is omnipresent and not spacialised
Head locked music is vital in most experiences. It helps to make a sonic distinction between the spatialised and the non-spatialised. It also leverages our pre-existing understanding from cinema that music is an accompaniment, a marker, a signifier, but not from the movie 'world'. A useful and important exception to make here is when it is from the movie world! The old trick of moving a song from diegetic treatment (background music in a pub) to non-diegetic, full score works incredibly well in VR. Is there a hi-fi set-up in the opening scene? Great - take the intro music and 'move' it to the in-shot speakers (with appropriate eq treatment, of course) as we fade up on the scene.

references & acknowledgements

References:

‘Plen An Gwari - The Playing Places of Cornwall’. Coleman, W. Golden Tree Productions. 2015
‘Directing for Cinematic Virtual Reality: how the traditional film director’s craft applies to immersive environments and notions of presence’. Mateer, J. Journal of Media Practice Vol 18 Issue 1. 2017
‘Directing Attention in 360-Degree Video’. Sheikh, A. Brown, A. Watson, Z & Evans, M. BBC Research & Development, UK. 2016
‘Film Theory: An Introduction through the Senses’. Elsaesser, T. Haggler, M. Routledge 2015
‘The Assembly: Telling Stories in Virtual Reality’ nDreams website. 2016 (accessed 28/5/18)
‘The Beginners Guide to VR Scriptwriting & Storytelling’. Damiani, J. VR Scout website. 2017 (accessed 4/6/18)
‘These VR Film Tips Show How To Direct Audience Attention’ Dwight, L. uploadvr.com website. 2016 (accessed 4/6/18)
‘VR Cinema is Here - and Audiences are in the Driving Seat’. The Conversation website. 2016. (accessed 18/6/18)
‘Towards a Narrative Theory of Virtual Reality’. {S.Louchart, R.S.Aylett} @salford.ac.uk
https://techcrunch.com/2017/11/21/with-dispatch-here-be-dragons-pushes-narrative-vr-storytelling-in-bold-new-directions/ (accessed 5/7/18)
https://www.fastcompany.com/40444098/how-facebook-used-ai-to-make-the-trippy-effects-in-this-vr-film (accessed 11/7/18)
https://variety.com/2017/film/asia/busan-virtual-reality-conference-draws-big-crowds-1202590306/ (accessed 11/7/18)

Images:

www.bewnanskernow.org/cornish-culture-blog/invitation-to-the-launch-of-plen-an-gwarri-the-playing-places-of-cornwall-project
www.wareable.com/vr/best-playstation-vr-games-2144
uploadvr.com/vr-film-tips-guiding-attention/
www.pinterest.com/pin/200550989641855551
‘Dispatch’. Here Be Dragons
‘Allumette’. Penrose Studios
‘Dear Angelica’. Oculus Story Studio
‘Alteration’. Okio-Studio – Arte – Saint-George – Metronomic
‘We Wait’ Aardman Animations
‘The Monster & The Girl’ 1941. Paramount Pictures https://markdavidwelsh.wordpress.com/tag/george-zucco/
Image & data from ‘Directing Attention in 360-Degree Video’. Sheikh, A. Brown, A. Watson, Z & Evans, M. BBC Research & Development, UK. 2016
‘AA033’ (Thinking). Gray, D. www.xplaner.com 2018.

Thanks to:

Darren Dubiki. Aardman Animations, UK
Ashley Cornes. Aardman Animations, UK
Logan Dwight. The Soap Collective, LA
Benedict Green. Ecco VR, LA
Dan Sheridan. nDreams, UK
Pete Billington. Fable Studio, LA
Prof. Chris Morris. Falmouth University
Denzil Monk. Falmouth University
David Smithers. Falmouth University
Andy Smith & Doug Brown. Falmouth University
Alex Falconer. UK