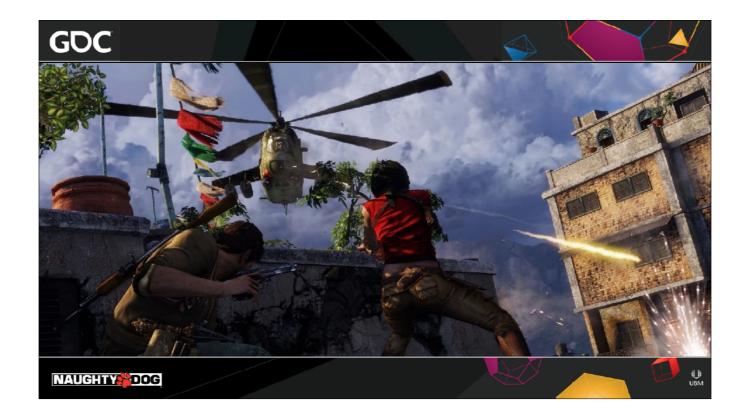
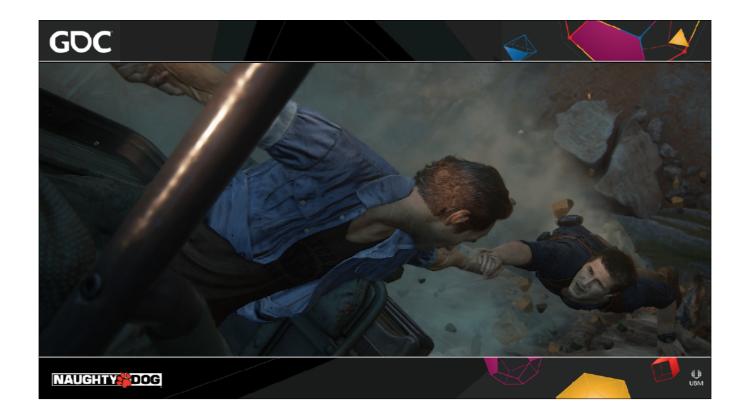


Good afternoon, and welcome to Realistic Performances in Games.

Fair warning, I've been told this talk is a bit "chewy", so...STRAP IN!



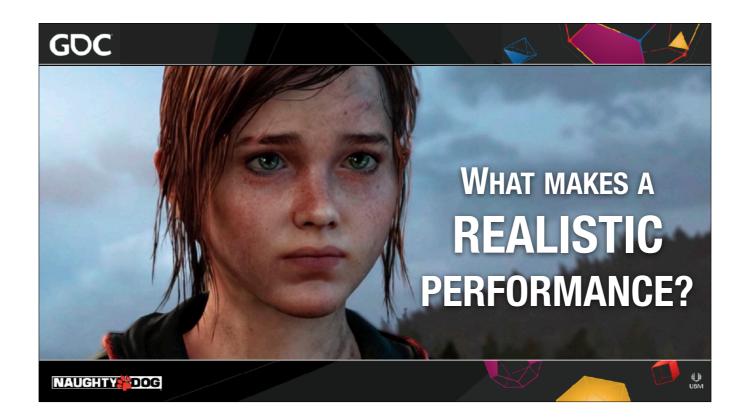
There's a perception out there that the average game is mindless entertainment, with at best a thin story, cardboard cutout characters and stilted acting.



But many companies have been working tirelessly against that stigma. At Naughty Dog, where I've been for over 7 years, we work hard to ensure our performances are believable, emotionally relevant to the Player, and blend with our gameplay as seamlessly as possible.



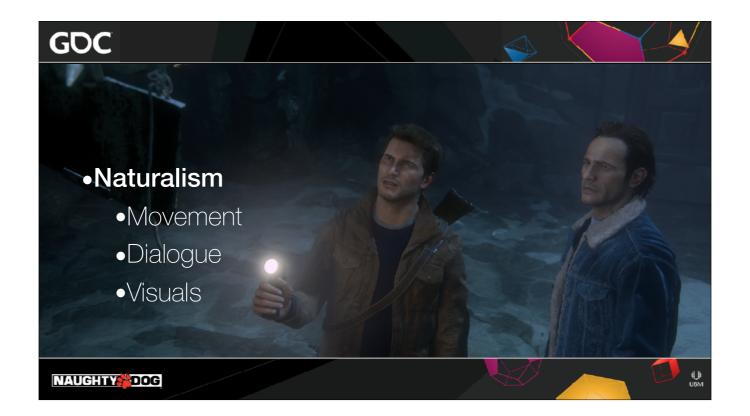
I'm Ryan James, and I work at Naughty Dog in a somewhat unique position—as Lead Editor. Aside from writing and directing some of our content, it's my job to help the rest of the team create and maintain performances that feel like what you'd expect from an edited movie—even when they're fully interactive.



But what makes a realistic performance?

$\{ \hbox{\bf SLIDE} \}$

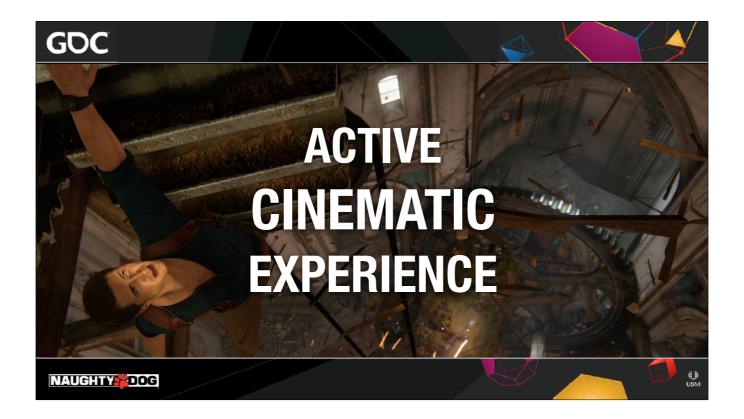
At Naughty Dog, it means a character needs to feel vulnerable, nuanced, and human.



Which means we're often focused on how **NATURAL** every moment feels.

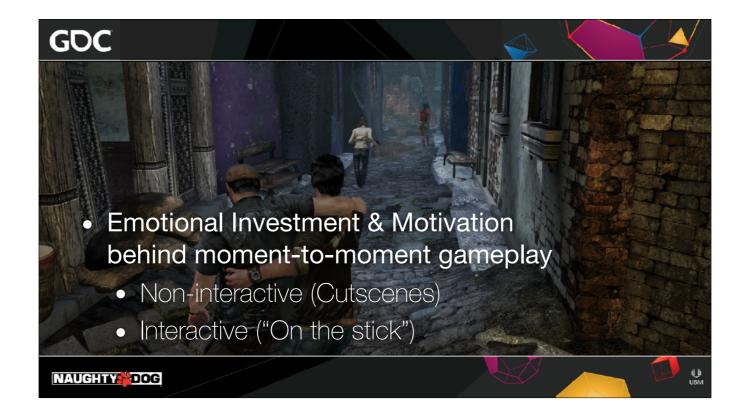
Meaning how the character MOVES, and how they SPEAK, needs to feel physically and emotionally relatable.

And of course, they also have to $\underline{\mathsf{LOOK}}$ realistic—which takes labor from $\mathbf{everyone}$ to get right.



Which is our entire goal. Naughty Dog's games strive to intertwine story and gameplay, so the Player experiences a "seamless" performance, which we call the **ACTIVE CINEMATIC EXPERIENCE**.

{beat}

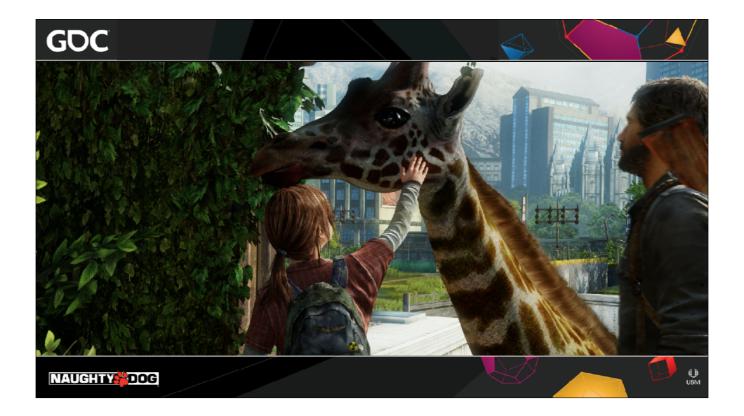


We like to provide the Player with **EMOTIONAL INVESTMENT & MOTIVATION** behind the moment-to-moment gameplay.

This is largely done through **NON-INTERACTIVE CUTSCENES**, but because we strive to keep the Player in control as often as possible, a lot of those drama and character moments end up paying off "**ON THE STICK**", as we call it.

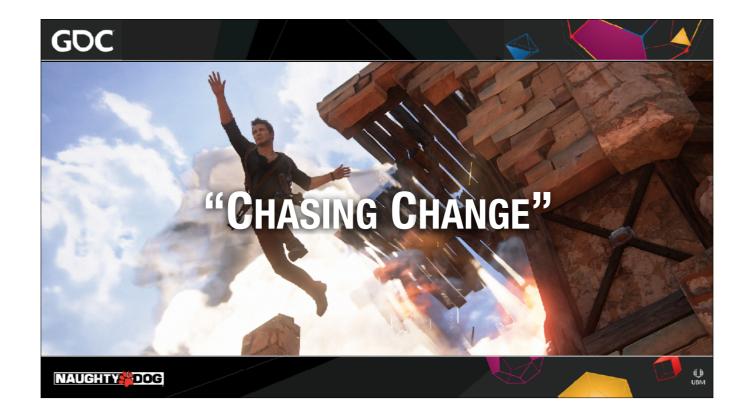


But figuring out which beats belong where—and pulling off the execution of them—requires a lot of iteration. And at Naughty Dog, one of our biggest strengths is our willingness to re-tool or even throw away content—including captured performances—if it will make the Player's experience better.

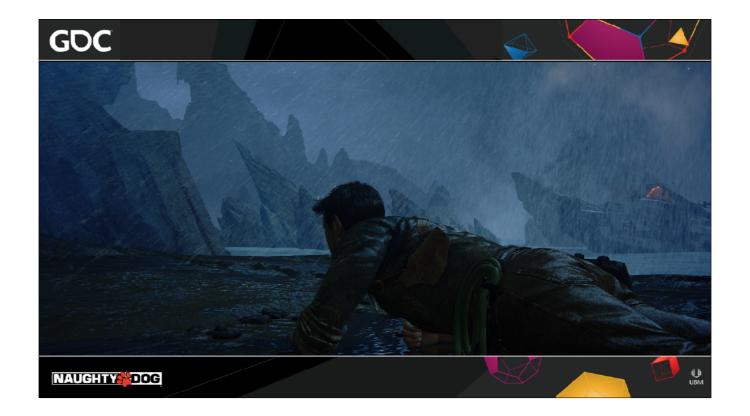


Even when deadlines are extremely tight, our company culture encourages people to speak up if they think something isn't working, and take advantage of our tools which enable quick alterations to try out new ideas.

That means all of the narrative content that we capture is prone to changing, even after it goes into the game. And, sometimes these alterations happen so quickly that their trickle-down effects get lost in the daily shuffle.



Thus it is up to the Editorial team to constantly <u>CHASE CHANGE</u>, tracking and managing each edited performance in its various evolutions over the course of development. In order to ensure that when the Player gets their hands on the game, our character's interactions still feel like the authentic, human moments we'd intended them to be when we first captured them.



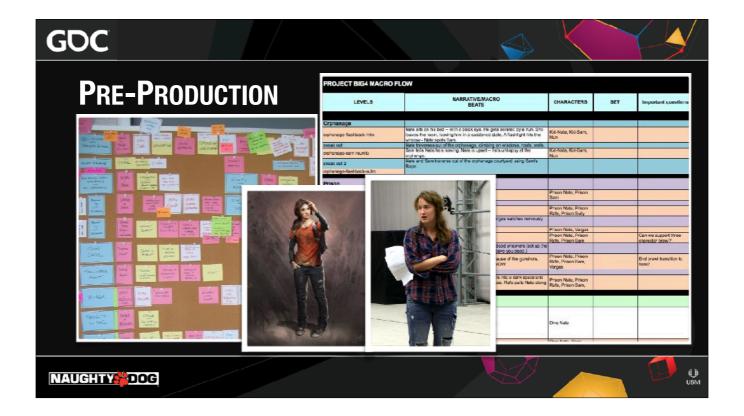
So today I wanted to share some of the techniques we've used—and challenges we've encountered—when creating realistic performances in games, while allowing them to change over the course of development.



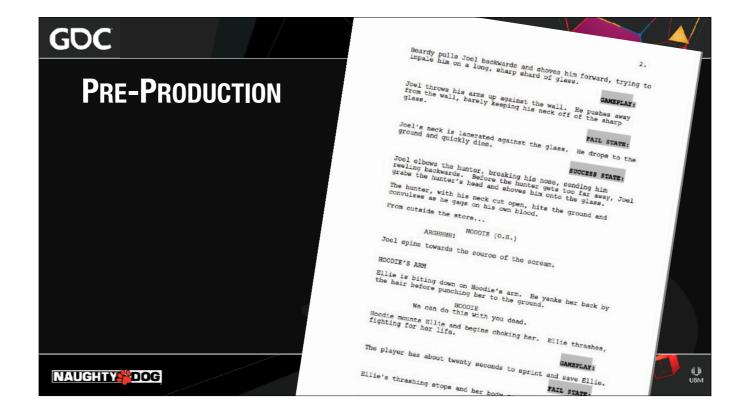
Oh, and just a heads up: <u>SPOILER ALERT</u> for several of our games.



Here at Naughty Dog we don't develop the story second. Since we shoot and implement narrative material in small batches—capturing progressively with the Actors for easily a year or more—our story in fact comes together right alongside Design.



Of course, it takes months of Pre-Production before we can shoot anything. Our directors and writers first juggle developing **Story Macros** with Design, overseeing the **creation of Characters** that offer varied gameplay mechanics, and **casting Actors** to embody those characters.



But as soon as we have enough material written, we jump in head first with our Actors to capture story beats via motion-capture and dialogue recording.

Often, we focus on a **Demo** of some kind: a small slice of the game that establishes our tone and explores the game's developing mechanics.



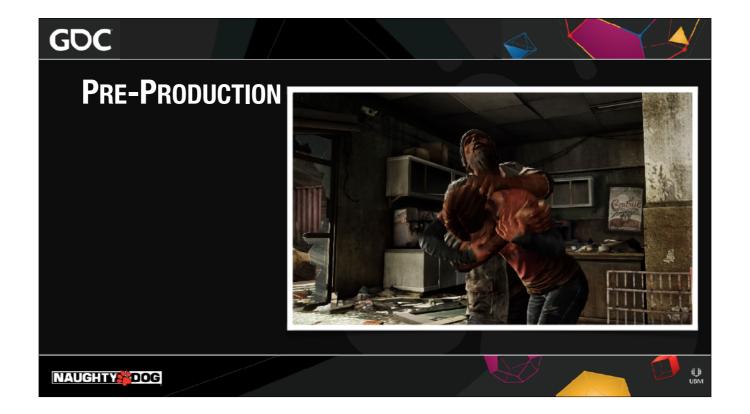
Right away, the game starts evolving because of these initial shoots. Not only do they help our principle Actors start infusing their personality into the characters, but their choices this early on allow the team to adapt to discoveries we've made with the Actors on stage.



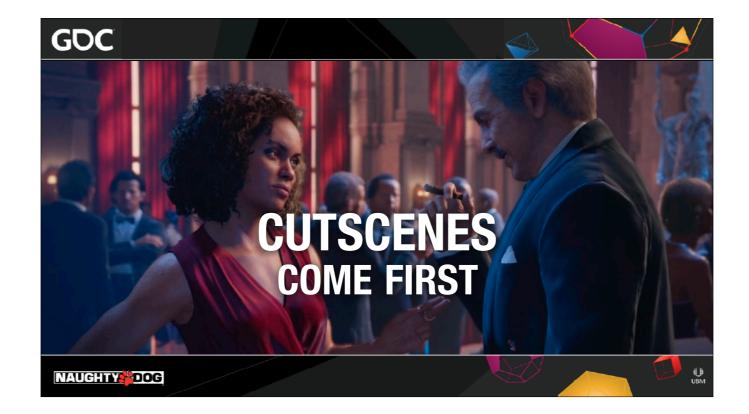
• For instance, on the very first day of shooting The Last of Us, we began with a scene where Joel and Ellie are ambushed during their cross-country journey. Ellie was originally pulled from their pickup truck and pinned helplessly by an enemy, making her a quasi-damsel in distress for Joel to save.

[SLIDE]

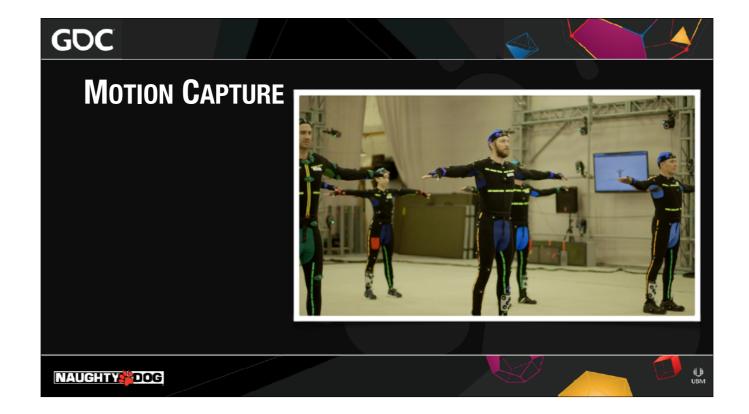
But Ashley Johnson (who plays Ellie) had a desire to "fight back" against her attacker. So we shot it that way, leading our Directors to go back to Design and rethink the character's overall vulnerability over the rest of the game.



Though we wanted to reinforce the need for the Player, as Joel, to protect Ellie from harm, it still worked better for her to became far more capable in combat situations than we'd originally conceived.



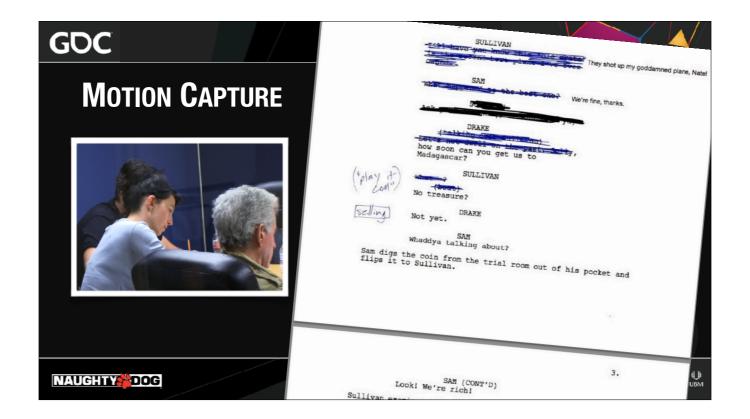
As we continue to shoot, **Cutscenes** are initially our top priority. Since these are **non-interactive beats**, reserved for the **more delicate emotional and character moments**, they require as much Animation time as possible to hit the high level of polish required.



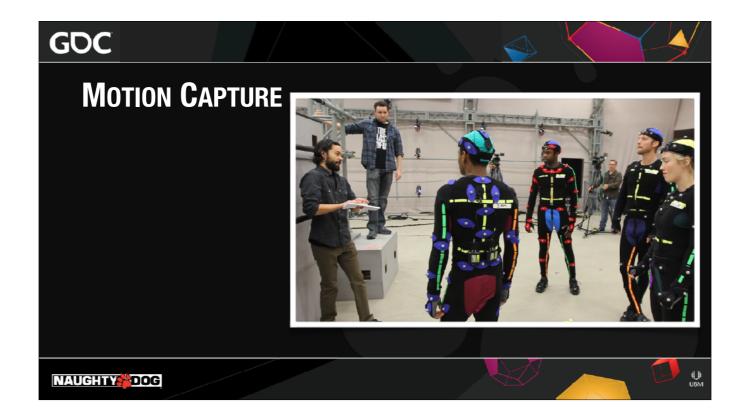
We've spoken about our capturing process before, but there's a couple things I feel we do that are essential to capturing the best performance possible. Because, as an Editor, if all I had to start working with was **stale or unnatural material**, **there'd only be so much we could do** in Post to make it better.



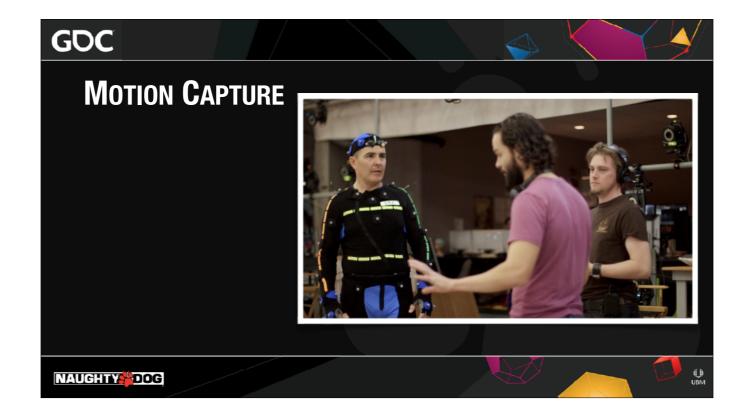
A lot goes into how we create scripts that give Actors honest material to play, but I sadly don't have time to touch on that today. However, I'd like to point out one strength in that process: every phase the material passes through is a collaborative effort. From the start, our Creative Directors encourage input from the team, and our Actors, either before a shoot or during our full-day rehearsals.



Especially with our Actors. We even do a table read to break down the script before we rehearse. If something doesn't feel right, the Director and Actors work out what sounds the most natural and try that. Sometimes when we're shooting, we try a few variations of takes, and see what we like best in the Edit Bay later.

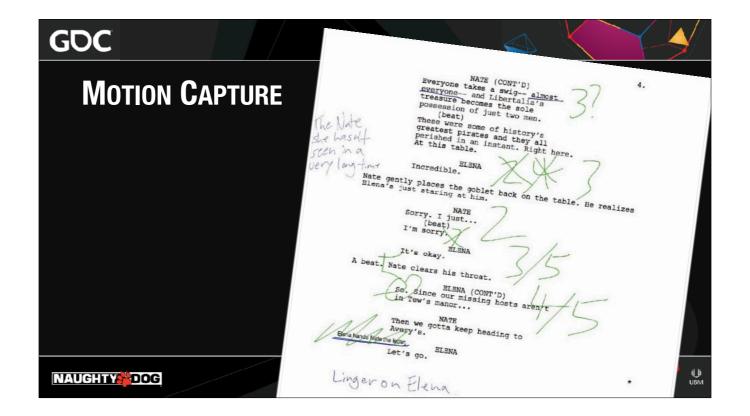


When shooting a scene, we primarily capture it in one continuous take, unless we need to break it up for additional set builds or to accommodate a stunt. Lighting, camera angles, we don't worry about any of this until Post-Production. This lets us focus on getting that "perfect master take" with the Actors, encouraging the spontaneity and improvisation that brings life and genuine emotion to the scenes.

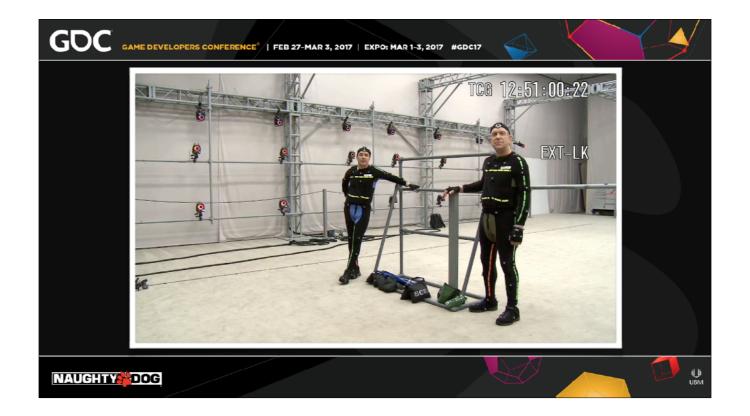


During the shoots, the Creative Director acts as the Director on set, partially because that's who's worked most closely with Design throughout the Production process, and therefore who can give the most context to the Actors.

Meanwhile, the Editor's job at a shoot is twofold. First, **staying aware of the Director's notes**, since then later they can be more helpful at personally executing—or communicating—the Directors' intention behind each moment.



And secondly, the Editor ensures we have the material we need to craft the dialogue performances later. This includes noting which takes were the best performances of a particular line...



...as well as asking for a Wild Take to get a clean read of a line when the performance is still fresh, rather than trying to recreate it months later in ADR. This is especially useful for Efforts—those "grunts" you hear when characters exert themselves—which often get spoiled by set, prop, or suit noise while the Actors are moving around.

{PLAY}

https://youtu.be/ulbSVcgeXU0



Once the footage of our Actors is "in the can," the real work begins: breaking it into pieces for the Design, Animation and Audio teams. Back at the office, the Editor, Lead Animator and Creative Director start editing as soon as possible.



We've found that cutting together reference videos is the fastest and most efficient way to communicate the eventual shape of a performance.

After choosing which takes of mocap we want to use, we create 4-paneled videos called "4ups" that show reference footage of the Actors alongside the intended camera angles.



As we experiment with cameras (and massaging the blocking for the best composition), you end up having to watch them with a split focus, looking at face here, body and camera there, and sometimes even listening to audio that doesn't match the Actor's lips.



This is because **we can take advantage of the fact that this isn't live action**. We can swap in different audio takes if we need to fix something—or use those Wild Takes to cover spoiled dialogue. The facial motion capture is even ordered separately after the edit is locked, and conformed to match. **So we truly have the freedom to change anything**.

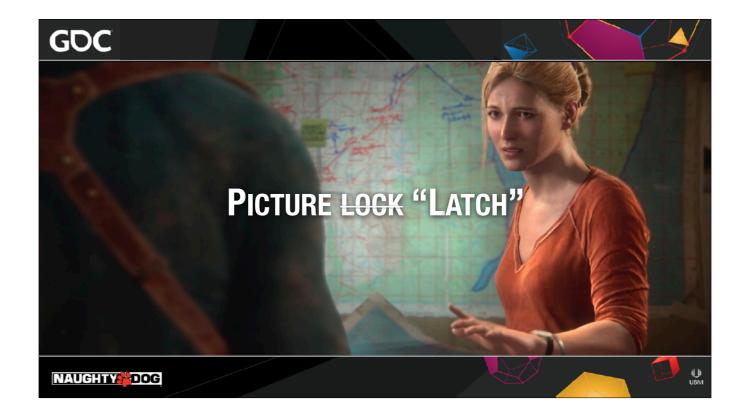


{SILENT} https://youtu.be/OOzGVQcjDPI

- Like swapping only one actor's mocap for an alternate take to get a cool gesture they did in there.
- Or cutting between two different performances in the middle of a shot. Normally we hide these edits on a camera cut just like live action—but if we don't, our Animators then have to clean up the resulting "pop" in the mocap.
- Or, at the most extreme, grabbing the face, audio and body all from different sources to cobble together a moment we don't have. Usually that's because we discover a way something could play better in the edit—and since we didn't get it originally we've got to fake it.

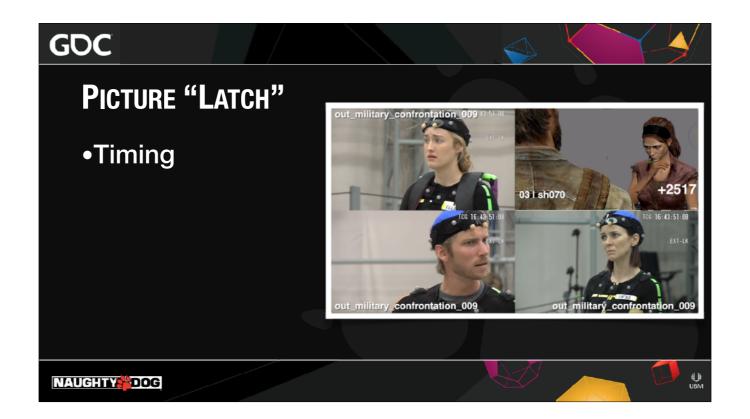


But usually the most we do is swap reads of some lines to increase clarity or subtlety. There's definitely magic in a performance that can easily be shattered if you stray too far from the source.



Once a scene's edit is approved, that 4up reference video is funneled to every department that will be involved in the further translation of these performances. This includes Animation, Props, Environment, Lighting, Sound, and Design.

But they all know the scene isn't "picture locked"—more like "picture latched"...because, as I warned, these edits may very well have to change.



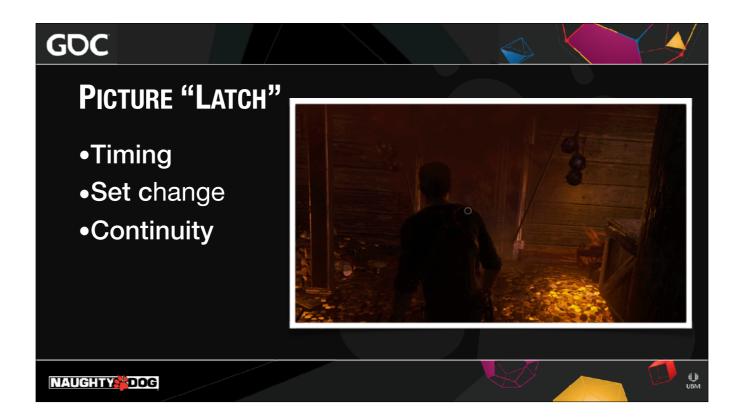
{SILENT} https://youtu.be/bQ6f18Ov9PE

• Sometimes, a change is required because of a mistake in the edit. Nailing the timing of a moment can sometimes be difficult, since you're essentially looking at mannequins in one pane, and trying to sell an emotional transition from the Actor in another. When the facial animation is applied and we look at the scene full-frame, things might move just a little too quickly, so we have to adjust—and then communicate those timing changes to all the affected departments.



{SILENT} https://youtu.be/363iQ2xUnFc

• Another change that comes up is that sets are often altered to accommodate gameplay revisions. The entrance or exit to a space had to move. There's a wall in front of that camera, or a cover object sitting right in the middle of a Character. Usually, these require minor Animation fixes, like moving characters or cameras. At worst, the environment might need to be reverted by Art and Design.



{SILENT} https://youtu.be/OH2-6OzTrPk

• The biggest changes come in maintaining continuity with gameplay. Since the transition of characters entering or leaving a scene needs to feel "seamless," their animations have to match what's happening in gameplay.

Design helps ensure this by creating "pinch points" that force a certain action from the Player—like opening a door or vaulting over something—that will blend into the start of a scene. And when it's over, we often blend characters' movements into regular gameplay animations to match on the way out.



But, when gameplay changes, so do these entrances and exits. Often this means re-capturing or at least re-animating that portion of the material.

{SILENT} https://youtu.be/Kbpulp7ZiLs

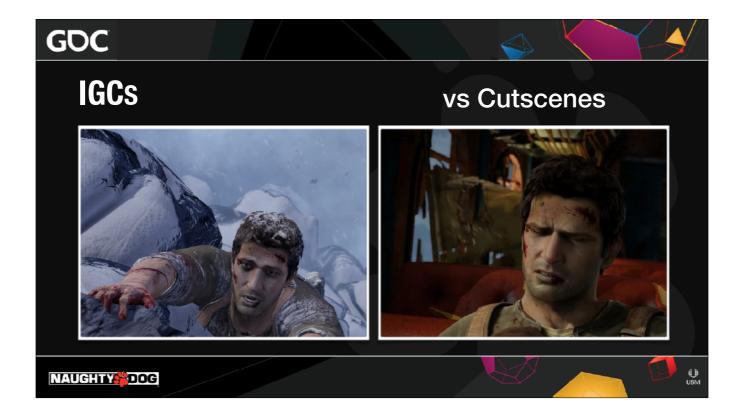
• Like this scene in The Last of Us, which first had to have the camera corrected to not be behind the Player—that would be too similar to gameplay. Then the level changed so that Joel would be in the lead rather than Ellie, so we needed to pickup a new opening where Joel went in first.



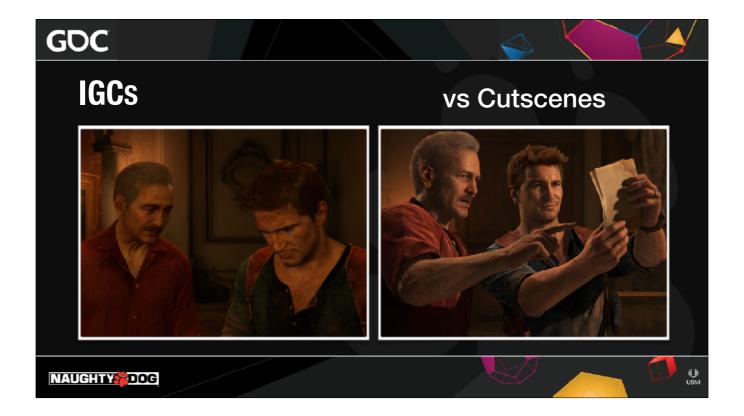
As these scenes change, we have to keep careful watch so nothing slips through the cracks. Sometimes **certain scenes linger in a "partially" done state for a long time**, waiting on a character model or a set that didn't exist when it was shot.



However, Cutscenes are just the tip of the iceberg. Since we like to keep the Player in control as often as possible, we've increasingly relied upon methods other than Cutscenes for implementing story beats. One major way is through animations we call IGCs.



"IGC" stands for "In-Game Cinematic" because—until Uncharted 4—all our Cutscenes were pre-rendered movies, so IGCs define the distinction between animations that happen live in game, versus more traditional non-interactive narrative moments like Cutscenes.



Now that everything plays back in real-time, the biggest distinction is that **some aspect of the animation is usually interactive**. In the very least, the Player might have camera control—but that all depends on the particular IGC.



Some are as insignificant as opening a door, or walking into frame at the start of a level.



Others are full-blown mini-Cutscenes, sandwiched between gameplay moments with no camera cuts, like during this puzzle in Uncharted 4.



{SILENT} https://youtu.be/kRePtBAAZhE

IGCs are also a clever method we use nowadays to **blend into and out of our Cutscenes**. The first or last shot of some Cutscenes on Uncharted 4 were IGCs, so the camera could blend smoothly, transitioning on a camera cut to the rest of the non-interactive Cutscene.



There are hundreds of these littered all throughout the game, and **they often will change multiple times**, being shortened, lengthened, or recaptured, until Design thinks they "feel" right, within the pace of the gameplay around them.



Editorial, like with Cutscenes, needs to keep up with any timing adjustments, moving parts of the audio around to match the animation as it changes.

But even though the interactivity of an IGC offers great advantages for gameplay, the way their audio works sometimes offers additional challenges.



IGC dialogue is similar to a Cutscene in that it plays from the first frame of the animation until the end, usually containing several lines of dialogue and/or efforts precisely timed to the animation.

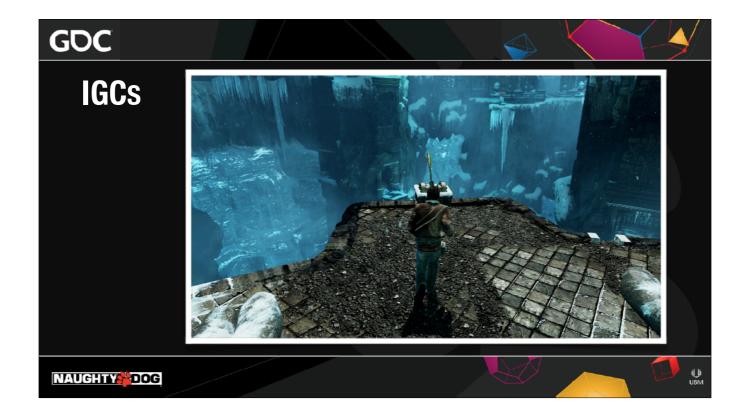


However, that's where the similarities end. Unlike Cutscenes—which play back like a movie, where every volume level or perspective shift is crafted and controlled—IGCs often have to account for a real-time, interactive camera.

So, IGC dialogue plays directly from the "voice box" of each character in 3D space. This creates different mix issues, because...



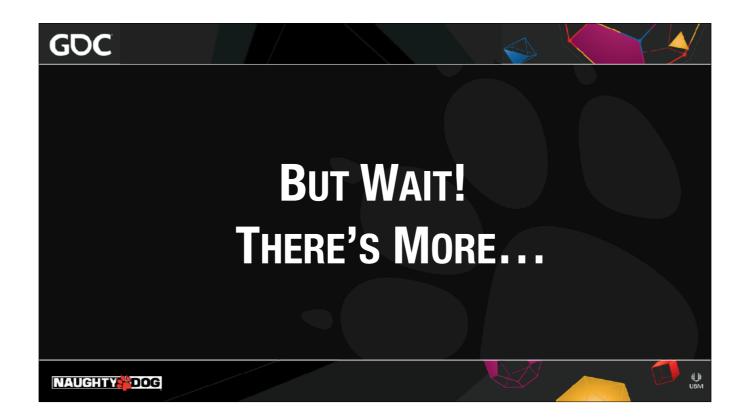
...our Camera is the Listener (i.e. microphone) by default. The benefit is that the position of the character's voice automatically matches what you're seeing on screen without any post-production work—Instant Realism!



But the downside is that if the camera is really far away and we want you to hear the speaker clearly, we have to cheat how the line plays. Or if the camera cuts while someone's speaking, suddenly the person you hear in your left ear is now talking on your right mid-word, which can be distracting.

• In this moment from Uncharted 2, we dealt with both problems. It took a lot of trickery to make Nate's voice audible from across the Ice Cave. But when the camera cuts across the space, you'll notice we couldn't quite fix the "pop" in Nate's voice as the camera cuts.

{PLAY} https://youtu.be/TTX47865_OQ



But wait, there's more!

Cutscenes and IGCs are only half the battle. Our performances would not feel seamless without a lot of other systems that play in between them.

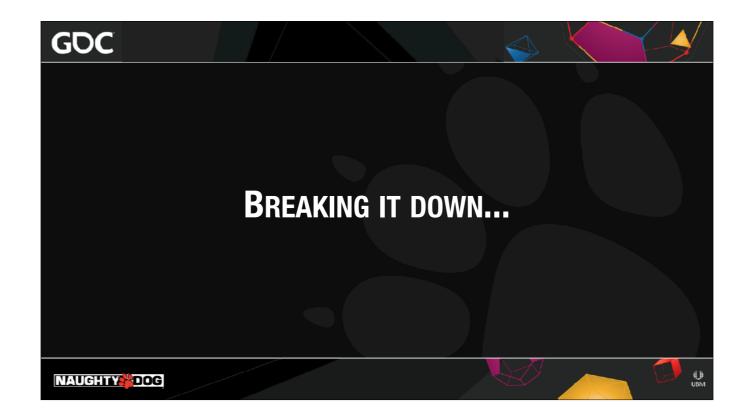
Before I break those down further, I'd like to look at an example clip of what it all looks like when everything comes together.



During this beat of Uncharted 4, Nate and Sam are just starting the next leg of their journey—exploring the ruins of a Scottish monastery that their rival, Rafe, and the team of Mercenaries he's employed are trying to excavate with dynamite.

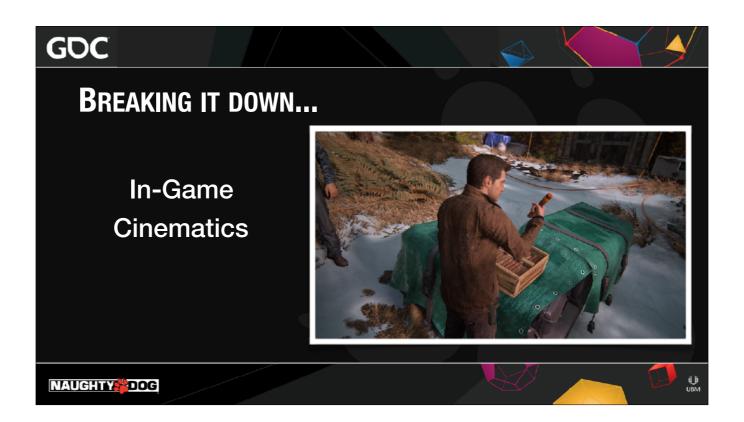
The Drake brothers meet the Shoreline mercs for the first time, learn the new dynamite gameplay mechanic, and solve the first puzzle of the level: finding a crate to climb past a section of ruins. That is, until another team of mercenaries arrive...

{PLAY} https://youtu.be/DgM8vrAtz_Q

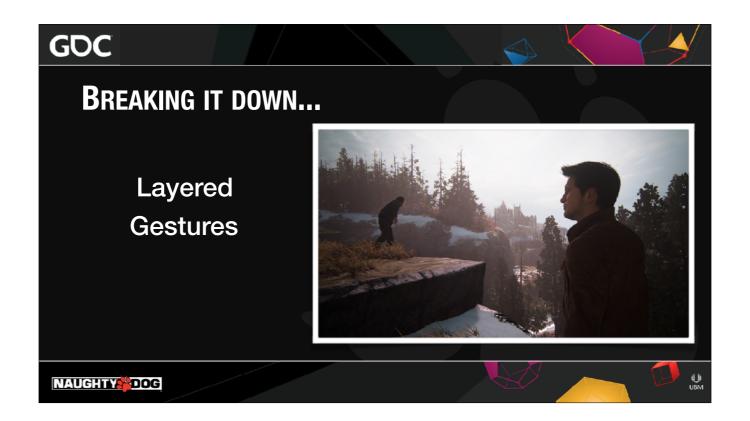


There's a lot going here, with every element helping convey our story beats.

At a glance, these elements include:



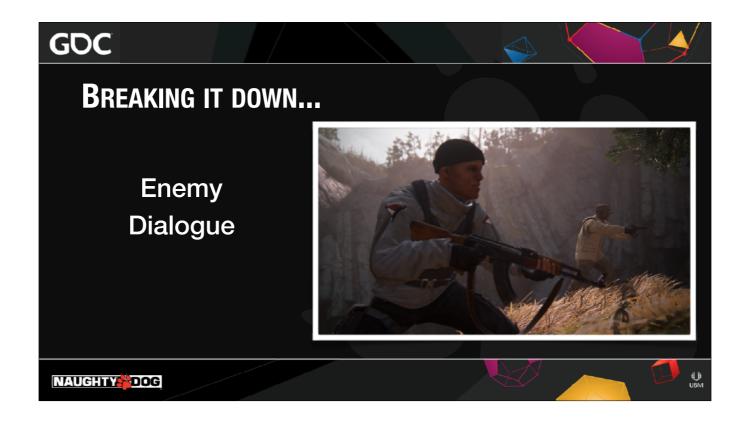
* IGCs when Nate uses his rope, and later when he picks up the dynamite.



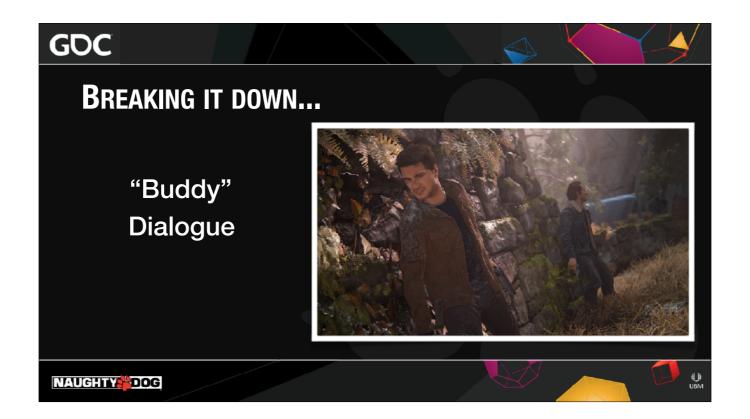
* ANIMATED GESTURE LAYERS on Nate and Sam as they explore and converse.



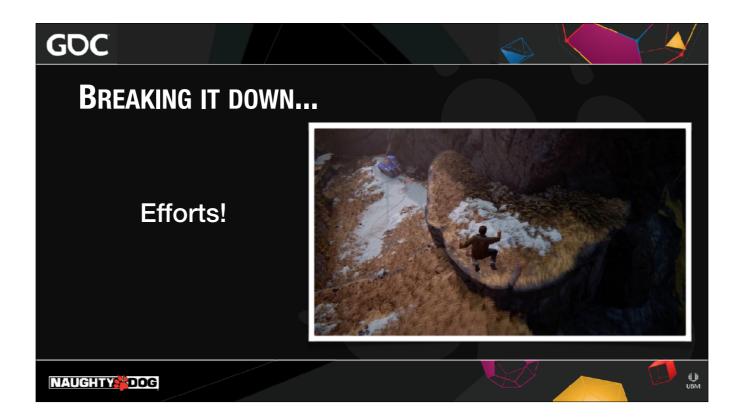
 $[\]mbox{\ensuremath{^{\star}}}$ a brief $\mbox{\ensuremath{^{\mathsf{CUTSCENE}}}}$ of Nate & Sam finding the Shoreline equipment before they're ambushed.



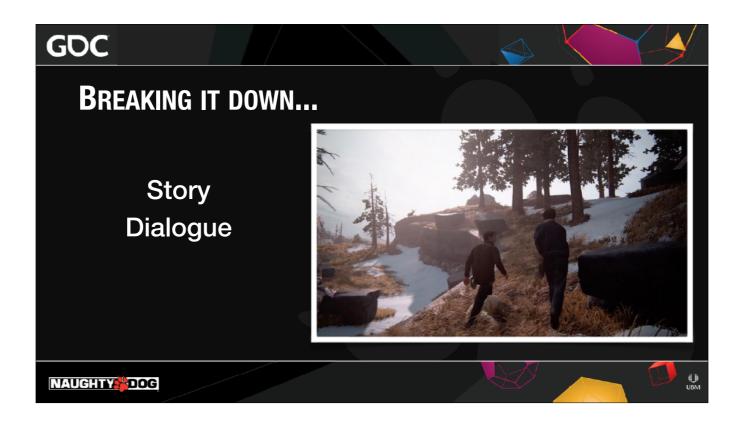
* SYSTEMIC ENEMY DIALOGUE from our Shoreline Mercenaries as they attack the Player.



* SYSTEMIC BUDDY DIALOGUE from Sam when the Player successfully stays in cover.



 $^{^{\}star}$ **EFFORTS** as Nate knocks out the last Merc, and then lands from a long drop near the ruins.



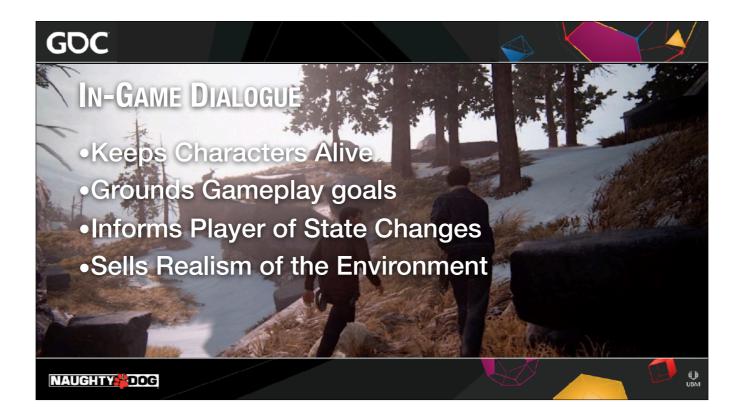
^{*} And, of course, before and after all of this there's also crap-tons of **STORY DIALOGUE** playing.



So, about that **STORY DIALOGUE**...

Most of the dialogue you hear in our games are **individual lines scripted by hand, with specific timing**. We even have a Designer and a Scripter 100% dedicated to creating systems for hooking it up. They can't complete a game entirely on their own, since by the time the game is finished, there's too much material for them cover alone. But they are *usually* the ones **triggering lines to play based on numerous conditions**—some as simple as when a Player walks through a defined region, others much more complex.

Rather than play with rigid timing like Cutscenes and IGCs, we cut our dialogue into pieces that can flexibly respond to the Player's actions. This also means we can adjust the execution more easily as things change during development.



These in-game dialogue moments actually make up a larger percentage of narrative beats than our Cutscenes. We feel this type of material is crucial because:

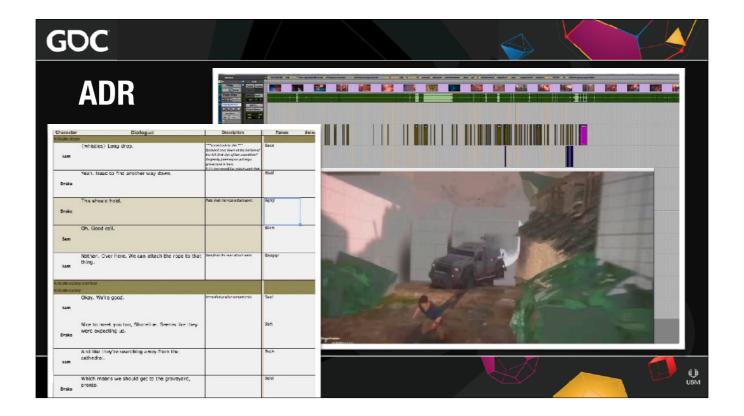
- It carries the story forward by keeping the characters alive between Cutscenes.
- It can justify practical, gameplay-driven goals within the context of the narrative.
- Likewise, it informs the Player of any changes in tension (like exploration or combat states).
- It's also a vital tool for selling the realism of the environment in the Player's mind when the characters respond contextually to it.



Unlike the cutscenes and many of the IGCs, most of this dialogue doesn't come from the mocap stage. That's because it primarily requires the characters speaking while in the Player's in full control.

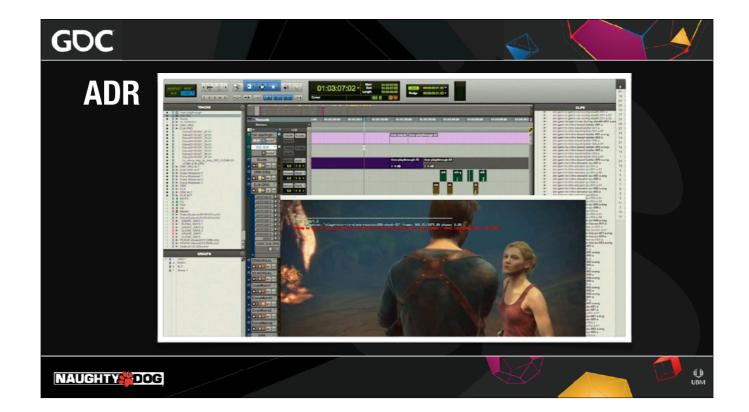


So, we record the lion's share of this content at an ADR studio. The Actors are usually paired up so we profit from their chemistry, just like at mocap. Also, they wear the same type of microphones so the audio quality matches that of the stage.

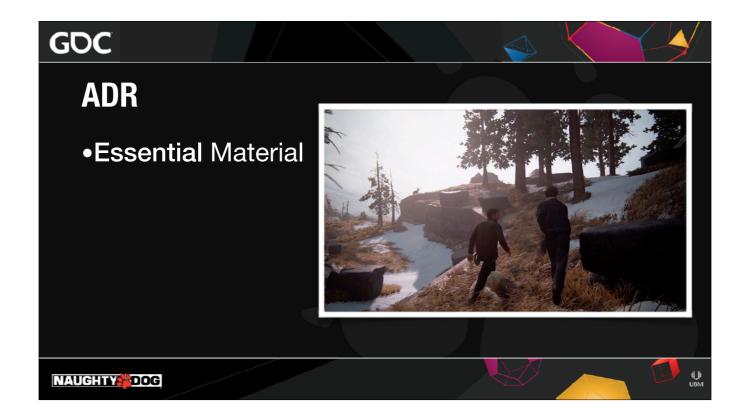


The scripts for ADR sessions come online rather late—as the levels they reflect often are still being locked down by Design.

This means that most of the time, the Actors are reading their lines without any time to prepare. So we bring a video capture of the level—in its most current state—to give them context.

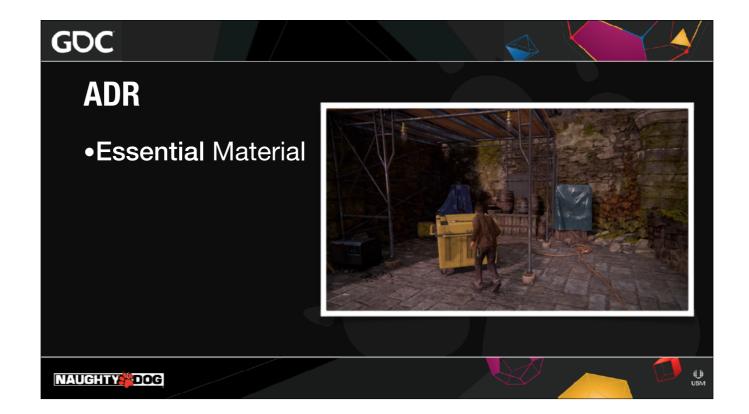


And any time there's a Cutscene that we've already shot and implemented, we watch those down with the Actors to remind them of the character's emotional state that they need to match to keep the tone consistent.



Because the game is still in flux, we record a variety of material in anticipation of ways in which the game might change.

• <u>Gameplay Conversations</u> make up the bulk of our recording. These aren't always straightforward, because **sometimes parts of them are optional**. The Actors will perform one version of the conversation all the way through, then pick up the other versions right away to keep the energy consistent.

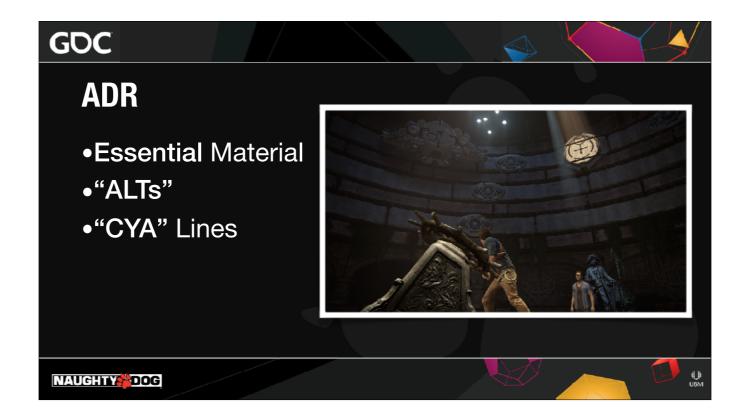


••• Such as: in that Scotland clip, different lines would play out depending on when the Player interacted with that yellow crate.

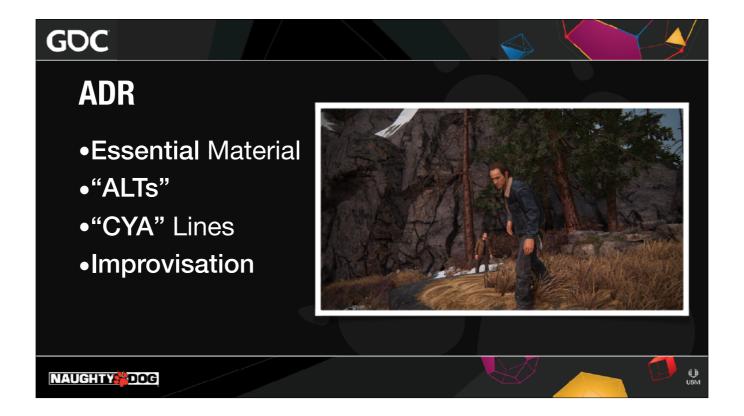
{PLAY} https://youtu.be/x3XKtjW5kbc



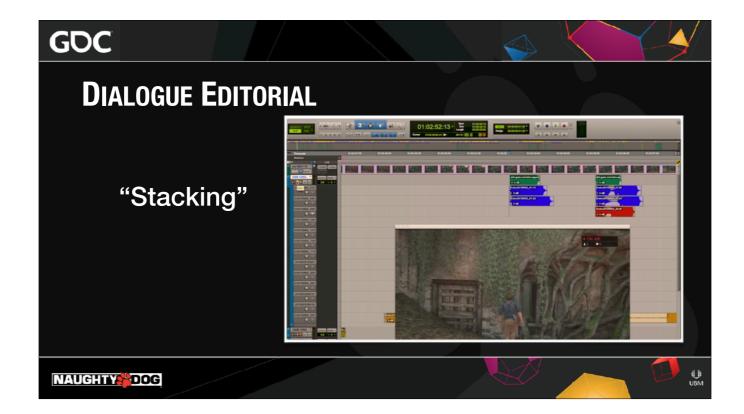
- * Sometimes we also pick up **alternate versions of lines**—"ALTs" for short—especially when the level is in a rougher state and we need to have multiple options as Design locks things down later. This frequently happens for puzzle spaces, among others—since they require tons of specific VO and animations to cover what the Player does (and in what order).
- *** On Uncharted 4, we even kept a "wish list" of ALT lines that Design thought they might need, and tried to cover them all as pickups near the end of recording. This resulted in much better coverage for the final game than we've ever had before.



* We also have a whole category we call "CYA Lines" (as in "cover your ass"), which are generic phrases like "Hey, look at this" and "Come on, this way"—which we record in multiple emotional states and contexts **Just In Case** we need them at a later date—either to help with gameplay changes or maybe sell a character beat. We record these with our heroes and all their Buddy allies.



- *Just like at mocap, our Actors are encouraged to improvise during ADR. Sometimes this will inspire us to rewrite—and re-record—a new version of a conversations on the spot.
- * And at the end of most sessions, we also record "Wild" passes of our levels, where the Actors watch the playthrough video again and are encouraged to riff on top of it. Though our Writers try to cover everything we need in a given section of the game, this often provides alternate versions of lines—or completely new conversations—that add enormous amounts of spark and character. We extract these "gems" and use them whenever possible.
- ***For instance, Sam's "location, location, location" line in that Scotland clip was an improv that Troy Baker threw in and we kept.

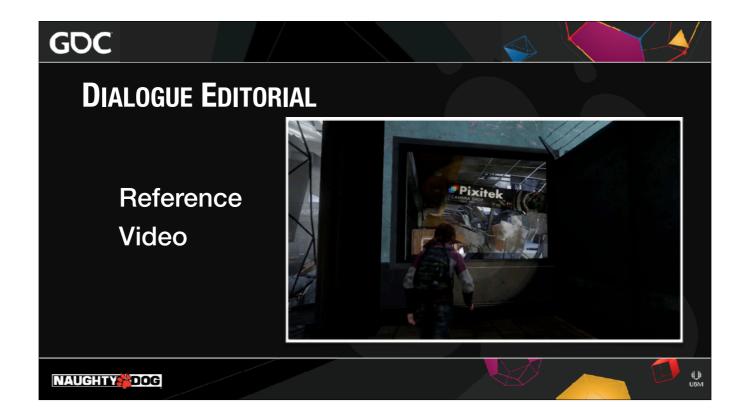


All our recorded dialogue is then brought back to the office and "stacked" in time to the video. This means each take of a line is edited in a ProTools playlist and arranged in columns, so we can quickly listen down to all the takes of a line when making our selects.

{PLAY} https://youtu.be/_2KO3EQZLFE



These picked lines are then edited into individual files with unique names, and uploaded to our dialogue database tool. **Editors also distribute scripts** from the database to our **Designers**, so they have a list of what content they need to implement.



Along with the scripts, we also supply another form of **reference movie**. Unlike our Cutscene 4ups, which are tied to motion capture data, these are instead exports of the very movies the Actors ADRed to, with dialogue placed to match the **timing** of the Actors' original deliveries.

{PLAY} https://youtu.be/PfwFQPlpJB8

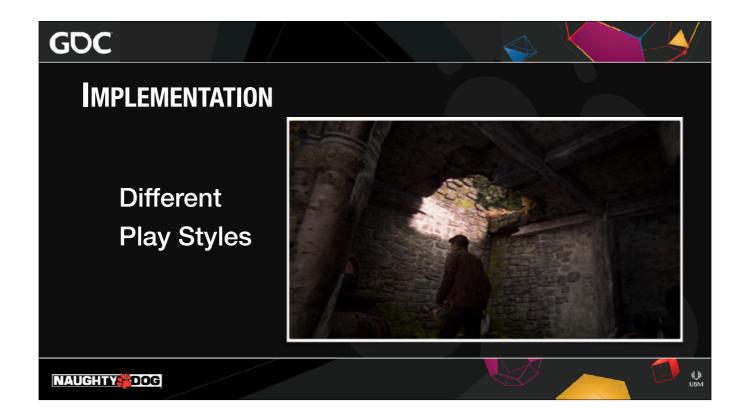
These ref vids help show where and when the individual lines of dialogue should to be scripted play, so that it "feels"—and sounds—like a realistic performance. Designers use it as a guideline as they translate all this cinematic pacing and subtlety into scripted math.

```
(wait 1.0)
(wait-say 'samuel 'vox-igc-sam-gam-sco-hillside-ruins-003)
;More Shoreline equipment.
(wait-say 'player 'vox-igc-drk-gam-sco-hillside-ruins-002)
;At least no Shoreliners.
(set-boolean 'shoreline-vox-done #t 'ss-sco-hillside-valley-sam-controller-1)
(wait (random-loat-min-max 3.0 6.0))
(send-event 'd INPERMENTATION
(wait 1.25)
(wait-say 'player 'vox-igc-drk-gam-sco-hillside-ruins-003)
;Maybe we'll get lucky and they'll blow themselves up.
(wait 0.75)
(wait-say 'samuel 'vox-igc-sam-gam-sco-hillside-climb-006)
;Ha
(set-boolean 'ruins-vox-done #t 'ss-sco-hillside-valley-sam-controller-1)
(send-event 'vox-done 'ss-sco-hillside-cathedral-dynamite-fx-1)
(go 'hillside-ruins-convo)
```

The goal is to include proper beats for a breath or natural pause between lines—either for the speaker to have a new thought, or for the companion character to listen before responding. When we don't hit this, the moment can feel rushed, disjointed, or flat-out confusing.



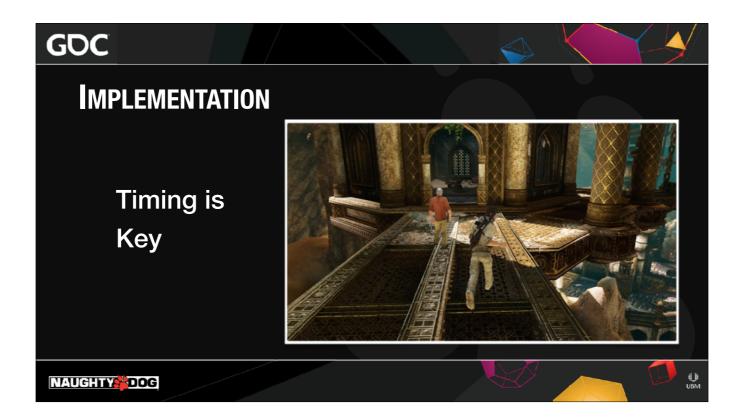
It sometimes takes several passes to get timing right in-game. **The Designers can only use these movies as a guideline,** because often the video's timing was too fast or slow, and **not representative of how everyone might play the game**. They use their best judgment, trying to keep **3 different potential play styles** in mind:



- * Some Players like to **Explore**, taking their time, stopping to hear every lovingly crafted line so they don't miss a thing.
- * But if we were to time out everything to that play style, it would ruin the experience of more **Average** Players, who like to enjoy the story but keep things moving.
- *** Where we can, the simplest solution is to just have certain lines **not play** if the Player is going to fast. Like in the Scotland ruins, the line specifically pointing out that a ledge is "too high for a boost" doesn't play if the Player rushes out of there before the previous conversation is finished.

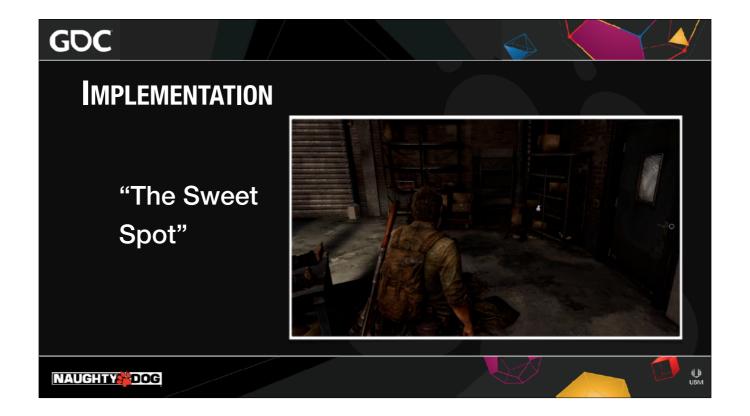
{PLAY} https://youtu.be/EF4diXE1TUk

* And then there's the **Speedy** Players, who dash through the puzzles to get to the next combat setup. We do our best to get our conversations playing naturally—but quickly—so they don't get lost, but sometimes it can't be helped.



*** Once in awhile, we have an area too small for the amount of VO that has to play, and end up shipping with moments that can feel like a string of lines back-to-back, depending on how quickly the Player is going. Like this moment in Uncharted 3 that just ended up a little too tight, compared to the pacing of the cutscene that preceded it.

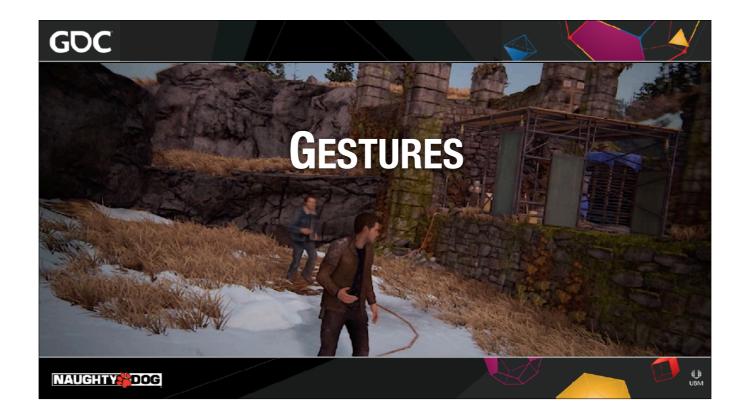
 $\label{eq:play} $$ \underline{\text{PLAY}}$ $$ \underline{\text{https://youtu.be/0wkLdee5cMY}} $$$



If we get the timing right, it can convey subtlety without needing a specific body gesture or facial performance to sell a moment.

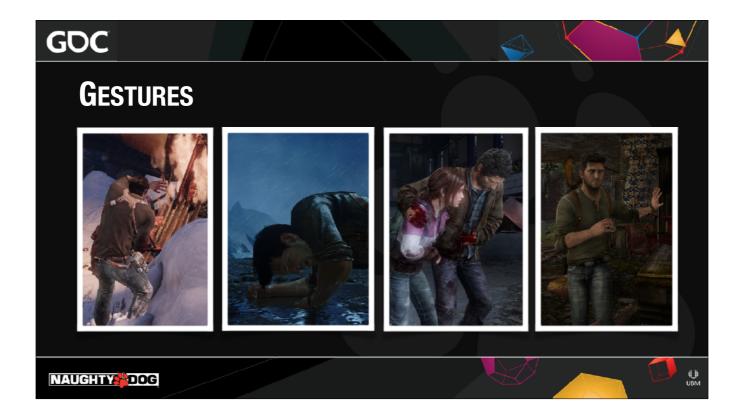
*** Like in this bit from The Last of Us, the pause before Ellie's "oh" line conveys a turn: her realization about Joel's darker past. Her follow-up question then gives Joel an awkward pause of his own. All of this emotional content manages to fit in one small stairwell, even if the Player never stops.

{PLAY} https://youtu.be/xG2aLMeW-cw

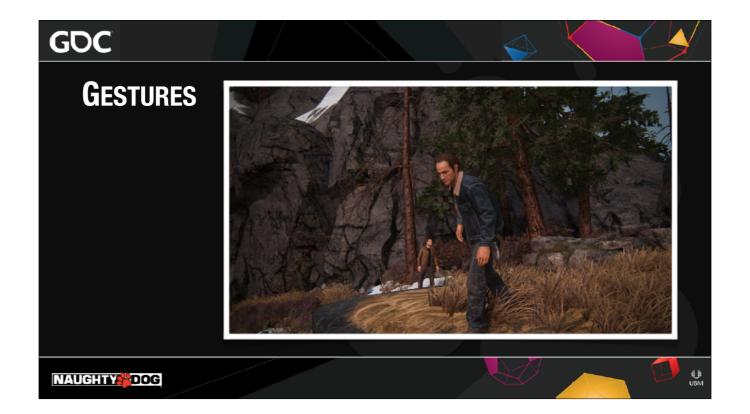


That moment worked on dialogue Timing alone. But when we need more, we can dynamically **add extra body motions** to our Story dialogue to increase physical—and emotional—realism.

*** In that Scotland clip, when you saw Nate waving his arm and actually looking at Sam while talking to him, that was an example of our **Gesture System**.



In previous games, we've had all sorts of **Contextual Move-sets & Animation** that helped sell the reality of the Hero's world by having him/her react to their environment. Whether it be recoiling from fire, stumbling while exhausted, bracing themselves against walls, etc.



Our designers are also able to play **additional unique animations** on our AI characters to bring them to life contextually. That's what allows Sam to look around and dynamically check out the world around him.



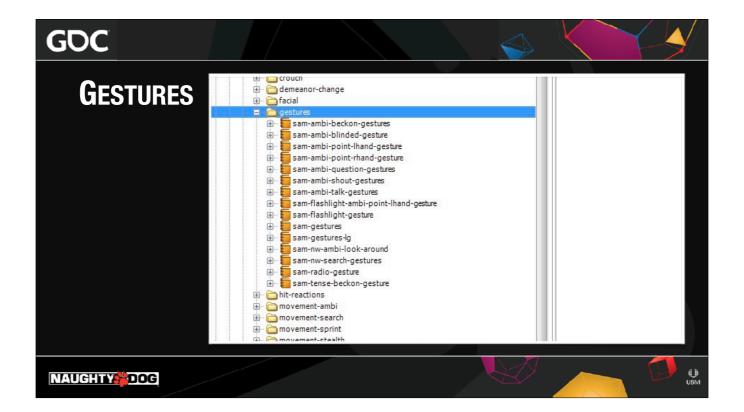
{SILENT} https://youtu.be/tOEE_Oft0vg

But now we employ an additional level of detail, **Gestures**. These are partial animations that play on top of normal gameplay animations, like a head shake, an arm pointing, a shrug, etc. The Player can be running, climbing, or crouching and these layers will usually still play.

***In case you're wondering why sometimes an arm isn't moving in that clip, that's because it's not part of the Animation's layer. Meaning the Player can be carrying a gun or other object, while the rest of the body does the Gesture.



One of the ways we "sell" these moments is by hooking a specific Gesture up to a line of dialogue in script. So when each line plays, the Gestures play on top of whatever the character is doing. Programming and scripting magic also helps the speaking characters actually look at each other dynamically, when appropriate.



Since we don't record specific Gestures for every line, the final performance isn't as rich as a Cutscene might be. But the result is still an added level of life that doesn't restrict the Player's movement as they roam around the game world.



Though we can also hook up custom **facial** gestures to lines in the same way, we usually don't go to the extra work—since the camera is normally behind the Player character's head, and far enough away from other speaking characters. By default we employ a neat **procedural system**, which generates lip sync and facial animations using the audio file's amplitude and subtitle text.

{PLAY} https://youtu.be/qHm-WOf2Zlo



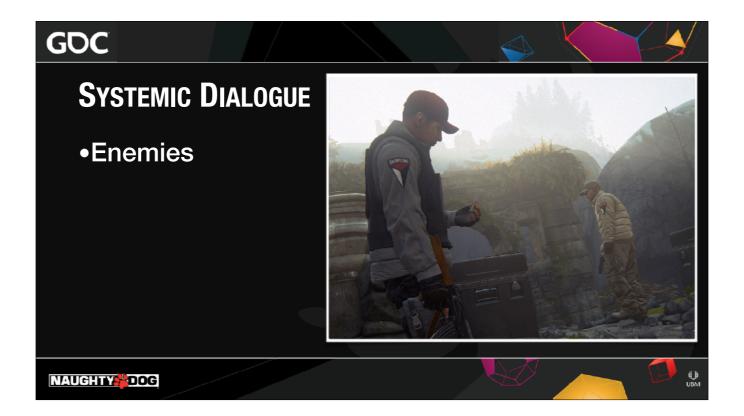
Besides the story dialogue, we have another category of lines called **Systemic VO**. These are **groups of lines with similar context that can play when triggered by Al-driven events**. This material helps to add life, reinforce character relationships, and provide feedback in various situations throughout the game—mostly during types of gameplay scenarios that use repeated systems, like Combat.



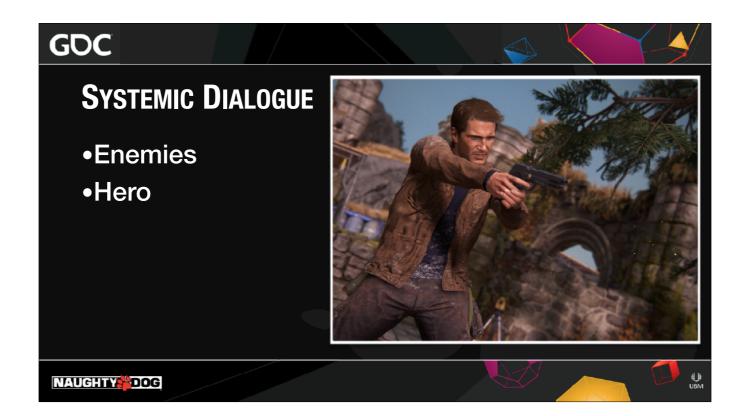
When an Enemy spots the Player, for instance, they have specific lines they use to call that out. ("There he is!"/"I see him!") These "Buckets", as we call them, offer multiple varieties that play based on the same situation, so the Player doesn't just hear "I see them!" over and over and over.

***At the end of our Scotland demo, Sam called out one of **his** Systemic lines to indicate the Player is still hidden from the Enemy. The entire bucket consists of these lines:

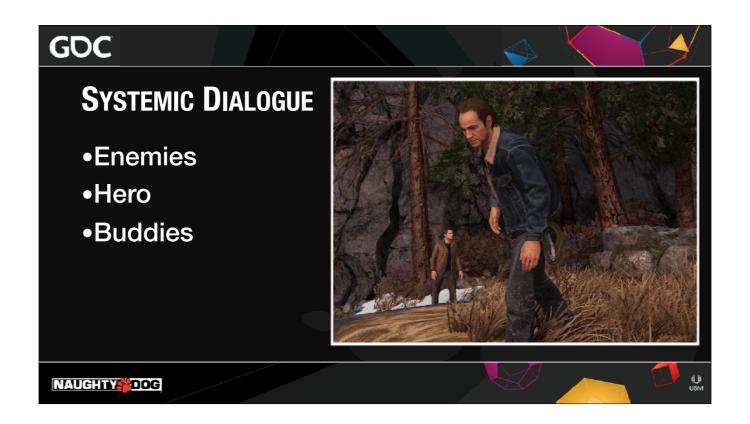
{PLAY} https://youtu.be/CGKsgNMdqWQ



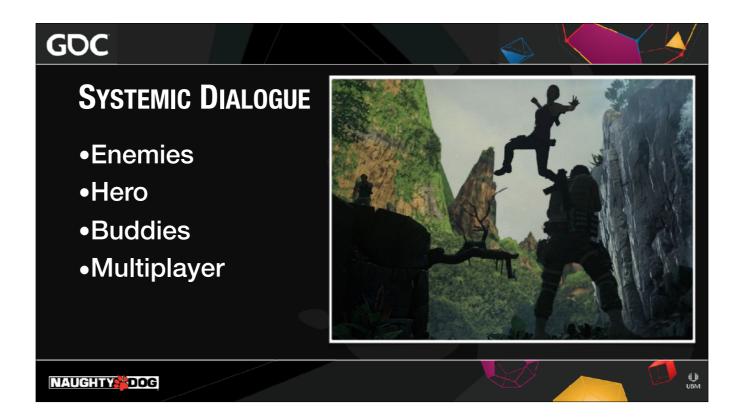
This sort of dialogue serves a practical function in terms of Player feedback, but it also keeps the characters alive and reacting to what's going on in the same way the Story lines do. Which is why we don't only use Systemic dialogue for our Enemies....



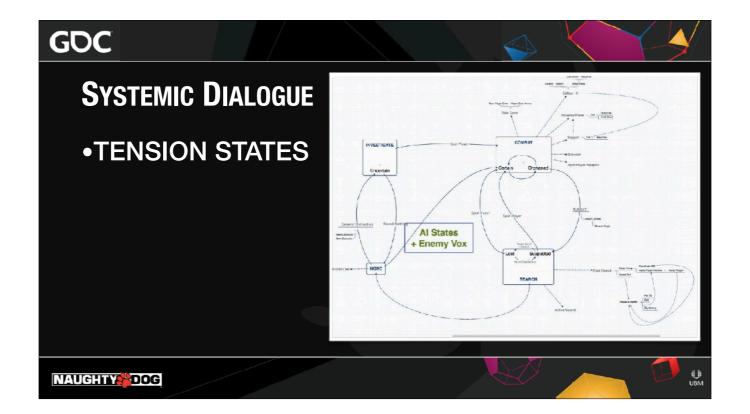
Any dialogue triggered by AI, rather than just script, falls into this category—including Buckets for any major characters: the Hero,...



...their Al-controlled "Buddies",...



...and of course all of our Multiplayer characters (since that mode is pretty much all combat mechanics).

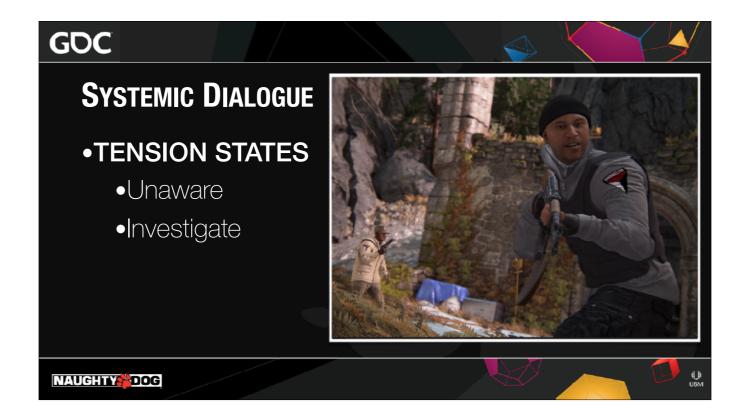


The system is controlled by the global **Tension State**, which we break down into 4 stages: **Unaware**, **Investigate**, **Search**, and **Combat**.

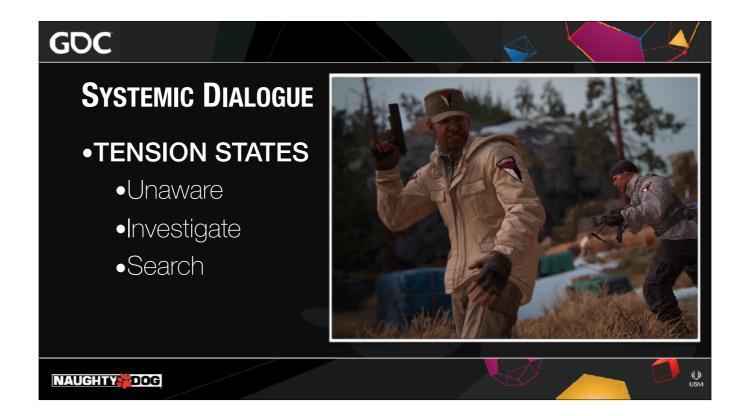


In Unaware, any threats nearby have no knowledge of the Player or their companions.

- Enemies can converse with custom, story-based dialogue, like they did in the end of that Scotland clip.
- The Player and his Buddies, meanwhile, will likely be in standard "Explore" mode. This is the natural state of the game. If they get close to Enemies in this state, Buddies and the Hero will use the Whispered versions of their various buckets, so the characters sound like they're trying to stay hidden.



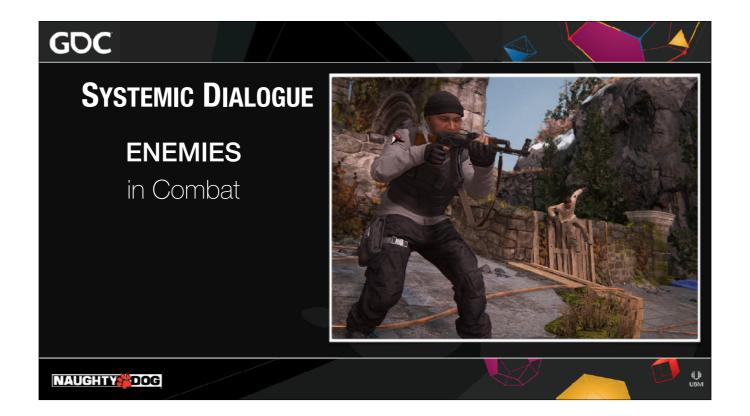
In **Investigate**, tension rises a bit. The Enemies heard or saw some evidence of Player interference. They try to uncover the source of the distraction—now playing their Systemic buckets regarding what they saw, checking in if anyone saw anything, etc. **This is the first level of Tension where the Player would feel threatened**, and that they should either fight—or get the hell out of there.



In **Search**, tension is near its breaking point because Enemies saw the Player, then lost track of them. They speak buckets of lines so the Player can audibly locate the Enemies attempting to converge on their position. Feedback from the Heroes remains whispered—including Buddies pointing out where the nearest Enemy is ("On your left." "Behind you.") or warning them if someone approaching is carrying a powerful weapon ("Lookout. Shotgun.").



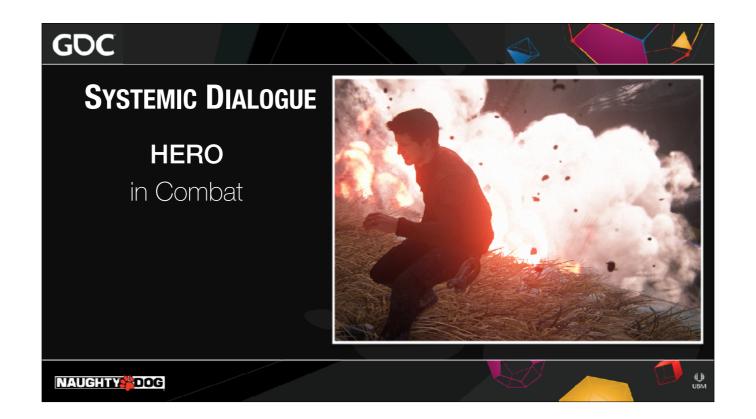
Combat is when all hell breaks loose, and in our games uses almost exclusively Systemic dialogue. But since we want it to feel as nuanced as any quieter story moments, we employ a lot of buckets to cover a variety of potential situations.



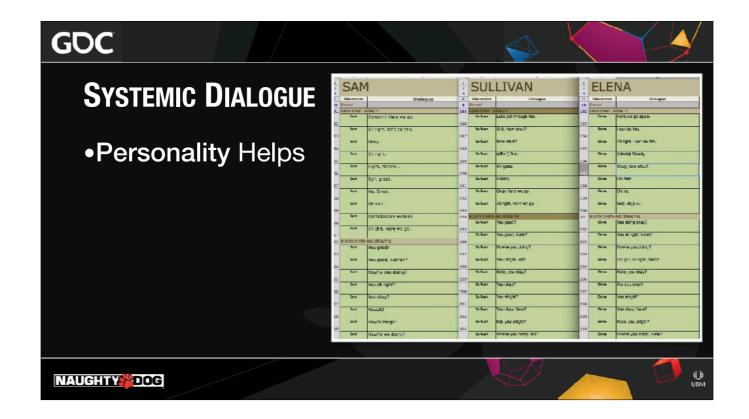
* Enemies are aware of the Player, and actively trying to attack him/her with guns, fists, or whatever other means they have. They have lines announcing when they are flanking, pointing out the Player's exact location, or warning each other if the Player has a power weapon like an RPG or Molotov Cocktail.



* Buddies continue to warn the Player the same as they do in Stealth, but instead of whispering, their lines are shouted over gunfire. We also have dialogue that helps emphasize their relationship to the Hero—i.e. that they care about them. Buddies ask if the Hero is okay after they've been shot stumbled by an explosion, shout out that they're coming to the rescue if the Player is grabbed by an Enemy, and say "thank you" to the Player if ever they're rescued in return.



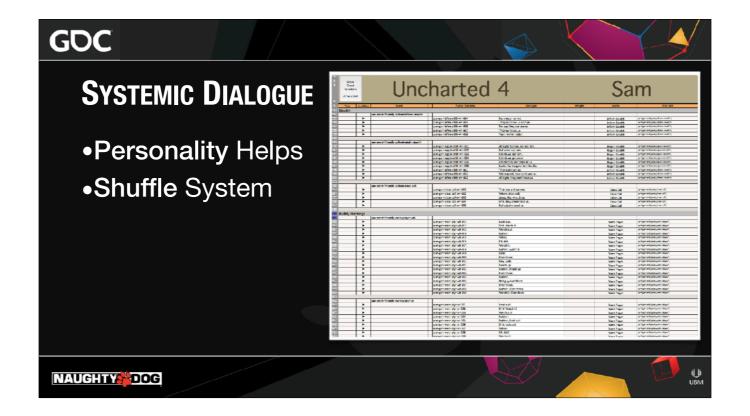
* The Hero also has dialogue to connect the Player to their emotional state, like the stress of diving out of the way of a grenade ("Oh god!") or steeling themselves up while reloading in cover ("Okay, here we go"). There's usually something that plays after taking down an enemy as well. In Uncharted, it's usually a funny quip that suits the tone ("Lights out!"), whereas in The Last of Us it matched the world's grittier nature (like an "Okay.").



Since Systemic lines have a more functional purpose, they sometimes lack the subtlety or subtext that our longer conversations do. But they always maintain the same personality as the character delivering them, so that the Player doesn't become weary of hearing the exact same thing from multiple characters over and over.

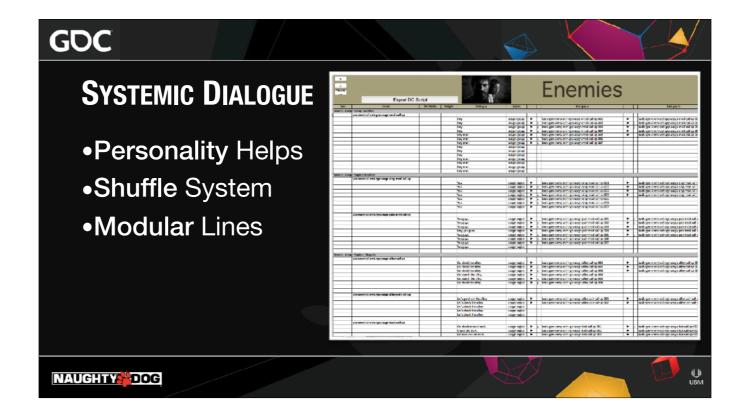
{SLIDE}

*** For example, Sam calls our main character "Nathan" or "little brother" in a lot of his systemic warnings—and is more ball-busting in his tone—while Sully calls him "kid" and maintains a more paternal tone. Elena calls him "Nate," and has her own brand of...loving exasperation with our hero in her pattern of speech.



Even with personality thrown in, the lines could still sound somewhat repetitive. So we have a "Shuffle" system which keeps track of every line said by every AI group, advancing the queue for any one bucket across the board when they speak, ensuring that the same exact line will not be heard frequently.

*** If Enemies spot the Player, one might yell "I see him!" The next time that happens, rather than hear "I see him!" again, the Player will hear "It's him!" or some other variant. Not until the bucket has been exhausted by the shuffle system does it reset to the top again. We try to record 10-20 different lines per category that might be heard the most to avoid the Player hearing a repetition in any one play session.



Sometimes, the lines are written to be modular, so that they can be stitched together based on situation. Not only does this allow for variety in the simulation, but it **makes the Enemies seem alive and intelligent.**

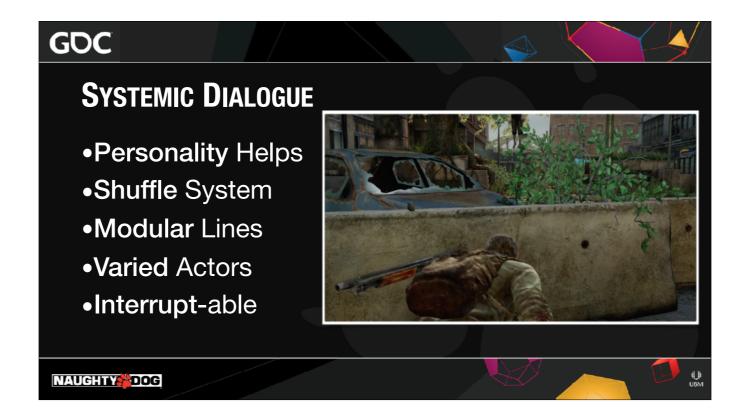


*** In The Last of Us, an Enemy might say "Hey. You two. Go search over there." Each of those is a separate bucket of lines that plays back to back, utilizing natural pauses to make them sound like complete, varied thoughts. In another scenario, it might play as "Hey. You. Go check the alley." or some other variant.



One subtle way we keep things feeling more real is by avoiding repetition not just in lines, but in the voices themselves. Since we have a limited number of Actors voicing our NPCs, we try to make sure they have different voice types (medium, low, etc) so that the Player won't hear voices that sound the same over and over.

Also, as I mentioned, Enemies are often initially encountered while they're having a Story conversation customized to that particular group or area. **We keep track of which Actor's voice is being used in each conversation**, then try our best to not use that Actor in the next setup—further avoiding repetition.

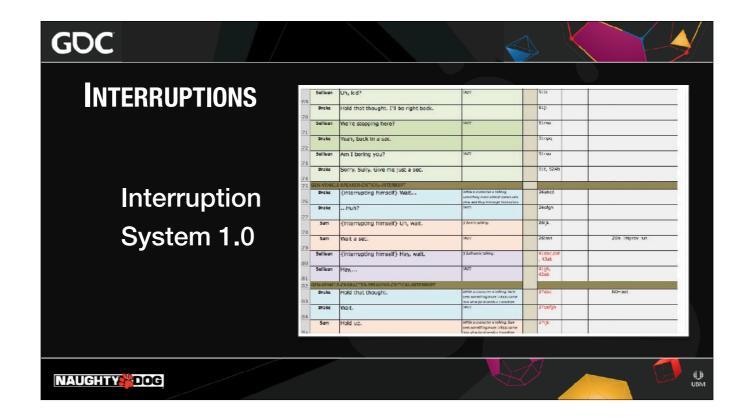


A tricky situation that then comes up with these Enemy conversations is **how to interrupt these lines that are already playing**. If the Player shows up in their vision cones, they have to switch to their Systemic dialogue and go into Combat. So we cut off a line, wait a split second, then just play the Systemic dialogue as if the Enemy had interrupted themselves voluntarily.

{PLAY} https://youtu.be/Zn1jhCQ4CXI

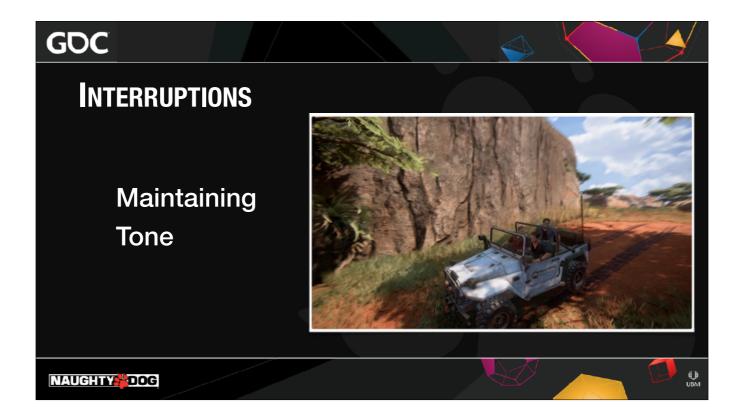


For a long time, this rough method was the only way we had to interrupt conversations. However, the addition of a drive-able vehicle in Uncharted 4—which the Player could exit at any time—meant entire sequences of lines might be skipped just because the Player wanted to explore a little.



This forced us to come up with a more elegant solution for interrupting—and resuming—dialogue between Nate and his Buddies, ensuring conversations couldn't so easily be lost forever.

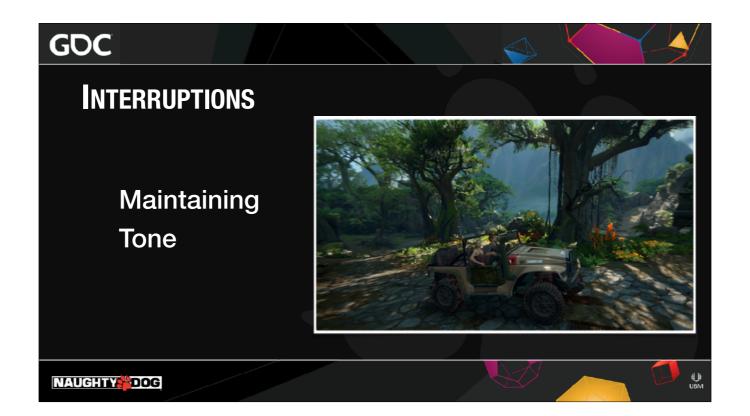
• We recorded more "buckets" of lines regarding—and responding to—the Player stopping the 4x4 and exiting, then attempting to resume the conversation upon return.



• To keep things in line with our story dialogue, the tone of these interrupts had to be different, depending on the nature of the characters' current relationships.

*** In the plains of Madagascar, Nate is comfortably exploring with Sully and Sam, so his tone is more casual.

 $\left\{ \begin{array}{c} PLAY \end{array} \right\} \ {\scriptstyle https://youtu.be/B_rL8xH0uD4}$



*** On the Libertalia island, Nate and Elena's relationship is strained, so he's more subdued when he hops out of the car.

{PLAY} https://youtu.be/enfZxvYGqvU



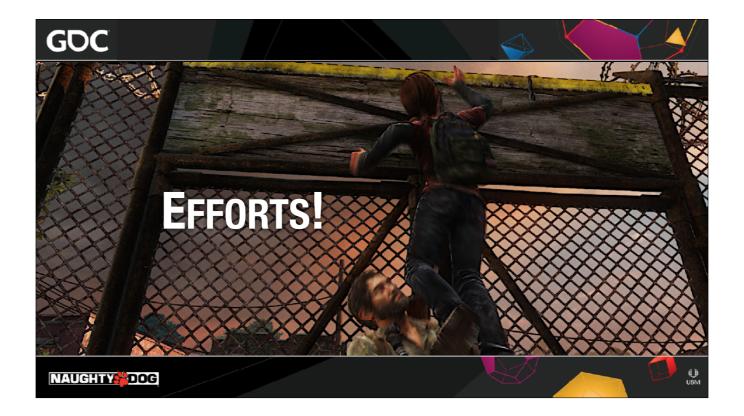
• To make it sound natural when the character stopped speaking, we had the Actors record "glottal stops"—the sound a person's throat makes when it stops speaking. These would play to cut off their lines in as natural a way as possible.

{PLAY} https://youtu.be/Y_Y-GS0OwfA



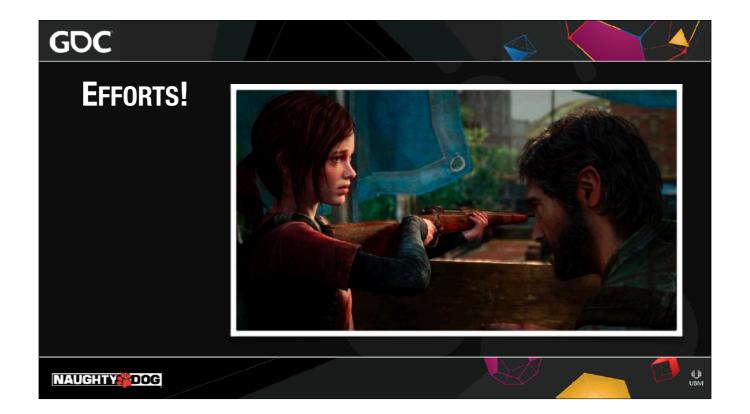
• Sometimes we edited alternate versions of lines that would flow more properly when coming out of an interruption. This meant combing through an entire conversation, and marking what lines needed ALTs in case the conversation resumed on that line.

In retrospect, we wish we had purposefully recorded alternate versions to restart a conversation, rather than just edit the originals. This system is still very new—and we're looking forward to building on it more for future releases.

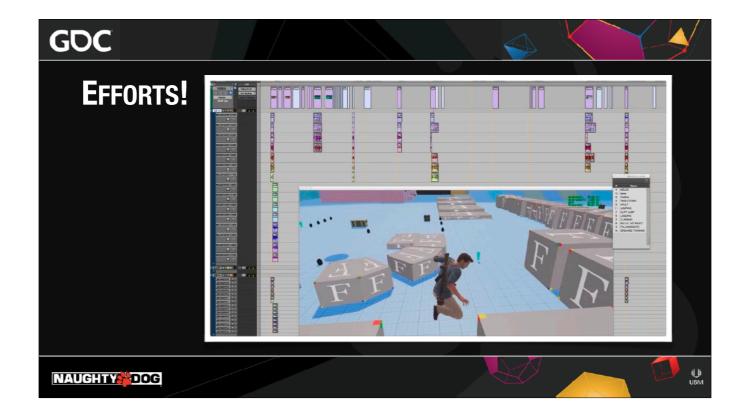


There's one last category of dialogue to cover briefly. **Efforts**! Besides during those Wild Takes we get at mocap, we record lots and lots and LOTS of Efforts—of various types—with each of our Actors.

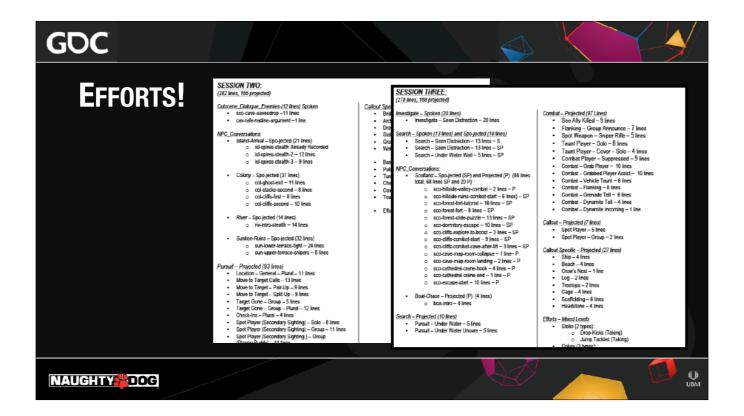
• Mostly because they give feedback to the Player while also helping the characters come alive, selling all those physical situations like climbing, fighting and...well, dying that happen in our games.



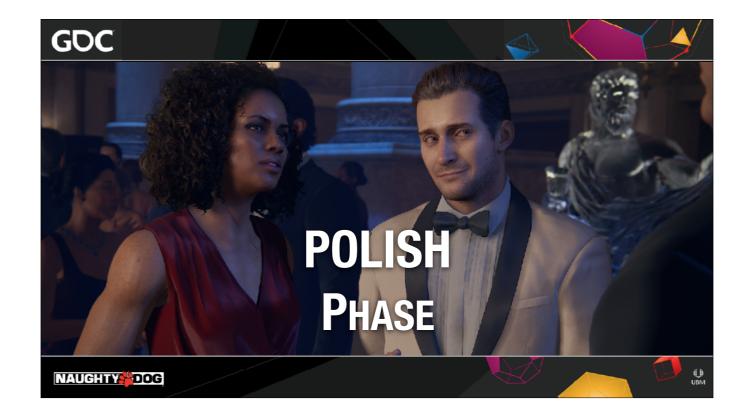
• Just having breathing from the characters, especially in situations that are tense, or where they're fatigued, also helps give a sense of what a character is feeling without them having to state it.



When recording, we have Actors chase videos of the Player grappling, fighting and falling—so they have context for the sounds they're making. We often get different sizes—small, medium and/or large—of each sound, and play the appropriate ones based on the intensity of the animation.



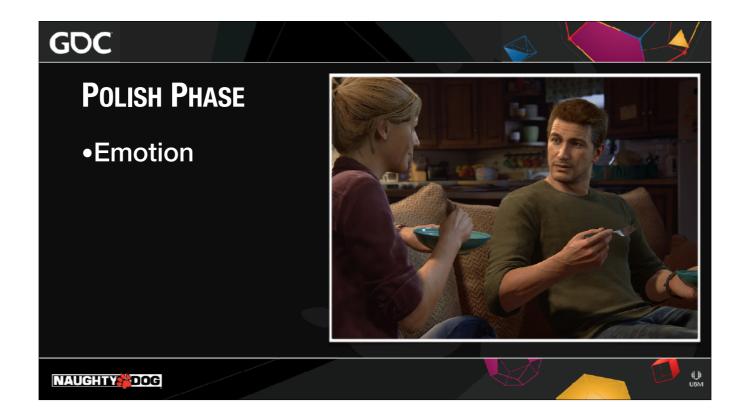
And, as most efforts are vocally stressful, we always record them at the end of a session. In fact, we do the same with any shouted material, so we don't blow out an actor's voice in the first hour.



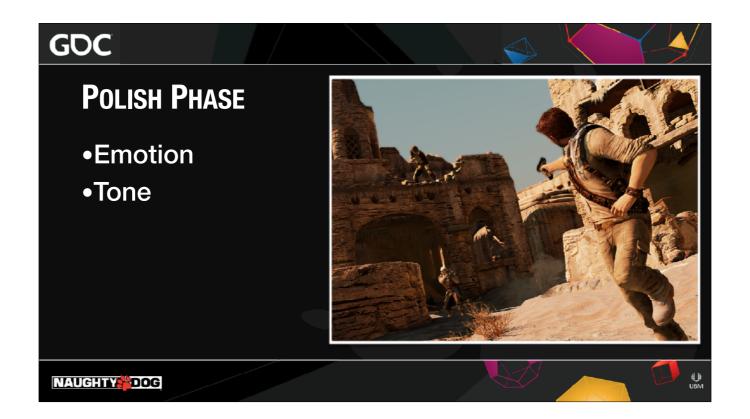
So! As all of these systems and animations start coming together, that's when the glut of minor [and major] changes truly start. A Cutscene or a conversation might work in a vacuum, but it's not until we get these pieces into the game can we truly judge what works.



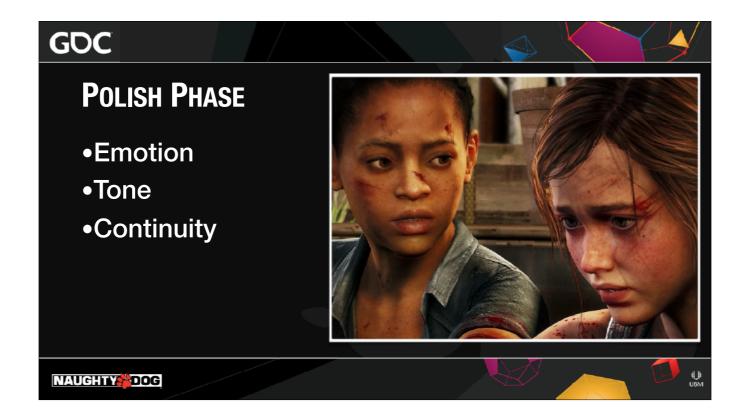
It's about getting it "on the stick" and feeling it out, and then making sure the performances remain consistent despite the adjustments. We continually tweak, swap, and rearrange our dialogue and animation—sometimes over the course of years—all to improve the emotional connection between the Player and what's going on around them.



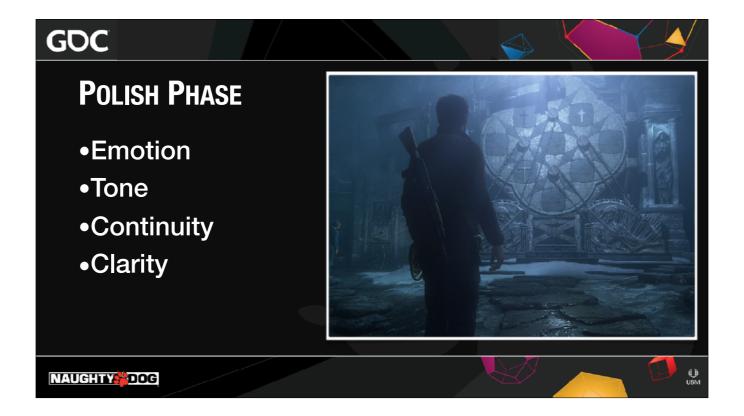
• First and foremost, **not all the content we write and record "works"** in the final context of the game. Even if nothing has changed in Design or Animation, **sometimes the emotional beat feels false**. IF possible, we'll alter that VO with an edit. By adding pauses, or stealing a breath, sigh or laugh from somewhere else and cobbling it onto the line, it can really help **make moments more natural**.



• Our Systemic dialogue might have similar problems when working in context. Sometimes an **AI bucket will fire in a situation that's tonally inappropriate**. Like Nathan Drake's bucket of quips that he might spout off anytime he knocks a dude out. If the tone of the game is more serious in that moment, our Designers flag the area to suppress all that VO until it's appropriate for it to play later in the game. The same goes for the Buddies's systemic dialogue.



• Another possibility is that a small—or not-so-small—bit of exposition might need to change. Sometimes it's contained in a cutscene, others it's a gameplay line whose context was altered in Design. Usually something this specific can't be fixed with an edit—a new line needs to be recorded.

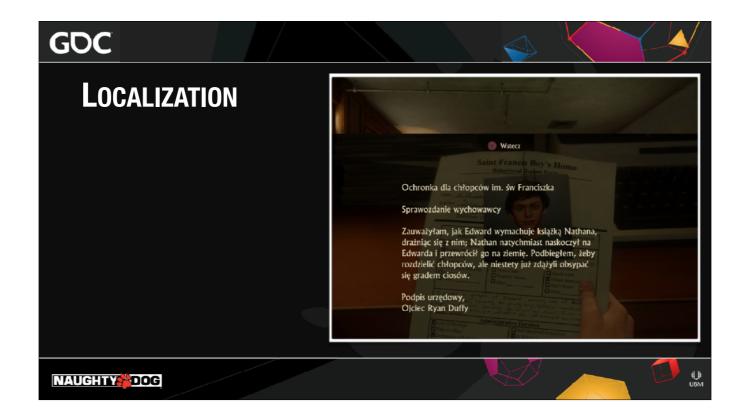


• There are also times when a line is making things confusing, so it **just has to go**. We hold many Focus Tests to determine where the pacing drags, augmenting or cutting as needed to hit the ideal amount of gameplay between story beats.

Since we over-write our levels anyway, we can usually afford to cut part—or all—of an exchange, and move another bit of dialogue in to fill the hole if necessary.



All of these adjustments make the game better. But...we also have limits on what we can change. As a project draws to a close, the first hard wall we hit is Localization.

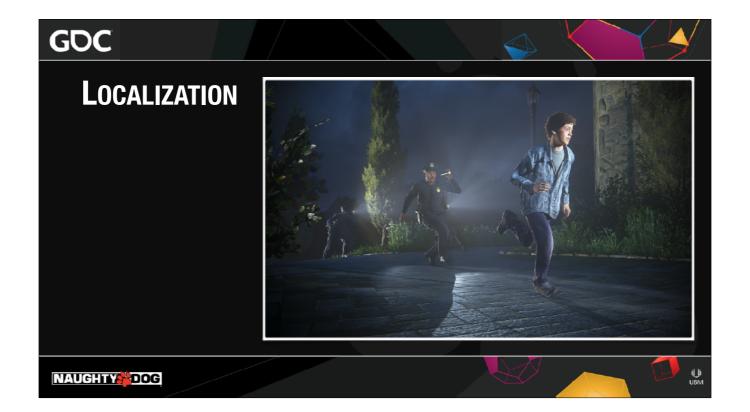


Localization, for those that might not know, is the process of translating and re-recording all text and dialogue into dozens of languages for international Players.

Games at some point have to lock, recording no more new lines and writing no more new text so that the existing material can go through this process in time for shipping.

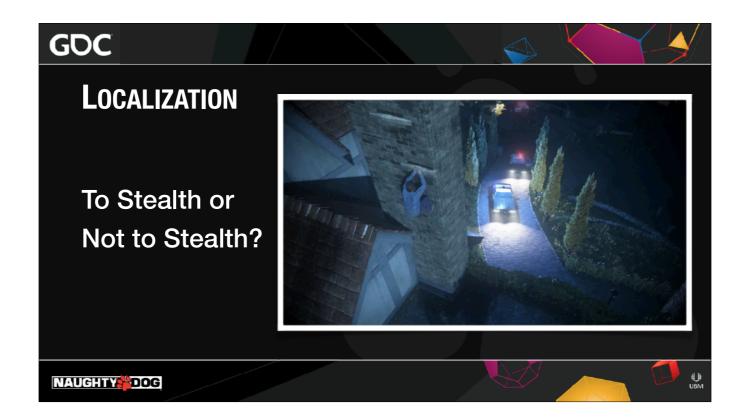


The problem is, nowadays the deadline for Localization is not after the game is finished. It's at least a few months prior. Which—especially since recorded dialogue is often so hyper-specific to the actions happening onscreen—leads to a greater problem when things in a game evolve after that date. **Because once we've reached the Loc deadline, nothing new can be recorded.**



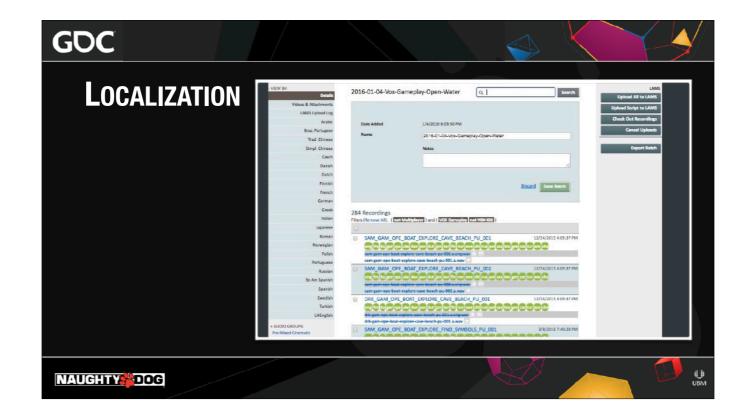
- This phase is when those generic "CYA" lines become **really** handy. We can use them to "patch" areas of the game that evolved after the deadline.
- ••• Need to direct the Player's attention at the solution to a puzzle? "Hey, look at this."
- ••• Does a character shout out "This way! C'mon!", but you want to sell that their Buddy is following? Add in a "Right behind ya!" response.

This type of dialogue has become so essential, we intend to record even more lines like it on future projects.



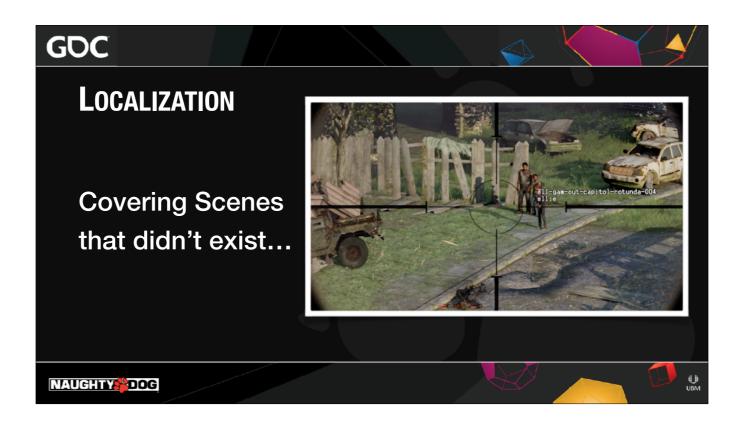
{SILENT} https://youtu.be/0QOy2QL0Rss

*** Like on Uncharted 4, we didn't know if the kids' escape in the "Manor" flashback would feel better as a tense stealth sequence, or a balls-to-the-wall chase from the police. We had recorded dialogue to support both scenarios (as time was running out), but most of it was for the stealth version. Turns out, what worked best was the latter. We ended up covering the majority of the chase using CYA Lines.



• But sometimes, CYA lines aren't enough. We need something specific, and have to borrow lines from another part of the game.

So we go digging through our database to see what we have that can cover the request. When picking the line that will be duplicated, we take into consideration how far apart the new instance is from the original place the line plays, so that—again—the Player won't notice a repeat.



*** Like this moment in The Last of Us that was covered entirely through CYA lines, re-used Systemic dialogue, and a Story line of Ellie's from earlier in the game.

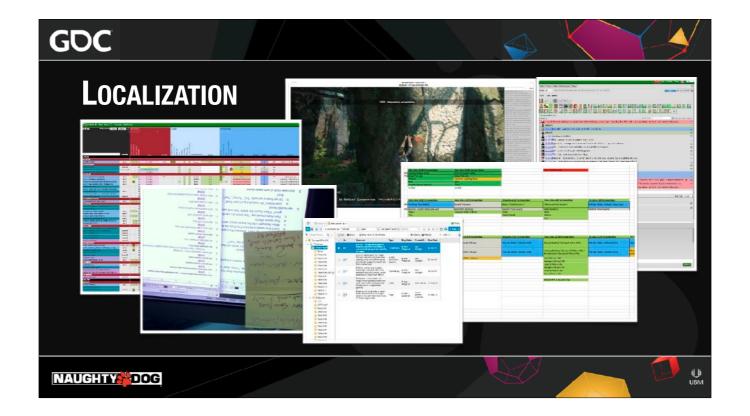
{PLAY} https://youtu.be/QdBWbQiRsU8



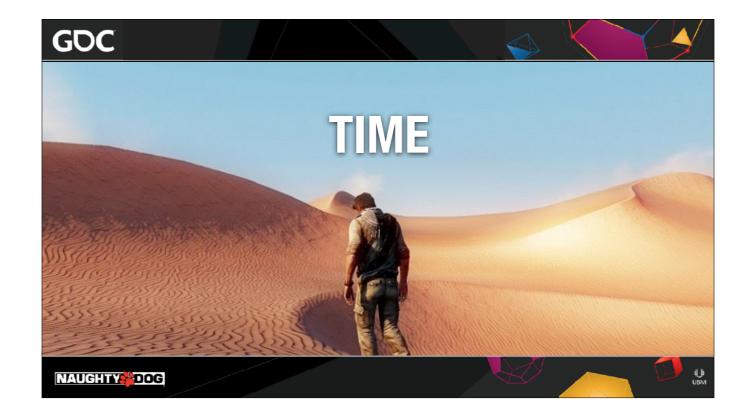
• There are even times where we've had to change the precise timing of certain IGCs or Cutscenes last-minute, meaning the audio files returned from all the territories have to be re-edited by hand on our end.



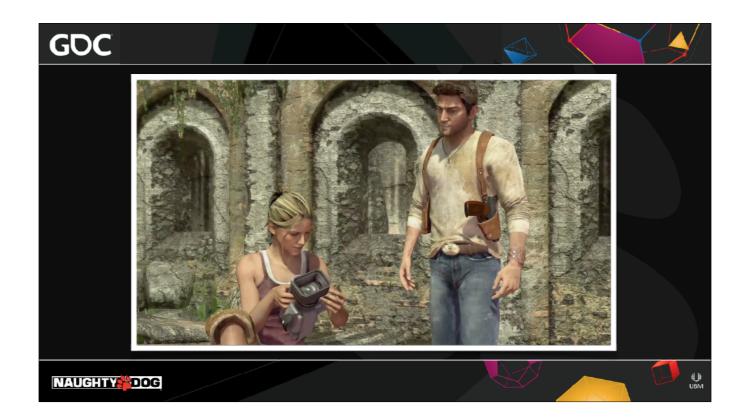
(Our Localization Manager keeps painstaking track of all of these changes so we know which files need the extra love.)



••• And, by the way, tracking all this madness—by any means necessary—is the biggest challenge of making things the "Naughty Dog way." The efforts of our minimal Production staff and various Leads are sometimes heroic.



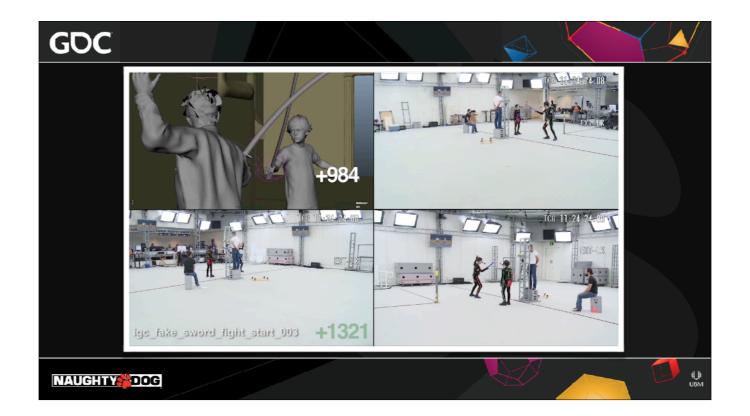
But as we near the end of production, the biggest limit on us is just time. At a certain point, we can only use animations that we've captured or lines that we've recorded. If we don't have the material, and can't make it up, sometimes that means living with tonal shifts, or big cuts of material that we painstakingly shot and mocap-ed.



{SILENT} https://youtu.be/YZ_2n3Y0WCA

*** Back on Uncharted 1, Design made a change to gameplay by **inserting combat into a sequence that was originally just quiet traversal**. Nate and Elena casually walked into an area and discovered a jet-ski. They weren't out of breath or recovering from a tense combat situation. But the gameplay **needed** a spike of energy and for the Player it worked, even though it caused a tonal mismatch with the Cutscene that followed.

Still, that late stage in the development process, we couldn't reshoot the scene. So this was logged as an example of how we could improve maintaining tone when going into a Cutscene in the future.

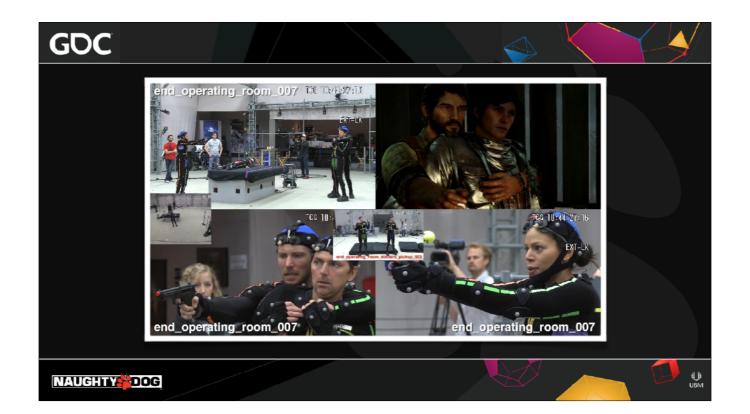


{PLAY} https://youtu.be/52SnBtlJioA

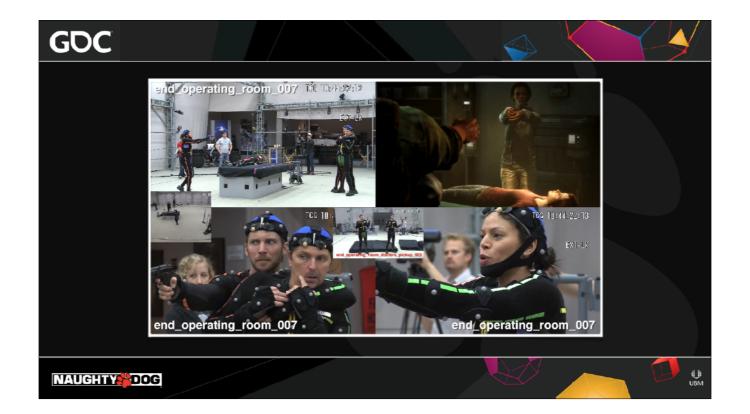
*** On Uncharted 4, we were working on an epic sword fight near the end of the game. And we'd planned to introduce those new mechanics during the Manor flashback with the kids. We even mocap-ed and recorded a bunch of material to make it happen. But when we looked at the amount of work it would take to polish and script all those animations, and time we had left to finish everything in the game, our Directors made a hard call and cut it. So the Player comes to the ending sword fight untrained, and can possibly be emotionally divested from the sequence as they struggle to learn new mechanics near the end of the game.



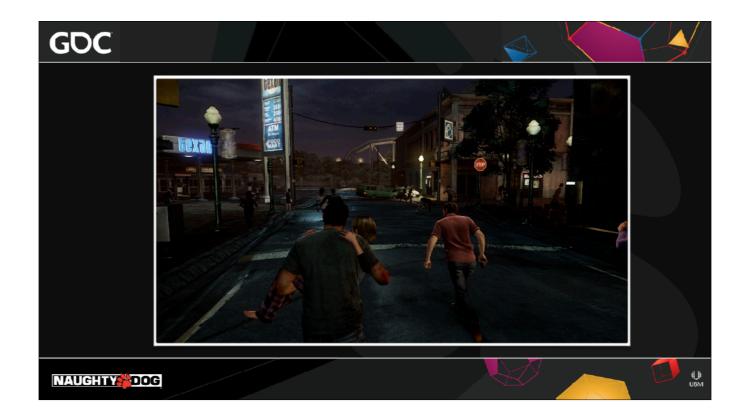
But, if we catch an issue early enough, it's been possible to make the hard choices required to maintain consistent tone and performances in a sequence that needs adjustment.



*** Nearing the end of The Last of Us, one of our Designers pointed out a **big flaw** we'd overlooked. The Firefly Hospital culminated with a Cutscene where Joel charged into the operating room to save Ellie...

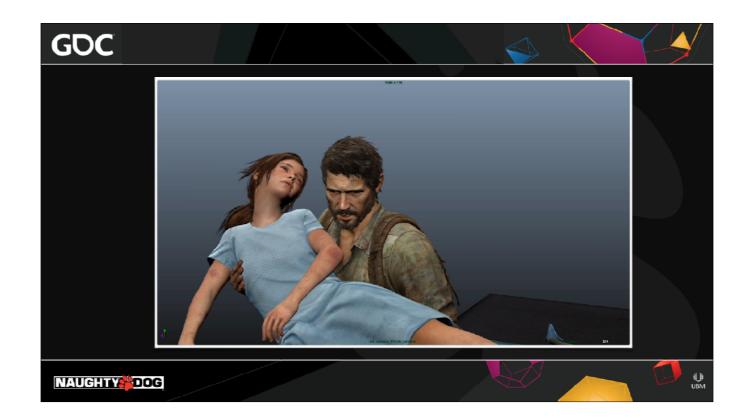


... In that scene, he kills the doctor (and eventually Marlene) right there, and then we smash-cut to him and Ellie driving away. It was a great Cutscene, and the performances we got of it, as-written, were excellent.

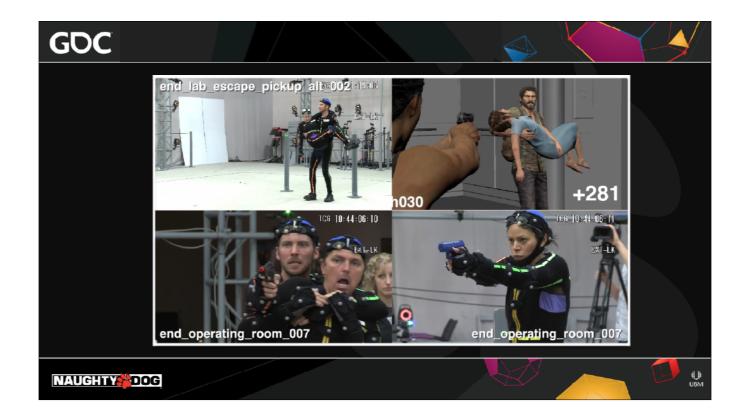


{SILENT} https://youtu.be/1Uu62po67H4

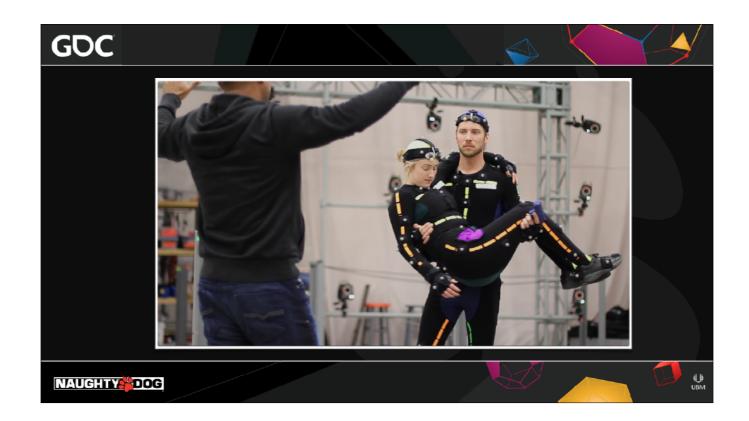
The problem was, as the Designer pointed out, that **the Player missed out on a big emotional payoff** that we hadn't even thought about. After carrying Joel's daughter at the beginning of the game, **the Player was set up to expect an echo of that event:...**



...that Joel would be carrying Ellie to safety at the end, redeeming some aspect his former loss.



So, with some hesitation, we looked at **what it might take to salvage what we could** from the existing Cutscene.



By keeping the Marlene animation and just picking up Joel—now carrying Ellie—...

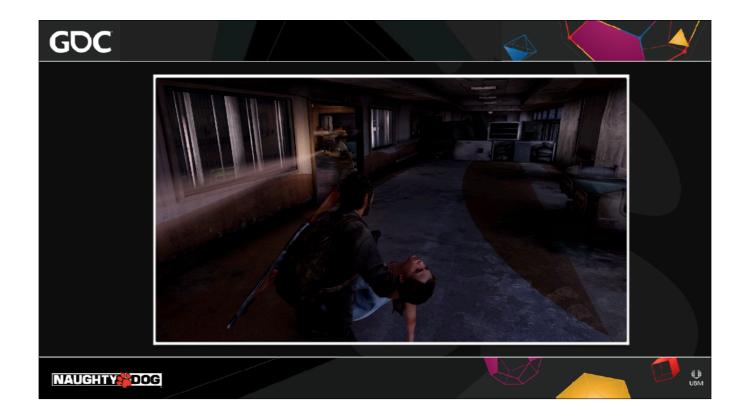


...we were able to move the main story beats of the Cutscene into the hospital parking garage.



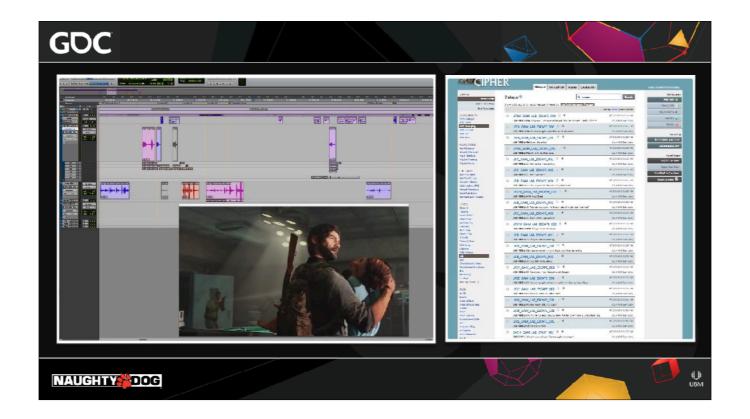
{SILENT} https://youtu.be/iBr61n6Oagl

This allowed for a brand new ending sequence, with Joel carrying Ellie past dozens of Firefly soldiers—who we were assured would refrain from shooting Joel because, seriously, he was carrying the cure for mankind!



So that's what we recorded all our dialogue for. And then, after Localization locked, it changed again! {SILENT} https://youtu.be/nGCqV/q_ME

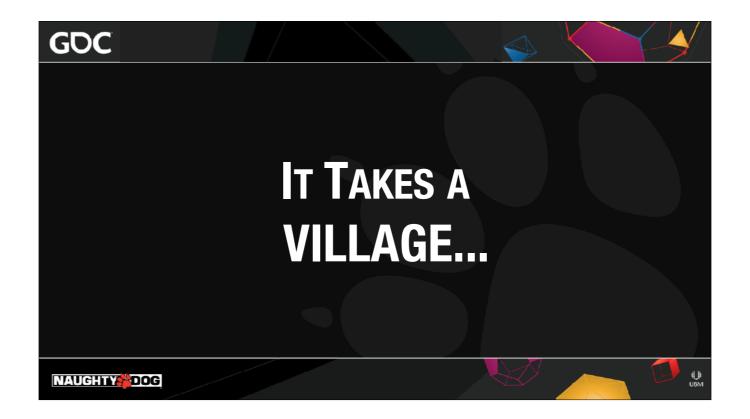
There wasn't enough tension if Joel could just walk out of the hospital with Ellie, so the actions of the Firefly soldiers was revamped so they could attack him. But that didn't match up with any of the dialogue we'd recorded!



It took digging through NPC material we had recorded for other parts of the game to make this moment—and the performances in it—feel **as equally polished** as the others in the game.



Ultimately, it was totally worth sacrificing some mocap and scrambling through folders of VO to create an experience that would capitalize on Player empathy to that degree.



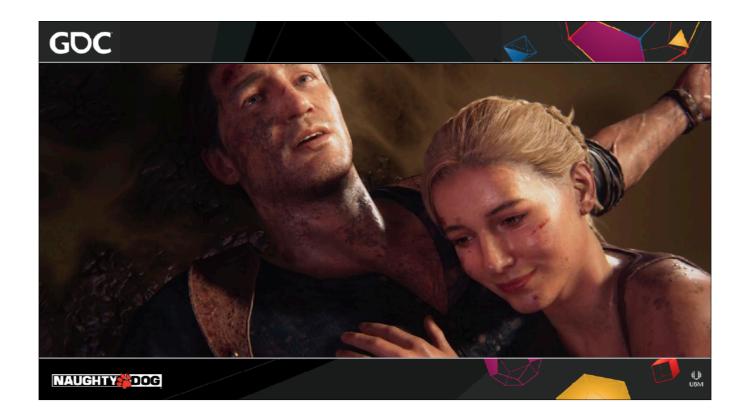
So...we do a lot of crazy shit!

And I'd just like to emphasize: the way we do it isn't the "only" way—or even the "right way"—to pull off realistic performances, especially depending on your Team size, budget, etc.

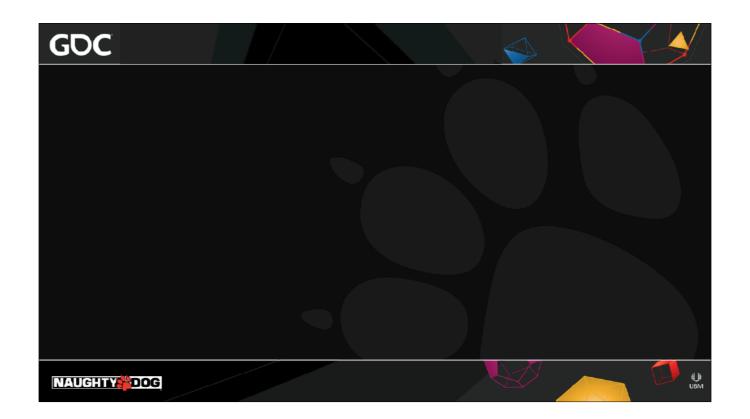
But, the level of flexibility we allow for story and gameplay to influence each other during development is why we feel we're able to strike a balance that we're proud of.



And it's also why executing these performances requires **constant collaboration and dedication from every discipline** on the Naughty Dog team. Writing, Acting, Rendering, Audio, Concept, Character, Programming, Animation, Design, Lighting, Production, even IT & Ops... it takes a veritable metropolis of talented men and women to make performances that the Player can *feel* when the controller is in their hands.



It's a constant struggle, but if we can truly move you—and maybe surprise you—then for us that's art worth making.



THANK YOU very much.