

The Trauma Floor

The secret lives of Facebook moderators in America

By Casey Newton Feb 25, 2019, 8:00am EST

Content warning: *This story contains discussion of serious mental health issues and racism.*

The panic attacks started after Chloe watched a man die.

She spent the past three and a half weeks in training, trying to harden herself against the daily onslaught of disturbing posts: the hate speech, the violent attacks, the graphic pornography. In a few more days, she will become a full-time Facebook content moderator, or what the company she works for, a professional services vendor named Cognizant, opaquely calls a “process executive.”

For this portion of her education, Chloe will have to moderate a Facebook post in front of her fellow trainees. When it’s her turn, she walks to the front of the room, where a monitor displays a video that has been posted to the world’s largest social network. None of the trainees have seen it before, Chloe included. She presses play.

The video depicts a man being murdered. Someone is stabbing him, dozens of times, while he screams and begs for his life. Chloe’s job is to tell the room whether this post should be removed. She knows that section 13 of the Facebook community standards prohibits videos that depict the murder of one or more people. When Chloe explains this to the class, she hears her voice shaking.

Returning to her seat, Chloe feels an overpowering urge to sob. Another trainee has gone up to review the next post, but Chloe cannot concentrate. She leaves the room, and begins to cry so hard that she has trouble breathing.

No one tries to comfort her. This is the job she was hired to do. And for the 1,000 people like Chloe moderating content for Facebook at the Phoenix site, and for 15,000 content reviewers around the world, today is just another day at the office.

Over the past three months, I interviewed a dozen current and former employees of Cognizant in Phoenix. All had signed non-disclosure agreements with Cognizant in which they pledged not to discuss their work for Facebook — or even acknowledge that Facebook is Cognizant’s client. The shroud of secrecy is meant to protect employees from users who may be angry about a content moderation decision and seek to resolve it with a known Facebook contractor. The NDAs are also meant to prevent contractors from sharing Facebook users’ personal information with the outside world, at a time of intense scrutiny over data privacy issues.

But the secrecy also insulates Cognizant and Facebook from criticism about their working conditions, moderators told me. They are pressured not to discuss the emotional toll that their job takes on them, even with loved ones, leading to increased feelings of isolation and anxiety. To protect them from potential retaliation, both from their employers and from Facebook users, I agreed to use pseudonyms for everyone named in this story except Cognizant's vice president of operations for business process services, Bob Duncan, and Facebook's director of global partner vendor management, Mark Davidson.

Collectively, the employees described a workplace that is perpetually teetering on the brink of chaos. It is an environment where workers cope by telling dark jokes about committing suicide, then smoke weed during breaks to numb their emotions. It's a place where employees can be fired for making just a few errors a week — and where those who remain live in fear of the former colleagues who return seeking vengeance.

It's a place where, in stark contrast to the perks lavished on Facebook employees, team leaders micromanage content moderators' every bathroom and prayer break; where employees, desperate for a dopamine rush amid the misery, have been found having sex inside stairwells and a room reserved for lactating mothers; where people develop severe anxiety while still in training, and continue to struggle with trauma symptoms long after they leave; and where the counseling that Cognizant offers them ends the moment they quit — or are simply let go.

The moderators told me it's a place where the conspiracy videos and memes that they see each day gradually lead them to embrace fringe views. One auditor walks the floor promoting the idea that the Earth is flat. A former employee told me he has begun to question certain aspects of the Holocaust. Another former employee, who told me he has mapped every escape route out of his house and sleeps with a gun at his side, said: "I no longer believe 9/11 was a terrorist attack."

Chloe cries for a while in the break room, and then in the bathroom, but begins to worry that she is missing too much training. She had been frantic for a job when she applied, as a recent college graduate with no other immediate prospects. When she becomes a full-time moderator, Chloe will make \$15 an hour — \$4 more than the minimum wage in Arizona, where she lives, and better than she can expect from most retail jobs.

The tears eventually stop coming, and her breathing returns to normal. When she goes back to the training room, one of her peers is discussing another violent video. She sees that a drone is shooting people from the air. Chloe watches the bodies go limp as they die.

She leaves the room again.

Eventually a supervisor finds her in the bathroom, and offers a weak hug. Cognizant makes a counselor available to employees, but only for part of the day, and he has yet to get to work. Chloe waits for him for the better part of an hour.

When the counselor sees her, he explains that she has had a panic attack. He tells her that, when she graduates, she will have more control over the Facebook videos than she had in the training

room. You will be able to pause the video, he tells her, or watch it without audio. Focus on your breathing, he says. Make sure you don't get too caught up in what you're watching.

"He said not to worry — that I could probably still do the job," Chloe says. Then she catches herself: "His *concern* was: don't worry, you can do the job."



On May 3, 2017, Mark Zuckerberg announced the expansion of Facebook's "community operations" team. The new employees, who would be added to 4,500 existing moderators, would be responsible for reviewing every piece of content reported for violating the company's community standards. By the end of 2018, in response to criticism of the prevalence of violent and exploitative content on the social network, Facebook had more than 30,000 employees working on safety and security — about half of whom were content moderators.

The moderators include some full-time employees, but Facebook relies heavily on contract labor to do the job. Ellen Silver, Facebook's vice president of operations, said in a blog post last year that the use of contract labor allowed Facebook to "scale globally" — to have content moderators working around the clock, evaluating posts in more than 50 languages, at more than 20 sites around the world.

The use of contract labor also has a practical benefit for Facebook: it is radically cheaper. The median Facebook employee earns \$240,000 annually in salary, bonuses, and stock options. A content moderator working for Cognizant in Arizona, on the other hand, will earn just \$28,800 per year. The arrangement helps Facebook maintain a high profit margin. In its most recent quarter, the company earned \$6.9 billion in profits, on \$16.9 billion in revenue. And while Zuckerberg had warned investors that Facebook's investment in security would reduce the company's profitability, profits were up 61 percent over the previous year.

Since 2014, when Adrian Chen detailed the harsh working conditions for content moderators at social networks for *Wired*, Facebook has been sensitive to the criticism that it is traumatizing some of its lowest-paid workers. In her blog post, Silver said that Facebook assesses potential moderators' "ability to deal with violent imagery," screening them for their coping skills.

Bob Duncan, who oversees Cognizant's content moderation operations in North America, says recruiters carefully explain the graphic nature of the job to applicants. "We share examples of the kinds of things you can see ... so that they have an understanding," he says. "The intention of all that is to ensure people understand it. And if they don't feel that work is potentially suited for them based on their situation, they can make those decisions as appropriate."

Until recently, most Facebook content moderation has been done outside the United States. But as Facebook's demand for labor has grown, it has expanded its domestic operations to include sites in California, Arizona, Texas, and Florida.

The United States is the company's home and one of the countries in which it is most popular, says Facebook's Davidson. American moderators are more likely to have the cultural context necessary to evaluate U.S. content that may involve bullying and hate speech, which often involve country-specific slang, he says.

Facebook also worked to build what Davidson calls "state-of-the-art facilities, so they replicated a Facebook office and had that Facebook look and feel to them. That was important because there's also a perception out there in the market sometimes ... that our people sit in very dark, dingy basements, lit only by a green screen. That's really not the case."

It is true that Cognizant's Phoenix location is neither dark nor dingy. And to the extent that it offers employees desks with computers on them, it may faintly resemble other Facebook offices. But while employees at Facebook's Menlo Park headquarters work in an airy, sunlit complex designed by Frank Gehry, its contractors in Arizona labor in an often cramped space where long lines for the few available bathroom stalls can take up most of employees' limited break time. And while Facebook employees enjoy a wide degree of freedom in how they manage their days, Cognizant workers' time is managed down to the second.

A content moderator named Miguel arrives for the day shift just before it begins, at 7 a.m. He's one of about 300 workers who will eventually filter into the workplace, which occupies two floors in a Phoenix office park.

Security personnel keep watch over the entrance, on the lookout for disgruntled ex-employees and Facebook users who might confront moderators over removed posts. Miguel badges in to the office and heads to the lockers. There are barely enough lockers to go around, so some employees have taken to keeping items in them overnight to ensure they will have one the next day.

The lockers occupy a narrow hallway that, during breaks, becomes choked with people. To protect the privacy of the Facebook users whose posts they review, workers are required to store their phones in lockers while they work.

Writing utensils and paper are also not allowed, in case Miguel might be tempted to write down a Facebook user's personal information. This policy extends to small paper scraps, such as gum wrappers. Smaller items, like hand lotion, are required to be placed in clear plastic bags so they are always visible to managers.

To accommodate four daily shifts — and high employee turnover — most people will not be assigned a permanent desk on what Cognizant calls “the production floor.” Instead, Miguel finds an open workstation and logs in to a piece of software known as the Single Review Tool, or SRT. When he is ready to work, he clicks a button labeled “resume reviewing,” and dives into the queue of posts.

Last April, a year after many of the documents had been published in the *Guardian*, Facebook made public the community standards by which it attempts to govern its 2.3 billion monthly users. In the months afterward, *Motherboard* and *Radiolab* published detailed investigations into the challenges of moderating such a vast amount of speech.

Those challenges include the sheer volume of posts; the need to train a global army of low-paid workers to consistently apply a single set of rules; near-daily changes and clarifications to those rules; a lack of cultural or political context on the part of the moderators; missing context in posts that makes their meaning ambiguous; and frequent disagreements among moderators about whether the rules should apply in individual cases.

Despite the high degree of difficulty in applying such a policy, Facebook has instructed Cognizant and its other contractors to emphasize a metric called “accuracy” over all else. Accuracy, in this case, means that when Facebook audits a subset of contractors’ decisions, its full-time employees agree with the contractors. The company has set an accuracy target of 95 percent, a number that always seems just out of reach. Cognizant has never hit the target for a sustained period of time — it usually floats in the high 80s or low 90s, and was hovering around 92 at press time.

Miguel diligently applies the policy — even though, he tells me, it often makes no sense to him.

A post calling someone “my favorite n-----” is allowed to stay up, because under the policy it is considered “explicitly positive content.”

“Autistic people should be sterilized” seems offensive to him, but it stays up as well. Autism is not a “protected characteristic” the way race and gender are, and so it doesn’t violate the policy. (“Men should be sterilized” would be taken down.)

In January, Facebook distributes a policy update stating that moderators should take into account recent romantic upheaval when evaluating posts that express hatred toward a gender. “I hate all men” has always violated the policy. But “I just broke up with my boyfriend, and I hate all men” no longer does.

Miguel works the posts in his queue. They arrive in no particular order at all.

Here is a racist joke. Here is a man having sex with a farm animal. Here is a graphic video of murder recorded by a drug cartel. Some of the posts Miguel reviews are on Facebook, where he says bullying and hate speech are more common; others are on Instagram, where users can post under pseudonyms, and tend to share more violence, nudity, and sexual activity.

Each post presents Miguel with two separate but related tests. First, he must determine whether a post violates the community standards. Then, he must select the correct reason why it violates the standards. If he accurately recognizes that a post should be removed, but selects the “wrong” reason, this will count against his accuracy score.

Miguel is very good at his job. He will take the correct action on each of these posts, striving to purge Facebook of its worst content while protecting the maximum amount of legitimate (if uncomfortable) speech. He will spend less than 30 seconds on each item, and he will do this up to 400 times a day.

When Miguel has a question, he raises his hand, and a “subject matter expert” (SME) — a contractor expected to have more comprehensive knowledge of Facebook’s policies, who makes \$1 more per hour than Miguel does — will walk over and assist him. This will cost Miguel time, though, and while he does not have a quota of posts to review, managers monitor his productivity, and ask him to explain himself when the number slips into the 200s.

From Miguel’s 1,500 or so weekly decisions, Facebook will randomly select 50 or 60 to audit. These posts will be reviewed by a second Cognizant employee — a quality assurance worker, known internally as a QA, who also makes \$1 per hour more than Miguel. Full-time Facebook employees then audit a subset of QA decisions, and from these collective deliberations, an accuracy score is generated.

Miguel takes a dim view of the accuracy figure.

“Accuracy is only judged by agreement. If me and the auditor both allow the obvious sale of heroin, Cognizant was ‘correct,’ because we both agreed,” he says. “This number is fake.”

Facebook’s single-minded focus on accuracy developed after sustaining years of criticism over its handling of moderation issues. With billions of new posts arriving each day, Facebook feels pressure on all sides. In some cases, the company has been criticized for not doing enough — as when United Nations investigators found that it had been complicit in spreading hate speech during the genocide of the Rohingya community in Myanmar. In others, it has been criticized for overreach — as when a moderator removed a post that excerpted the Declaration of Independence. (Thomas Jefferson was ultimately granted a posthumous exemption to Facebook’s speech guidelines, which prohibit the use of the phrase “Indian savages.”)

One reason moderators struggle to hit their accuracy target is that for any given policy enforcement decision, they have several sources of truth to consider.

The canonical source for enforcement is Facebook’s public community guidelines — which consist of two sets of documents: the publicly posted ones, and the longer internal guidelines, which offer more granular detail on complex issues. These documents are further augmented by a 15,000-word secondary document, called “Known Questions,” which offers additional commentary and guidance on thorny questions of moderation — a kind of Talmud to the community guidelines’ Torah. Known Questions used to occupy a single lengthy document that

moderators had to cross-reference daily; last year it was incorporated into the internal community guidelines for easier searching.

A third major source of truth is the discussions moderators have among themselves. During breaking news events, such as a mass shooting, moderators will try to reach a consensus on whether a graphic image meets the criteria to be deleted or marked as disturbing. But sometimes they reach the wrong consensus, moderators said, and managers have to walk the floor explaining the correct decision.

The fourth source is perhaps the most problematic: Facebook's own internal tools for distributing information. While official policy changes typically arrive every other Wednesday, incremental guidance about developing issues is distributed on a near-daily basis. Often, this guidance is posted to Workplace, the enterprise version of Facebook that the company introduced in 2016. Like Facebook itself, Workplace has an algorithmic News Feed that displays posts based on engagement. During a breaking news event, such as a mass shooting, managers will often post conflicting information about how to moderate individual pieces of content, which then appear out of chronological order on Workplace. Six current and former employees told me that they had made moderation mistakes based on seeing an outdated post at the top of their feed. At times, it feels as if Facebook's own product is working against them. The irony is not lost on the moderators.

"It happened all the time," says Diana, a former moderator. "It was horrible — one of the worst things I had to personally deal with, to do my job properly." During times of national tragedy, such as the 2017 Las Vegas shooting, managers would tell moderators to remove a video — and then, in a separate post a few hours later, to leave it up. The moderators would make a decision based on whichever post Workplace served up.

"It was such a big mess," Diana says. "We're supposed to be up to par with our decision making, and it was messing up our numbers."

Workplace posts about policy changes are supplemented by occasional slide decks that are shared with Cognizant workers about special topics in moderation — often tied to grim anniversaries, such as the Parkland shooting. But these presentations and other supplementary materials often contain embarrassing errors, moderators told me. Over the past year, communications from Facebook incorrectly identified certain U.S. representatives as senators; misstated the date of an election; and gave the wrong name for the high school at which the Parkland shooting took place. (It is Marjory Stoneman Douglas High School, not "Stoneham Douglas High School.")

Even with an ever-changing rulebook, moderators are granted only the slimmest margins of error. The job resembles a high-stakes video game in which you start out with 100 points — a perfect accuracy score — and then scratch and claw to keep as many of those points as you can. Because once you fall below 95, your job is at risk.

If a quality assurance manager marks Miguel's decision wrong, he can appeal the decision. Getting the QA to agree with you is known as "getting the point back." In the short term, an

“error” is whatever a QA says it is, and so moderators have good reason to appeal every time they are marked wrong. (Recently, Cognizant made it even harder to get a point back, by requiring moderators to first get a SME to approve their appeal before it would be forwarded to the QA.)

Sometimes, questions about confusing subjects are escalated to Facebook. But every moderator I asked about this said that Cognizant managers discourage employees from raising issues to the client, apparently out of fear that too many questions would annoy Facebook.

This has resulted in Cognizant inventing policy on the fly. When the community standards did not explicitly prohibit erotic asphyxiation, three former moderators told me, a team leader declared that images depicting choking would be permitted unless the fingers depressed the skin of the person being choked.

Before workers are fired, they are offered coaching and placed into a remedial program designed to make sure they master the policy. But often this serves as a pretext for managing workers out of the job, three former moderators told me. Other times, contractors who have missed too many points will escalate their appeals to Facebook for a final decision. But the company does not always get through the backlog of requests before the employee in question is fired, I was told.

Officially, moderators are prohibited from approaching QAs and lobbying them to reverse a decision. But it is still a regular occurrence, two former QAs told me.

One, named Randy, would sometimes return to his car at the end of a work day to find moderators waiting for him. Five or six times over the course of a year, someone would attempt to intimidate him into changing his ruling. “They would confront me in the parking lot and tell me they were going to beat the shit out of me,” he says. “There wasn’t even a single instance where it was respectful or nice. It was just, *You audited me wrong! That was a boob! That was full areola, come on man!*”

Fearing for his safety, Randy began bringing a concealed gun to work. Fired employees regularly threatened to return to work and harm their old colleagues, and Randy believed that some of them were serious. A former coworker told me she was aware that Randy brought a gun to work, and approved of it, fearing on-site security would not be sufficient in the case of an attack.

Cognizant’s Duncan told me the company would investigate the various safety and management issues that moderators had disclosed to me. He said bringing a gun to work was a violation of policy and that, had management been aware of it, they would have intervened and taken action against the employee.

Randy quit after a year. He never had occasion to fire the gun, but his anxiety lingers.

“Part of the reason I left was how unsafe I felt in my own home and my own skin,” he says.

Before Miguel can take a break, he clicks a browser extension to let Cognizant know he is leaving his desk. (“That’s a standard thing in this type of industry,” Facebook’s Davidson tells me. “To be able to track, so you know where your workforce is.”)

Miguel is allowed two 15-minute breaks, and one 30-minute lunch. During breaks, he often finds long lines for the restrooms. Hundreds of employees share just one urinal and two stalls in the men’s room, and three stalls in the women’s. Cognizant eventually allowed employees to use a restroom on another floor, but getting there and back will take Miguel precious minutes. By the time he has used the restroom and fought the crowd to his locker, he might have five minutes to look at his phone before returning to his desk.

Miguel is also allotted nine minutes per day of “wellness time,” which he is supposed to use if he feels traumatized and needs to step away from his desk. Several moderators told me that they routinely used their wellness time to go to the restroom when lines were shorter. But management eventually realized what they were doing, and ordered employees not to use wellness time to relieve themselves. (Recently a group of Facebook moderators hired through Accenture in Austin complained about “inhumane” conditions related to break periods; Facebook attributed the issue to a misunderstanding of its policies.)

At the Phoenix site, Muslim workers who used wellness time to perform one of their five daily prayers were told to stop the practice and do it on their other break time instead, current and former employees told me. It was unclear to the employees I spoke with why their managers did not consider prayer to be a valid use of the wellness program. (Cognizant did not offer a comment about these incidents, although a person familiar with one case told me a worker requested more than 40 minutes for daily prayer, which the company considered excessive.)

Cognizant employees are told to cope with the stress of the jobs by visiting counselors, when they are available; by calling a hotline; and by using an employee assistance program, which offers a handful of therapy sessions. More recently, yoga and other therapeutic activities have been added to the work week. But aside from occasional visits to the counselor, six employees I spoke with told me they found these resources inadequate. They told me they coped with the stress of the job in other ways: with sex, drugs, and offensive jokes.

Among the places that Cognizant employees have been found having sex at work: the bathroom stalls, the stairwells, the parking garage, and the room reserved for lactating mothers. In early 2018, the security team sent out a memo to managers alerting them to the behavior, a person familiar with the matter told me. The solution: management removed door locks from the mother’s room and from a handful of other private rooms. (The mother’s room now locks again, but would-be users must first check out a key from an administrator.)

A former moderator named Sara said that the secrecy around their work, coupled with the difficulty of the job, forged strong bonds between employees. “You get really close to your coworkers really quickly,” she says. “If you’re not allowed to talk to your friends or family about your job, that’s going to create some distance. You might feel closer to these people. It feels like an emotional connection, when in reality you’re just trauma bonding.”

Employees also cope using drugs and alcohol, both on and off campus. One former moderator, Li, told me he used marijuana on the job almost daily, through a vaporizer. During breaks, he says, small groups of employees often head outside and smoke. (Medical marijuana use is legal in Arizona.)

“I can’t even tell you how many people I’ve smoked with,” Li says. “It’s so sad, when I think back about it — it really does hurt my heart. We’d go down and get stoned and go back to work. That’s not professional. Knowing that the content moderators for the world’s biggest social media platform are doing this on the job, while they are moderating content ...”

He trailed off.

Li, who worked as a moderator for about a year, was one of several employees who said the workplace was rife with pitch-black humor. Employees would compete to send each other the most racist or offensive memes, he said, in an effort to lighten the mood. As an ethnic minority, Li was a frequent target of his coworkers, and he embraced what he saw as good-natured racist jokes at his expense, he says.

But over time, he grew concerned for his mental health.

“We were doing something that was darkening our soul — or whatever you call it,” he says. “What else do you do at that point? The one thing that makes us laugh is actually damaging us. I had to watch myself when I was joking around in public. I would accidentally say [offensive] things all the time — and then be like, *Oh shit, I’m at the grocery store. I cannot be talking like this.*”

Jokes about self-harm were also common. “Drinking to forget,” Sara heard a coworker once say, when the counselor asked him how he was doing. (The counselor did not invite the employee in for further discussion.) On bad days, Sara says, people would talk about it being “time to go hang out on the roof” — the joke being that Cognizant employees might one day throw themselves off it.

One day, Sara said, moderators looked up from their computers to see a man standing on top of the office building next door. Most of them had watched hundreds of suicides that began just this way. The moderators got up and hurried to the windows.

The man didn’t jump, though. Eventually everyone realized that he was a fellow employee, taking a break.

Like most of the former moderators I spoke with, Chloe quit after about a year.

Among other things, she had grown concerned about the spread of conspiracy theories among her colleagues. One QA often discussed his belief that the Earth is flat with colleagues, and “was actively trying to recruit other people” into believing, another moderator told me. One of Miguel’s colleagues once referred casually to “the Holofoax,” in what Miguel took as a signal that the man was a Holocaust denier.

Conspiracy theories were often well received on the production floor, six moderators told me. After the Parkland shooting last year, moderators were initially horrified by the attacks. But as more conspiracy content was posted to Facebook and Instagram, some of Chloe's colleagues began expressing doubts.

"People really started to believe these posts they were supposed to be moderating," she says. "They were saying, 'Oh gosh, they weren't really there. Look at this CNN video of David Hogg — he's too old to be in school.' People started Googling things instead of doing their jobs and looking into conspiracy theories about them. We were like, 'Guys, no, this is the crazy stuff we're supposed to be moderating. What are you doing?'"

Most of all, though, Chloe worried about the long-term impacts of the job on her mental health. Several moderators told me they experienced symptoms of secondary traumatic stress — a disorder that can result from observing firsthand trauma experienced by others. The disorder, whose symptoms can be identical to post-traumatic stress disorder, is often seen in physicians, psychotherapists, and social workers. People experiencing secondary traumatic stress report feelings of anxiety, sleep loss, loneliness, and dissociation, among other ailments.

Last year, a former Facebook moderator in California sued the company, saying her job as a contractor with the firm Pro Unlimited had left her with PTSD. In the complaint, her lawyers said she "seeks to protect herself from the dangers of psychological trauma resulting from Facebook's failure to provide a safe workplace for the thousands of contractors who are entrusted to provide the safest possible environment for Facebook users." (The suit is still unresolved.)

Chloe has experienced trauma symptoms in the months since leaving her job. She started to have a panic attack in a movie theater during the film *Mother!*, when a violent stabbing spree triggered a memory of that first video she moderated in front of her fellow trainees. Another time, she was sleeping on the couch when she heard machine gun fire, and had a panic attack. Someone in her house had turned on a violent TV show. She "started freaking out," she says. "I was begging them to shut it off."

The attacks make her think of her fellow trainees, especially the ones who fail out of the program before they can start. "A lot of people don't actually make it through the training," she says. "They go through those four weeks and then they get fired. They could have had that same experience that I did, and had absolutely no access to counselors after that."

Last week, Davidson told me, Facebook began surveying a test group of moderators to measure what the company calls their "resiliency" — their ability to bounce back from seeing traumatic content and continue doing their jobs. The company hopes to expand the test to all of its moderators globally, he said.

Randy also left after about a year. Like Chloe, he had been traumatized by a video of a stabbing. The victim had been about his age, and he remembers hearing the man crying for his mother as he died.

“Every day I see that,” Randy says, “I have a genuine fear over knives. I like cooking — getting back into the kitchen and being around the knives is really hard for me.”

The job also changed the way he saw the world. After he saw so many videos saying that 9/11 was not a terrorist attack, he came to believe them. Conspiracy videos about the Las Vegas massacre were also very persuasive, he says, and he now believes that multiple shooters were responsible for the attack. (The FBI found that the massacre was the work of a single gunman.)

Randy now sleeps with a gun at his side. He runs mental drills about how he would escape his home in the event that it were attacked. When he wakes up in the morning, he sweeps the house with his gun raised, looking for invaders.

He has recently begun seeing a new therapist, after being diagnosed with PTSD and generalized anxiety disorder.

“I’m fucked up, man,” Randy says. “My mental health — it’s just so up and down. One day I can be really happy, and doing really good. The next day, I’m more or less of a zombie. It’s not that I’m depressed. I’m just stuck.”

He adds: “I don’t think it’s possible to do the job and not come out of it with some acute stress disorder or PTSD.”

A common complaint of the moderators I spoke with was that the on-site counselors were largely passive, relying on workers to recognize the signs of anxiety and depression and seek help.

“There was nothing that they were doing for us,” Li says, “other than expecting us to be able to identify when we’re broken. Most of the people there that are deteriorating — they don’t even see it. And that’s what kills me.”

Last week, after I told Facebook about my conversations with moderators, the company invited me to Phoenix to see the site for myself. It is the first time Facebook has allowed a reporter to visit an American content moderation site since the company began building dedicated facilities here two years ago. A spokeswoman who met me at the site says that the stories I have been told do not reflect the day-to-day experiences of most of its contractors, either at Phoenix or at its other sites around the world.

The day before I arrived at the office park where Cognizant resides, one source tells me, new motivational posters were hung up on the walls. On the whole, the space is much more colorful than I expect. A neon wall chart outlines the month’s activities, which read like a cross between the activities at summer camp and a senior center: yoga, pet therapy, meditation, and a *Mean Girls*-inspired event called On Wednesdays We Wear Pink. The day I was there marked the end of Random Acts of Kindness Week, in which employees were encouraged to write inspirational messages on colorful cards, and attach them to a wall with a piece of candy.

After meetings with executives from Cognizant and Facebook, I interview five workers who had volunteered to speak with me. They stream into a conference room, along with the man who is

responsible for running the site. With their boss sitting at their side, employees acknowledge the challenges of the job but tell me they feel safe, supported, and believe the job will lead to better-paying opportunities — within Cognizant, if not Facebook.

Brad, who holds the title of policy manager, tells me that the majority of content that he and his colleagues review is essentially benign, and warns me against overstating the mental health risks of doing the job.

“There’s this perception that we’re bombarded by these graphic images and content all the time, when in fact the opposite is the truth,” says Brad, who has worked on the site for nearly two years. “Most of the stuff we see is mild, very mild. It’s people going on rants. It’s people reporting photos or videos simply because they don’t want to see it — not because there’s any issue with the content. That’s really the majority of the stuff that we see.”

When I ask about the high difficulty of applying the policy, a reviewer named Michael says that he regularly finds himself stumped by tricky decisions. “There is an infinite possibility of what’s gonna be the next job, and that does create an essence of chaos,” he says. “But it also keeps it interesting. You’re never going to go an entire shift already knowing the answer to every question.”

In any case, Michael says, he enjoys the work better than he did at his last job, at Walmart, where he was often berated by customers. “I do not have people yelling in my face,” he says.

The moderators stream out, and I’m introduced to two counselors on the site, including the doctor who started the on-site counseling program here. Both ask me not to use their real names. They tell me that they check in with every employee every day. They say that the combination of on-site services, a hotline, and an employee assistance program are sufficient to protect workers’ well-being.

When I ask about the risks of contractors developing PTSD, a counselor I’ll call Logan tells me about a different psychological phenomenon: “post-traumatic growth,” an effect whereby some trauma victims emerge from the experience feeling stronger than before. The example he gives me is that of Malala Yousafzai, the women’s education activist, who was shot in the head as a teenager by the Taliban.

“That’s an extremely traumatic event that she experienced in her life,” Logan says. “It seems like she came back extremely resilient and strong. She won a Nobel Peace Prize... So there are many examples of people that experience difficult times and come back stronger than before.”

The day ends with a tour, in which I walk the production floor and talk with other employees. I am struck by how young they are: almost everyone seems to be in their twenties or early thirties. All work stops while I’m on the floor, to ensure I do not see any Facebook user’s private information, and so employees chat amiably with their deskmates as I walk by. I take note of the posters. One, from Cognizant, bears the enigmatic slogan “empathy at scale.” Another, made famous by Facebook COO Sheryl Sandberg, reads “What would you do if you weren’t afraid?”

It makes me think of Randy and his gun.

Everyone I meet at the site expresses great care for the employees, and appears to be doing their best for them, within the context of the system they have all been plugged into. Facebook takes pride in the fact that it pays contractors at least 20 percent above minimum wage at all of its content review sites, provides full healthcare benefits, and offers mental health resources that far exceed that of the larger call center industry.

And yet the more moderators I spoke with, the more I came to doubt the use of the call center model for content moderation. This model has long been standard across big tech companies — it's also used by Twitter and Google, and therefore YouTube. Beyond cost savings, the benefit of outsourcing is that it allows tech companies to rapidly expand their services into new markets and languages. But it also entrusts essential questions of speech and safety to people who are paid as if they were handling customer service calls for Best Buy.

Every moderator I spoke with took great pride in their work, and talked about the job with profound seriousness. They wished only that Facebook employees would think of them as peers, and to treat them with something resembling equality.

“If we weren't there doing that job, Facebook would be so ugly,” Li says. “We're seeing all that stuff on their behalf. And hell yeah, we make some wrong calls. But people don't know that there's actually human beings behind those seats.”

That people don't know there are human beings doing this work is, of course, by design. Facebook would rather talk about its advancements in artificial intelligence, and dangle the prospect that its reliance on human moderators will decline over time.

But given the limits of the technology, and the infinite varieties of human speech, such a day appears to be very far away. In the meantime, the call center model of content moderation is taking an ugly toll on many of its workers. As first responders on platforms with billions of users, they are performing a critical function of modern civil society, while being paid less than half as much as many others who work on the front lines. They do the work as long as they can — and when they leave, an NDA ensures that they retreat even further into the shadows.

To Facebook, it will seem as if they never worked there at all. Technically, they never did.