Convo

Andrew Madrick

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Design

University of Washington 2022

Committee:

Karen Cheng

Audrey Desjardins

Program Authorized to Offer Degree: School Of Art + Art History + Design

University of Washington

Abstract

Convo

Andrew Madrick

Chair of the Supervisory Committee: Karen Cheng

Design

The way in which we talk to each other online is different from how we talk to each other in real life. Social media incentivizes conversations that are more emotional and tribal than in-person dialogue. (Brady et al., 2017; Rathje et al., 2021). In an effort to make online conversations less performative, I have created a tool, called Convo, that encourages individuals to have more logical, nuanced and civil conversations online. Through peer–to–peer feedback, Convo offers a crowdsourced and self–sustaining model for online interaction that benefits three key stakeholders: platforms, active posters and passive viewers. Convo encourages problematic online posters to either modify their posts or to take an "off-ramp" to direct messaging by showing peer feedback from other readers/viewers.

Table of Contents

Background

Like many Americans, the 2016 presidential election made me curious to know more about the relationship between social media and society. More than any other aspect of the problem space, I became obsessed with the way in which we talk to each other online. This is not limited to interaction on social media (Twitter, Facebook, etc.). Even in more benign corners of the internet (examples? NYTimes comments?), people seem to speak to one another online in ways that are very different than in person. Researchers believe that this kind of behavior is, at least in part, an outgrowth of the incentive structures that surround online conversation and the way in which platforms moderate these conversations (add citations)

Incentive Structures for Online Conversation

Two strong incentives that affect behavior online are the “Like” and “Share” features found on many social media platforms. These features allow for a public, quantifiable way to measure the wider public’s impression of a post online. Moreover, these features gamify the process of having a conversation, encouraging people to behave in ways that no longer resemble how individuals interact with one another in person. As Jonathan Haidt says, these kinds of features encourage individuals to be “guided not just by their true preferences but by their past experiences of reward and punishment, and their prediction of how others would react to each new action” (Haidt, 2022).

This makes intuitive sense: humans are social creatures that have built society on a foundation of cooperation (“The Cooperative Human,” 2018). We can extrapolate from this that being a cooperative species means that we largely behave in ways that will benefit our standing in a group setting. Quantifiable metrics for others approval of our behavior (Likes and Shares) allow for real time feedback about how others feel about our social performance.

Now, human interaction online has rarely been a source for optimism about our future as a species. However, introducing the Like and Share features online undoubtedly pushed online communication in a certain direction. A quick glance of your Twitter feed will likely include a mix of hyperbole and nervous energy. Researchers believe that posts that trigger emotions are most likely to be Liked or Shared (Brady et al., 2017), especially anger at out-groups (Rathje et al., 2021). Therefore , we can see that, in the quest for likes and shares, a poster (who relies on “their past experiences of reward and punishment”) will post something likely to spark emotion or out–group hate, often in the form of hyperbole. If the emotion or hate focuses on what the public deems as noble, the poster will achieve initial goal of internet fame and popularity. However, if they misfire or say the

wrong thing out of ignorance or a lapse of judgment, they can face “cancellation” in various manifestations (Haidt, 2022).

The Role of Moderation

Social media companies are businesses—big businesses that have “customer” bases that exceed the population of entire continents (*Facebook MAU Worldwide 2022*, n.d.). Platforms like Facebook and Twitter are largely monetized through advertising. Astonishingly, Facebook made over $100 billion from advertising alone in 2021 (*Facebook Ad Revenue 2009-2020*, n.d.). Although contentious posts, that might include out–group hate, for example, are effective at engaging users, a site dominated by this kind of content might not attract the revenue from advertisers.

As a result, moderation is often used to keep a baseline of civility and safety on social media platforms. To understand how Convo fits into the current online ecosystem, it’s crucial to understand the role that moderation currently plays in policing and promoting different kinds of content, regardless of its negative or positive valence.

At present, moderation exists as top–down and bottom–up systems (Roberts, 2017). The top–down version of moderation is often referred to as “supervisor moderation”. This system works by combining the reach of artificial intelligence with the capacity for nuance of assigned human moderators. Supervisor-based systems can be efficient and often allow a site to only host content that aligns with the business interests of the platform. However, as sites like Twitter are increasingly referred to as the “de facto public square”, the use of supervisors in free conversation seems at odds with the democratic basis of free and open discussion and debate. Moreover, supervisors (humans or AI), often lack an appropriate level of context for properly understanding the intricacies of various posts or discussions online. This may lead to awkward, incorrect and inequitable consequences for those whose behavior is flagged or deemed problematic (Roberts, 2017).

The bottom–up alternative to supervisor moderation is known as “distributed moderation”. This method works by allowing users to moderate each another with little to no intervention on the side of the hosting platform. The best-known version of distributed moderation is Wikipedia, where anyone can post anything. However, due to the sheer number of other editors on the platform, incorrect or unsubstantiated posts are flagged and eliminated relatively quickly. A distributed system is, by its very nature, a hyper–democratic form of moderation. Because it relies on real people, none more powerful than the next, equity is built into the system. Additionally, distributed moderation is relatively inexpensive and easy to set up if a robust user base exists. However, distributed moderation is not perfect. These systems are not explicitly beholden to a

shared set of values or rules, creating a kind of vacuum of meaning and direction(Roberts, 2017).

The Opportunity

Convo as a Solution

With my thesis, I propose an alternative model of how we interact with content online: **Convo**.

Convo is a tool that works with existing incentives and models of moderation to encourage meaningful conversations online. Through peer–to–peer feedback, Convo offers a crowdsourced and therefore self–sustaining model for online interaction that benefits three key stakeholders: platforms, active posters and passive viewers.

Stakeholders

There are three primary stakeholders that benefit when Convo is adopted: Platforms, Viewers and Posters.

Platforms are the online arenas in which public conversation and discourse is taking place. Convo is best suited not for the large social media platforms (such as Facebook, Twitter, Nextdoor, Quora, Reddit) but for editorial platforms such as newspaper article comments, or Facebook groups. By implementing Convo, editorial platforms can raise the quality of the comments that appear online by empowering people to moderate one another, signaling trust in the process. At the same time, the off–ramps offered by Convo keep people online. Back–channeling difficult conversations empowers and engages individuals that might otherwise risk cancellation or public shame.

Posters are the individuals who are actively participating in the conversation. Posters are enabled to have difficult conversations that might otherwise get them kicked off a platform. When the conversations get tough, Posters don't have to walk on eggshells, being afraid to say the wrong thing. Rather, Convo educates these individuals by enabling them to receive detailed feedback on why the conversation might not be going in a desirable direction. Again, the low–stakes off ramps offered by Convo (Rooms and Exploring) provides the posters with a way out of the game of posting online.

Viewers are those who watch a conversation from the outside looking in. By contributing meaningful feedback to others online, Viewers are empowered with increased autonomy and influence within the arena of conversation. Not only does Convo legitimize the voice

of viewers, it also makes their experience online more personal, customizable and interesting. By back–channeling bad conversations, Viewers don't have to wade through endless toxic or irrelevant conversations to find meaningful content. Further, public tags allow for easily searchable posts and feeds, again, increasing the ease of use and enjoyment that comes out of watching a conversation unfold.

How does Convo work?

Convo works by combining the both distributed and supervisor based moderation systems. As a starting point, Convo works alongside automated moderation tools like Google's Jigsaw or Openweb to ensure a baseline level of safety within comment sections.

However, safe conversations are not synonymous with engaging, stimulating or meaningful conversations. This is where we see the value of a tool like Convo. By providing a bottom–up, peer–to–peer form of moderation focused primarily on finding and facilitating meaningful conversations, Convo offers platforms the ability to drastically increase the quality of their comment threads.

The Convo Process

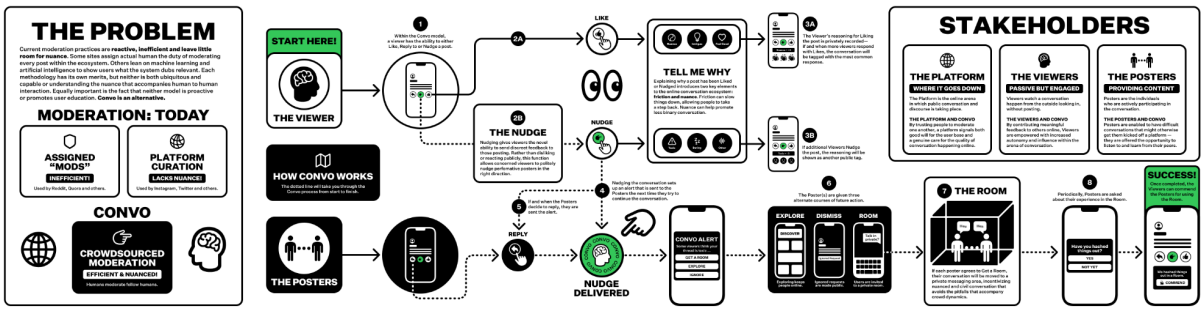

Figure 1: The Convo Process

ONE: Likes and Nudges

Within the Convo model, people who want to engage with a post but do not wish to reply publicly are given fairly standard response tools: Liking and Nudging. Liking a post is the positive response while Nudging is its negative counterpart. Besides the vocabulary of the "Nudge", these are standard online reactions not dissimilar to up and down voting.

TWO: Adding Richness

Where Convo breaks from standard response interactions is what happens after an individual presses the Like or Nudge button. Before either of these responses can be registered, an individual *must* give a reason for the given response. This process introduces two key elements to the conversational ecosystem: friction and nuance. Friction slows things down. Within the Convo model, an individual cannot mindlessly click one button and continue scrolling. Rather, this added step activates the viewer, prompting them, ideally, to think about why they responded the way they did.

Nuance, on the other hand, adds a richness to the kind of responses regularly registered online. By inviting people to qualify why exactly they responded the way they did, Convo encourages viewers to dig deeper than binomial, black and white modes of thinking. This results in a more organized, less hyperbolic public feed. As we'll see, introducing this added layer of data allows the platform to offer individuals the ability to sort posts and conversations in ways not currently offered by most major platforms (which often serve content based on personalized curation or chronological order).

THREE: Liking and Sorting

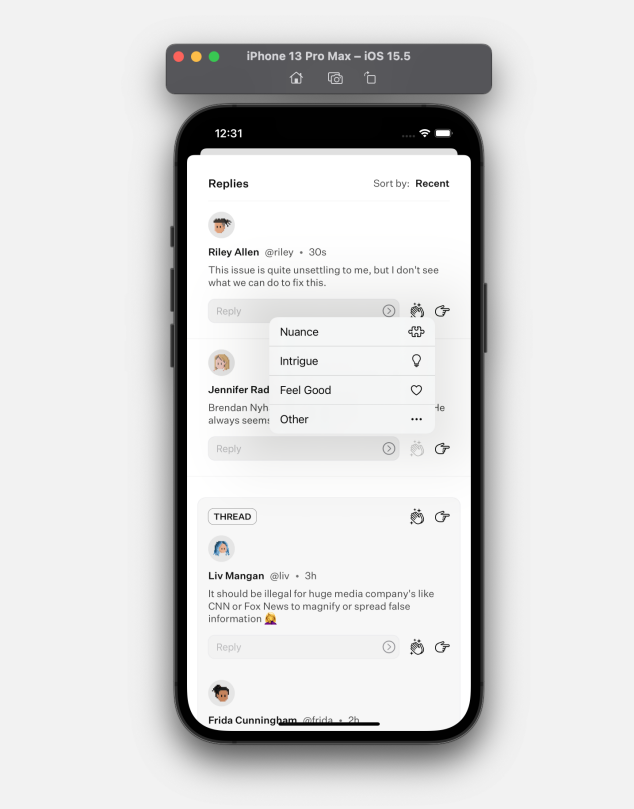

Figure 2: Liking

Moving on, if a viewer Likes a post or a conversation, and indicates why they did so, the reasoning is privately recorded. Again, herein lies a novel feature. Being socially motivated creatures, humans will modify their behavior based on the attitude of the crowd. If an individual sees that a post is gaining positive or negative momentum for one reason or another, it is likely that they will be inclined to pile on. For this reason, Conco doesn't show a number Likes or Nudges, it simply shows the most common response as a tag.

Posts or conversations will be publicly tagged with the most common reason for Liking the post. "Publicly tagged" in this sense, means that the post will include a small badge that indicates the reason why.

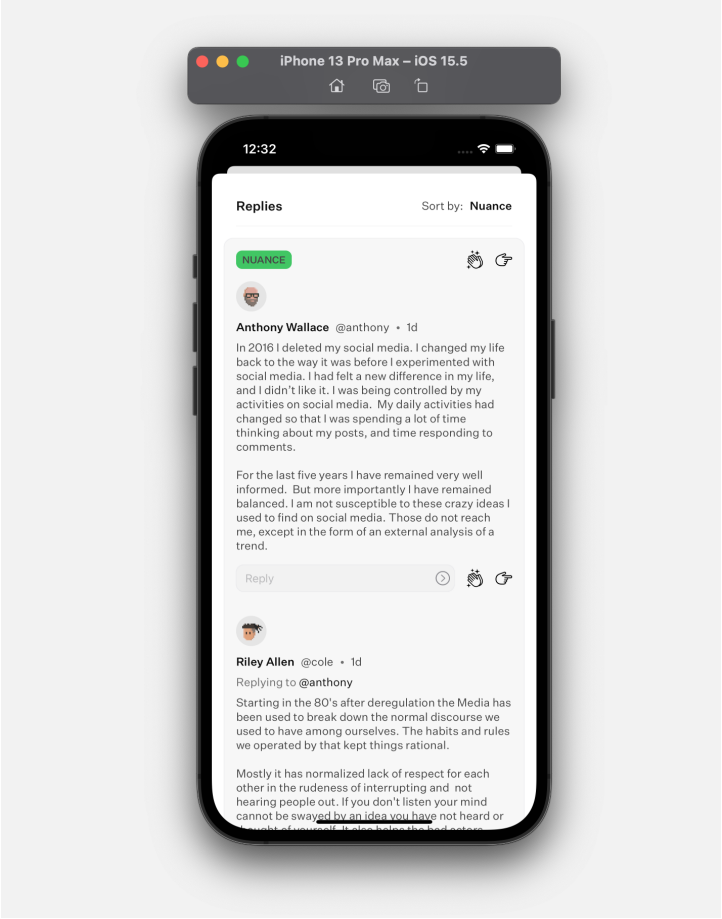

Figure 3: Sorting

Liking also provides an opportunity for another added benefit: allowing users to search for threads and comments based on positive attributes, rather than the more common

methods of temporal or popularity based sorting (as seen on platforms like Twitter and Reddit). This again, empowers the user. An individual is able to search for exactly what they might find interesting and avoid the pitfalls of sifting through endless bad content.

FOUR: The Nudge

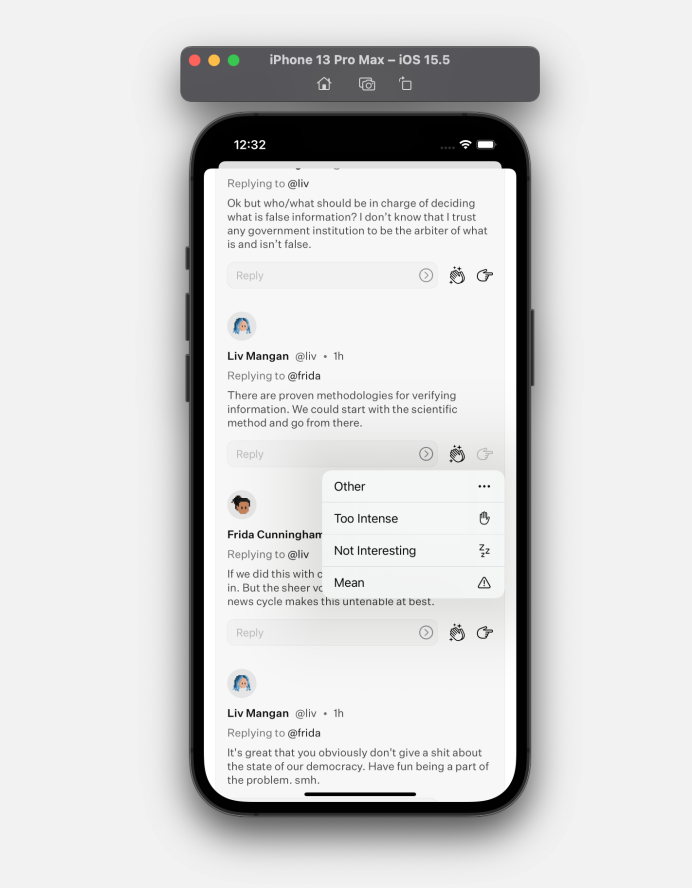

Figure 4: The Nudge

The other option for Viewers is to Nudge a post. A Nudge is more than just a dislike or a downvote. Rather, it is a suggestion from Viewers that the Poster should choose an alternate course of action. It indicates that the behavior of the Poster leaves something to be desired in the realm of public conversation. When Viewers Nudge a post, they are empowered to send private and discreet feedback to Posters, giving them a voice without having to speak up publicly.

As Nudges add up, the Poster is able to view the responses coming in. This creates a low–stakes form of social pressure on the Poster before things go public. If this individual

sees that a large number of peers would prefer them to behave differently, they might be pressured to modify their behavior without top–down, system–based interventions.

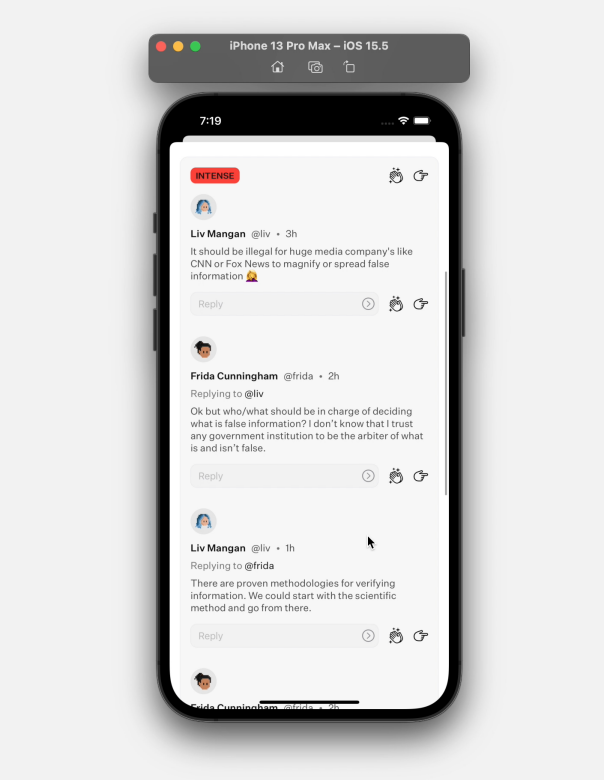

Figure 5: Post tagging

Just as with Likes, meeting certain thresholds of Nudges will trigger a public tag highlighting the primary reason why. This part of the intervention represents the first form of public, social pressure to modify behavior — if the Poster chooses to ignore the private requests for changing their behavior, they may be convinced to do so after a small (or large) dose of public shame.

FIVE: Poster's Response

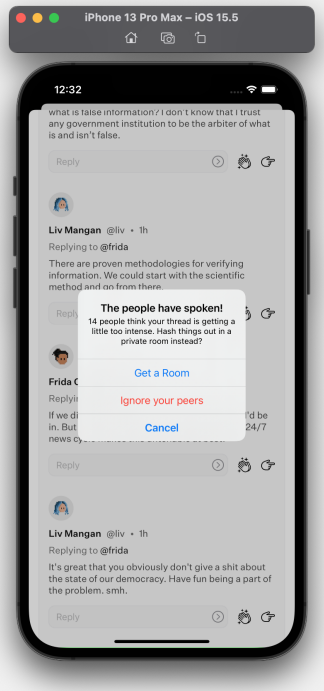

Figure 6A: Nudge delivered

Figure 6B: Entering the room

Sometimes, Posters will choose to ignore the mounting private and public pressure that results from a high volume of Nudges. In this case, a system based intervention becomes necessary to coax a behavioral modification. Within the Convo model, this intervention is called a Convo Alert.

A Convo Alert is deployed when the Poster decides to continue the conversation in question. When this individual attempts to reply, an alert is delivered the instant they press the "Reply" button. The alert contains three courses of future action: "Get a Room", "Explore Elsewhere", and "Ignore".

The inspiration for the Nudge button came directly from Thaler and Suntiens's work in "libertarian paternalism", an idea that was popularized by their book *Nudge.* Thaler and Sunstien's work explores exactly how to modify behavior without doing so heavy handedly. They unpack the idea of a nudge or "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing [...] incentives" (Thaler & Sunstien, 2008).

Within the Convo system, there are no heavy handed suggestions. Rather, people are gently, but firmly, encouraged to change their behavior. Moreover, these encouragements are scaffolded by peer pressure: by showing individuals how many other people want them to behave differently, perhaps they will be more inclined to do so.

Getting a Room

Getting a Room allows individuals to keep a conversation going without facing the pitfalls of public discourse and crowd dynamics. As I've touched on above, research shows that pairing bad incentive structures with the public nature of online discourse regularly leads to these kinds of conversations going off course. Moreover, humans simply behave differently in public and in private. A seminal study on behavior in public settings demonstrated that being subjected to the pressures of a group setting forces people to think less clearly and independently (Asch, 1951). In light of this, moving subpar conversations away from the public eye and into the privacy of a Room presents the opportunity to reset incentives and social pressures and offers interlocutors a lower stakes environment in which they might hash things out.

In terms of the mechanism of a Room, it's a rather simple feature. If the platform has native direct messaging capabilities, the Room will be indistinguishable from these existing features. However, in the case that a platform does not already have these structures in place, Convo includes a messaging space for the offending individuals.

Taking public conversations to a private space may seem antidemocratic in the sense that conversations become back–channeled. Convo responds to these kinds of claims by tagging the original thread with a badge that informs the public the back and forth has moved to a private space. The tag includes the ability for Viewers to "Give Props" to the Posters for moving to a room. "Props" are always public, playing into peer pressure as a behavioral modifier. This kind of feature represents a shared object through which the move to the Room becomes an opportunity for Convo to brand itself and gamify good behavior.

As the conversation in a Room is happening, the participating individuals are intermittently asked how things are going. If both individuals report that things are going well or that things have been resolved, the Room Tag is updated, informing the Viewers that the Room is working. Again, this offers a moment to brand the experience and further gamify the desired behavior.

Exploring

Should an individual choose to “Explore” they are shown aspects of the platform that are not connected with the thread in question. They will, instead, be navigated to a platform’s “Explore” or “Discover” page. If the site does not have these features built in, the Poster will be shown other content deemed relevant or interesting.

Exploring means keeping people online. In a perfect world, Convo would allow and encourage people to step away from their devices altogether. However, to do any good at all, Convo has to first be adopted by clients — it would be hard to sell Convo to any kind of platform if it encouraged the people using the platform to stop doing so.

Ignoring

Ignoring the Convo Alert means that a Poster is right back to where they started. Not everyone will choose to Get a Room or Explore, and that’s perfectly acceptable. Offering soft behavioral boundaries is essential in getting a wide swath of the public to adopt a novel system (Thaler & Sunstien, 2008).

Where does Convo work best?

Throughout the duration of the project, I’ve desired to make something with as far a reach as possible. This led me to design ideas that, upon completion, would exist as platform agnostic design recommendations for any network that utilizes online conversation. The Convo framework has the opportunity to work wherever online conversations are occurring. That means social media “proper” (Facebook, Twitter, etc.) as well as the comment threads that accompany news articles online, for example. Particularly, Convo might work in small conversational environments where asking someone to modify their behavior might make for awkward interactions between friends and peers.

Of course, there are trade offs to attempting to “design for everything”. The positive aspects of such an approach mean that anyone can implement and utilize this kind of tool. The negative side of the coin means that whatever is made cannot be made with as much detail and precision as one might achieve designing for a precise platform or tool.

This is a compromise I was and am still happy ro make. The speed with which technology changes is unparalleled. Not only does designing for the “right now” mean that an idea might be outdated in a matter of weeks, it also means that the design itself is weighted down by the rules and parameters of existing platforms and technologies.

Further, being wisely vague in the approach means that the focus on the ideas themselves can be teased out in more detail. Rather than sweating the small stuff (UI, microinteractions, etc.), the higher level interactions and strategy get a larger share of the designers attention — ultimately, in my opinion, resulting in better ideas.

Building Convo: The Prototype

SwiftUI

Testing and sharing the core ideas of Convo required a high fidelity prototype. As you read on, you'll see that initially, I wanted to build a prototype that stood on its own, separate from any existing mobile frameworks. When the Spring quarter rolled around, I found this idea to be untenable. For one, building every last detail of a mobile app, from scratch, is a monumental task — I decided to pivot.

One of my advisors encouraged me to look into using SwiftUI, a framework used for building Apple based interfaces, as a system with which I could build and test my ideas. Having a modest, but robust, background in web development and coding, I decided to dive in.

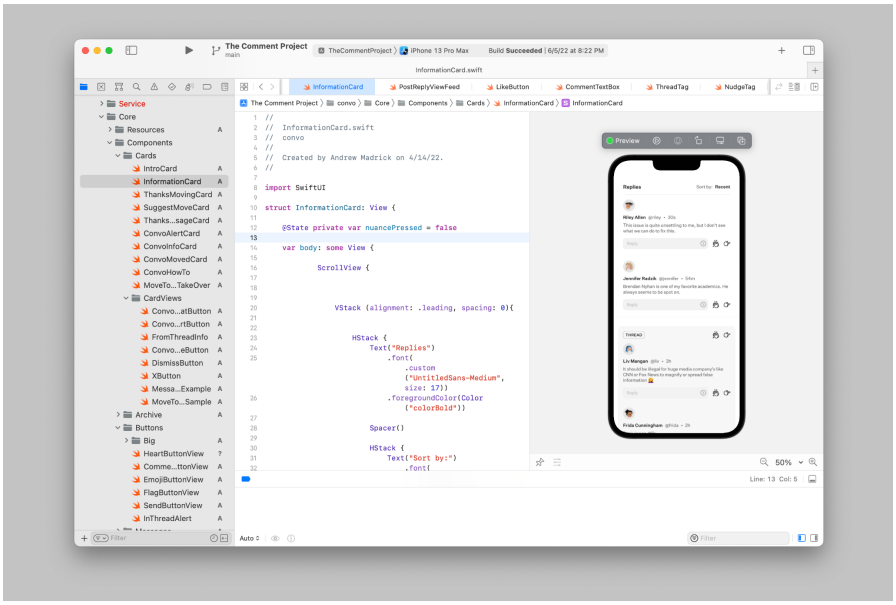

Figure 7: SwiftUI interface

Switching to SwiftUI meant more trade offs. I had to pivot from a completely custom interface to one that was built on and highly resembled existing interfaces that one might

be used to seeing if they have an iPhone or iPad. However, this new constraint offered some exciting opportunities.

Rather than craft every UI element and microinteraction, I was able to utilize those provided by SwiftUI. With a little customization (using Untitled Sans for the system typeface, rather than Apple's own SF Pro), I was able to create a prototype that relied heavily on Apple's native functionalities while maintaining a feel of individuality in the feel of my prototype.

Iterating and Prototyping

How to steal from existing UI patterns?

From the onset of this project, I was set on designing the perfect comment thread. Rather than learn from what exists in the world today, I would design the UI from the ground up. Below, you can see my first pass at this. It's clear I wasn't doing anything new quite yet — consciously or not, it looks like I just redesigned Reddit's UI.

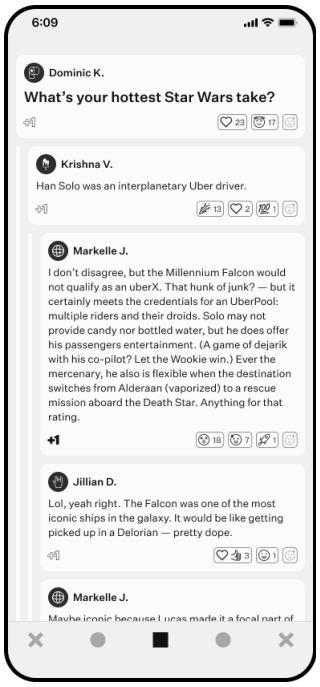

Figure 8: Initial UI

I presented the pictured iteration to an outside professional, a master when it comes to interface design. This person gave me sage (and commonsensical) advice: copy what's

been done before and improve it from there. He elaborated on this directive: “create a pixel–for–pixel redesign of TikTok’s home screen, Twitter or Reddit’s threading, etc.”

This advice set me, visually, on a new course entirely. Rather than toil away homing in on four- or eight-pixel corner radii (the answer is eight), for example, I could enter a cheat code (paste a screenshot into Figma) and have the answer in a matter of seconds.

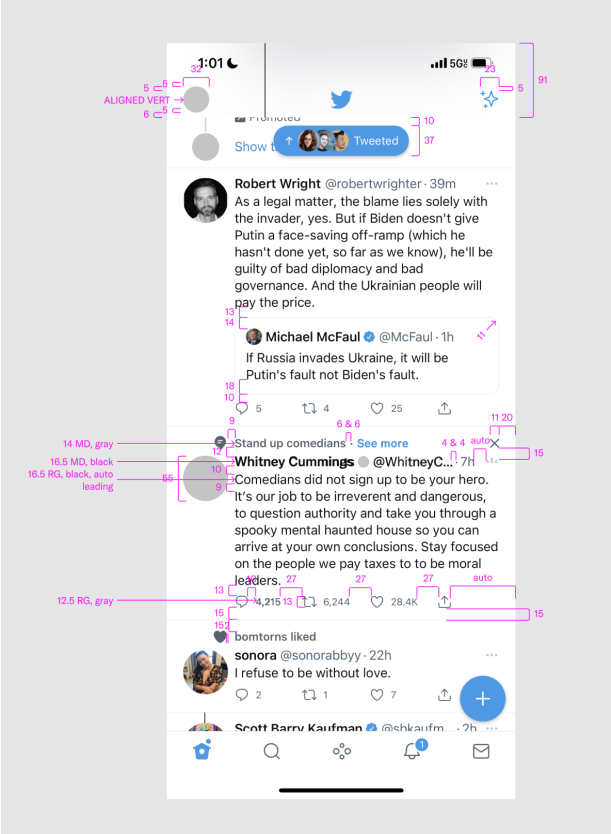

Figure 9: Analysis of Twitter UI

What I found during this process is that interfaces are beautiful objects in and of themselves. In a word, they are more designed than I could have imagined. I found Twitter’s visual design, for example, to be a demonstration of a mastery over color, composition, and typography (shout–out Grilli Type). Further, the macro and micro interaction patterns are beautiful, intuitive, and fun. Taking infinite screenshots and recreating them was a challenge, but one of the best lessons in product design to date.

Black & White to Color

Going back to the beginning of the quarter — one design parameter I set for myself was to make my interface rely on typography for hierarchical delineations. I simply refused to use color. This was a noble challenge, but one that ultimately fell flat (pun intended). The interface mockups that resulted from this exercise were a bland jumble of black, white, and gray. It was harder than necessary to follow and left something to be desired.

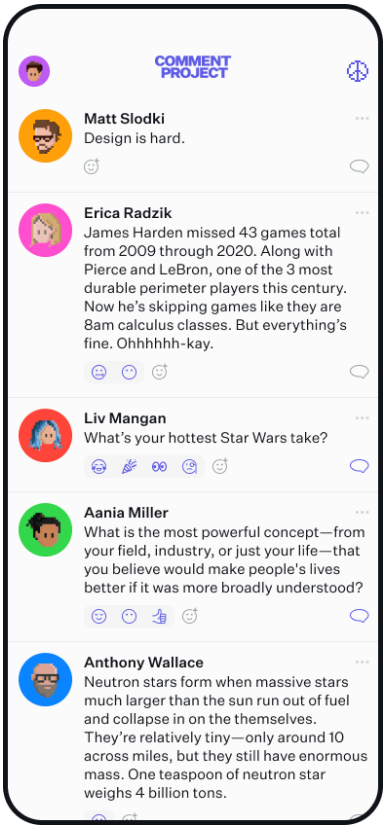

Figure 10: Adding color to the interface

To delineate quickly and easily who was saying what within the interface, I added color to the profile picture backgrounds. Rather than rely on names or icons to understand who was saying what, the instantaneous nature of color allowed for an immediate understanding of what was happening in the thread.

This splash of color started seeping into the overall branding of my project. I introduced a deep blue color for primary logos, labels, buttons and emojis. This added the kind of richness while aligning the visual tone of the project to the motivating ethos of the project.

How to show emojis?

When it comes to representing emojis, I'm admittedly unsatisfied with where I'm currently at. Initially, I wanted to organize emoji responses in word clouds — doing so would allow for a visually striking and novel way of doing something that's been done a million times before.

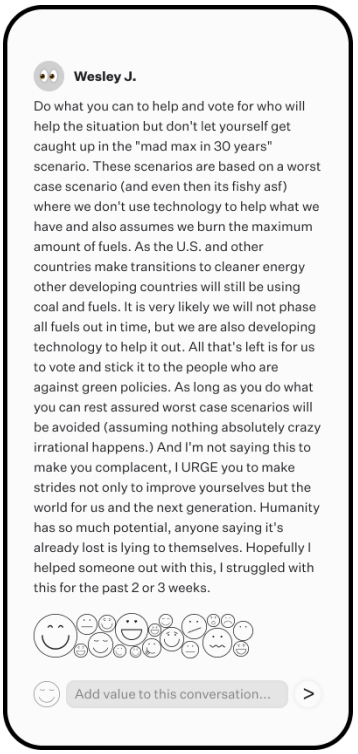

Figure 11: Emoji cloud

This proved to be interesting, but untenable. How would users interact with the cloud? The smallest emojis would be far too small to tap. Not to mention the kind of real estate these things would take up. For the moment, I've been working with a Discord inspired kind of "bucket" in which the emojis can hang out. Does this need to be revised? Yes — but that will be sorted later.

Typography

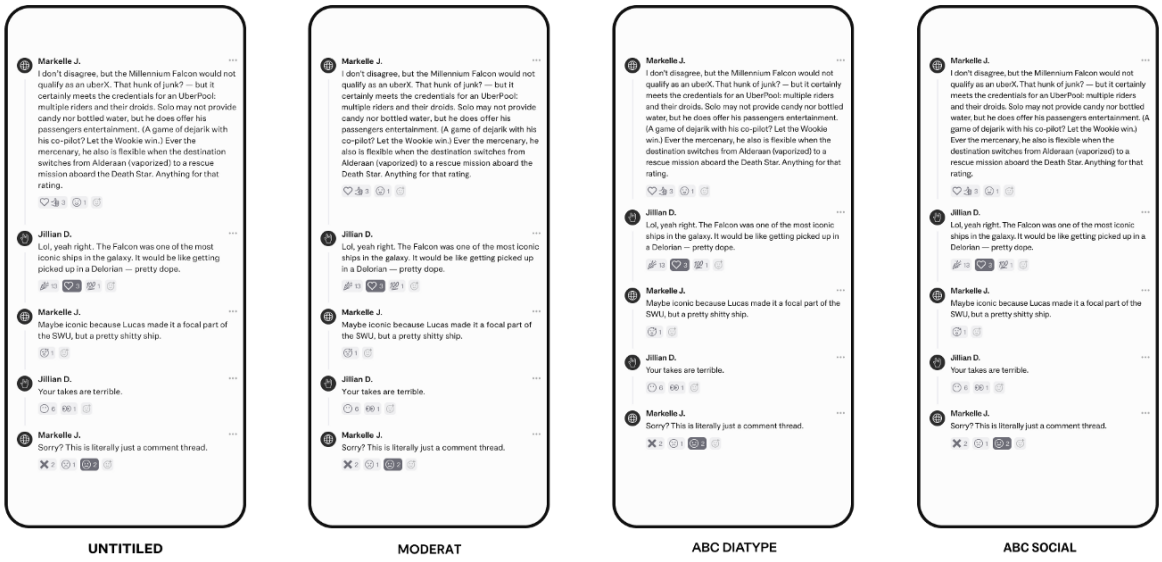

Figure 12: Typography options

From the onset of this project, I knew whatever I was going to make would be set entirely in Untitled Sans. However, I knew I would need to do my due diligence, and at least look at a few other sans–serif fonts. My criteria were short but specific. Whatever typeface I chose must come from an independent foundry, must have a high x–height and it must be visually interesting but subdued. I perused some of my favorites: ABC Diatype (ABC Dinamo), Founders Grotesk (Klim) Area (Blazetype), and Lars (Bold Decisions). I love these typefaces, but nothing spoke to me the way Chris Sowersby's reluctant masterpiece does. Untitled is the only typeface featured in my project.

Practitioner Input

Over the course of the year, I met with a handful of highly skilled interface and app designers. These meetings were off the record, so I will refrain from mentioning them by name.

One advisor gave me some simple but revelatory advice: "whatever you make, don't brand the thing." Up to that point, I had been positioning my interventions as standalone features, things to be added to an existing platform (see below for how this shifted). The genius in not branding these interventions is the fact that by leaning on existing patterns and features, whatever I make can slide effortlessly into existing platforms. I won't have to

argue that Twitter, for example, should adopt the Convo intervention and everything that comes with it. Rather, by using Twitter's existing Direct Messaging system, they can get all the added value of Convo without the extra calories.

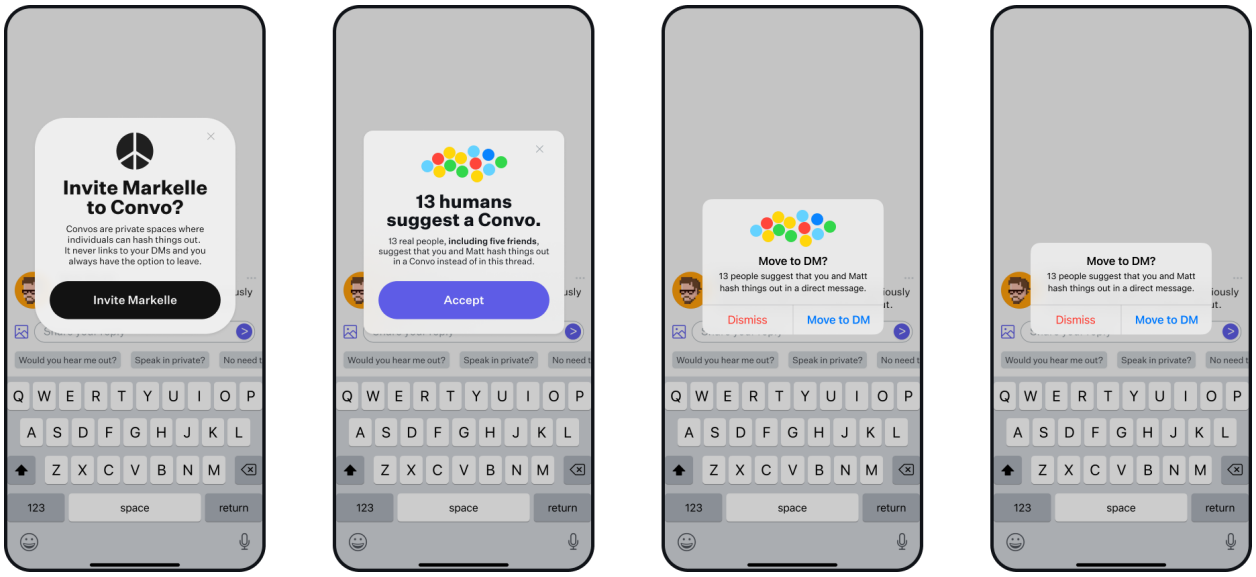

Figure 13: Alert options

I was also instructed to think about when to deploy the Convo notifications. Rather than alert the individual the exact moment Convo recommendations are made by peers, think about when the ideal moment of intervention might be. After some discussion, we reasoned that alerting might be best deployed right after someone "smashes" the send button.

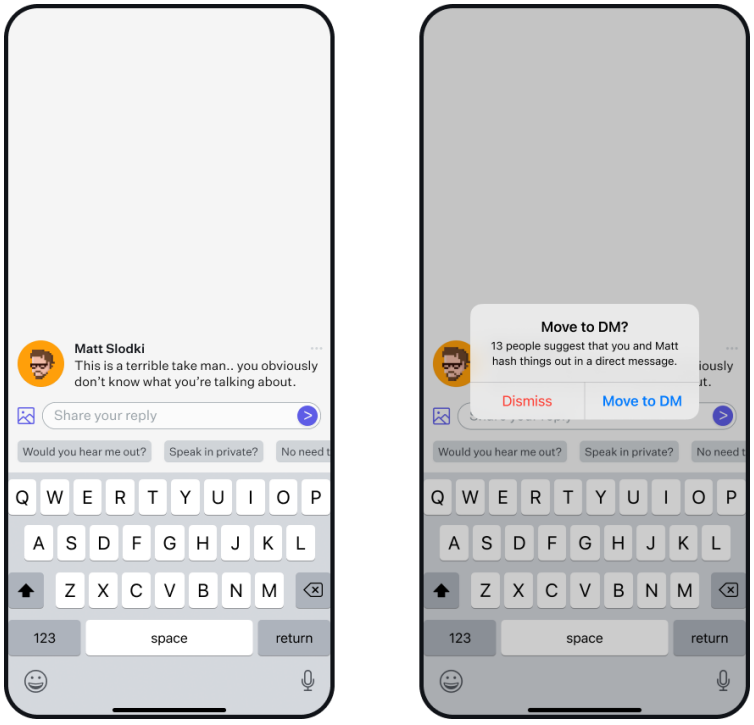

Figure 14: Standarized alert

This advisor went on, encouraging me to lean into the research I've conducted on nudging and social pressure when it comes to behavior modification. When alerting the subjects of a Convo request, perhaps they should see how many people want them to "get a room". What if their credentials were shared with the subject themselves? This is a thread I have begun to pick up, but one that needs further exploration.

Another outside advisor additionally pressed me to lean into the craft of my project. No matter how strong the concept might be, if this thing isn't made with an obsession over detail, it will fall flat. He went on to press the idea that when this is being tested, if the feel or look of the prototype are distracting, the user will be focused on that. Users need cognitive clarity to give insight into the conceptual underpinnings of the project.

With that, he instructed me to recreate the UI of existing platforms. As I mentioned above, this exercise was particularly helpful for my development as a digital designer. Platforms like Twitter (whose UI I'm particularly fond of) have already nailed the execution of visual organization and design. By recreating various UI's, I was reminded that as designers, we stand on the shoulders of giants. Reinventing the wheel just isn't always necessary.

Ideating and Downselection

When I began thinking about what kind of interventions to create, my methodology for ideating varied — some were random ideas that had lived in the back of my mind for months, while others were manufactured through classic human–centered design methodologies. Through these processes, I generated a number of ideas, four of which I thought had the most promise.

Figure 15: Nth degree responses

One: Nth degree responses.

Rather than responding to whole comments or posts, what if users could react to the minutiae of a post? The driving force of this idea would be the clarity and intrigue that comes when precision is brought to conversation.

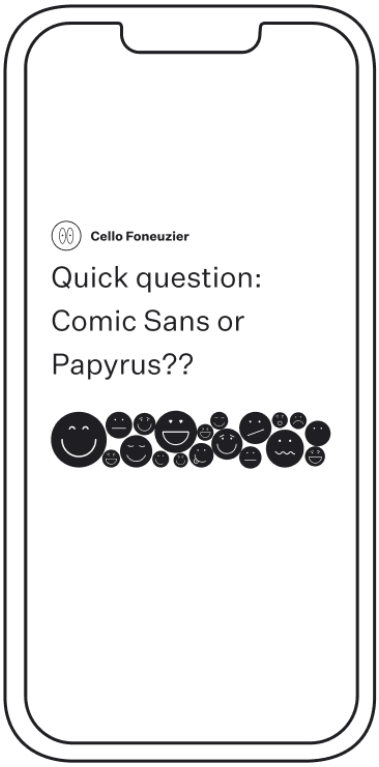

Figure 16: Lots and lots of emojis

Two: Lots and lots of emojis.

Emojis seem underutilized online. What if they became focal elements instead? By shifting a social platform to shift the way in which emojis are used — towards something closer to the model implemented by Slack, for example — more nuance could be added to a fundamental feature of social media.

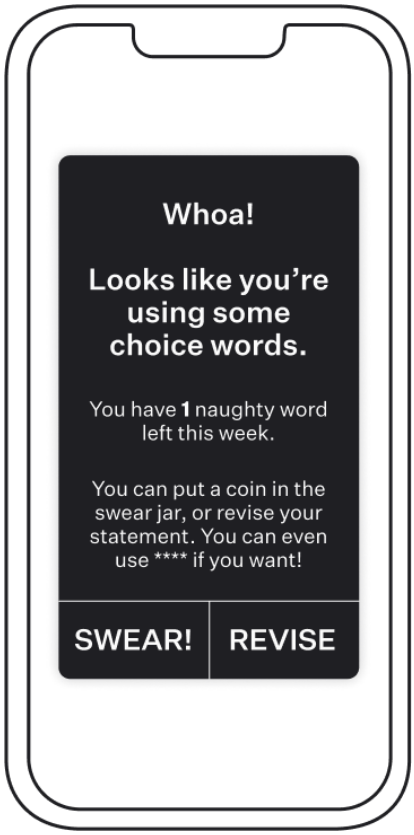

Figure 17: The fun(?) swear ja

Three: The fun(?) swear jar.

Crude language is inevitable and can heat things up quickly. What if language became gamified, making courtesy a desired outcome? Rather than being overly paternalistic, this idea intentionally allows individuals to be subversive and creative. Maybe they dodge the rules by playing tricks on the system, which is kind of the point. At first glance, threads would appear more civil, which is incredibly important in its own right.

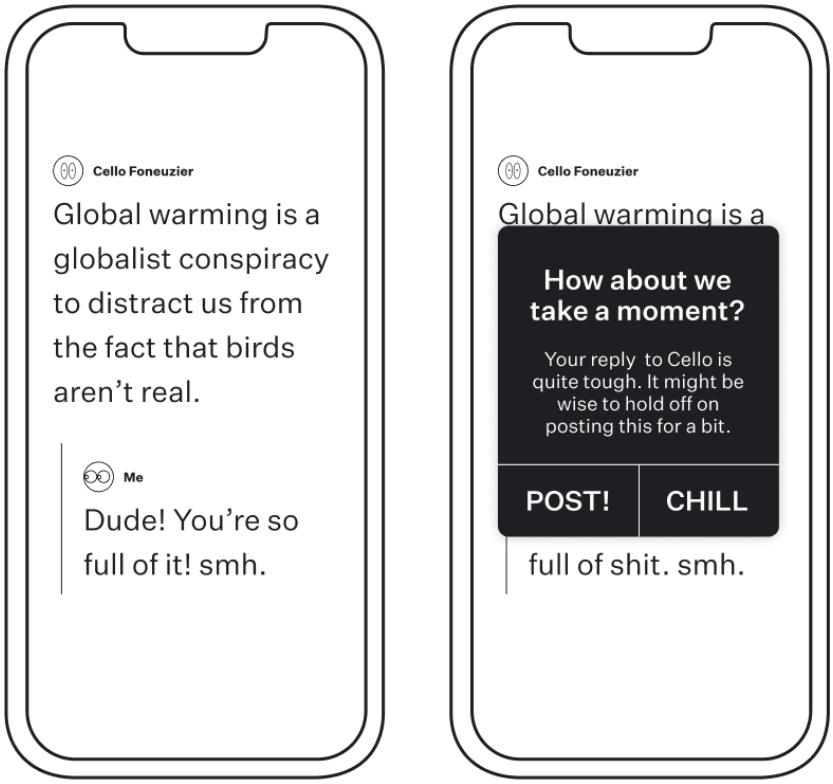

Figure 18: Put baby in the corner

Four: Put the baby in the corner.

Almost everyone has posted something they later regretted. This idea takes the immediacy out of the posting process. Individuals are able to take some time to think before the post actually goes live, decreasing the chance of hostile interactions.

Refinement

After discussing the ideas presented above with my committee as well as a number of outside advisors, I decided to move forward with three alternatives: "Convo", "Foreground", and a rethink of the way we use emojis to communicate in comment threads (which, unfortunately, never got a name).

Convo cleans up comment threads.

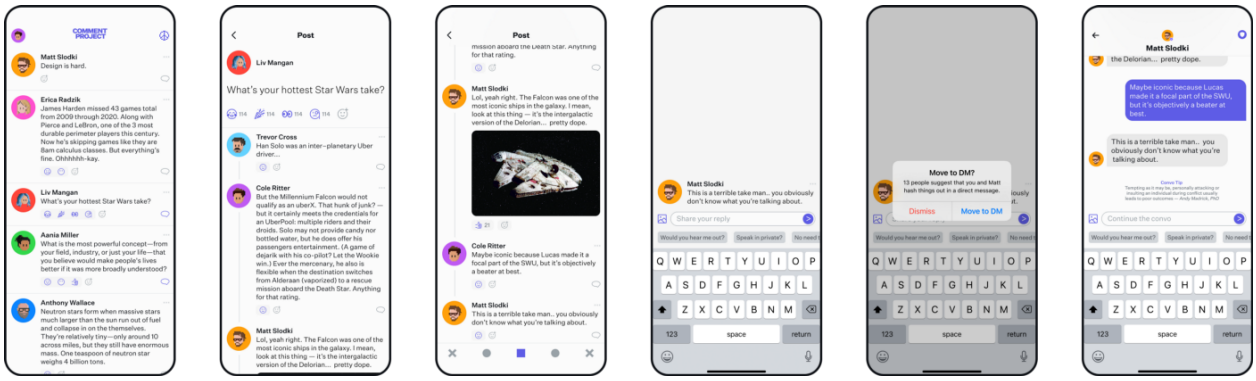

Figure 19: The Convo user flow

Convo is an outgrowth of the “Put Baby in the Corner” idea mentioned above. Situationally, this idea centers on two or more individuals “clogging up” a comment thread. This could be through a hostile back and forth, or simply a meaningless, boring, and unproductive dialogue. The heart of this feature lies in the novel ability of other users (peers) to recommend that the interlocutors move their conversation away from the thread to a direct message. Once recommendations pile up to a certain level, the interlocutors are alerted that their conversational peers want them to “get a room”. If accepted, the dialogue in questions will be moved from the thread and into a direct message, the chrome of which is populated with helpful conversational and confrontational tips and prompts.

Foreground finds meaning.

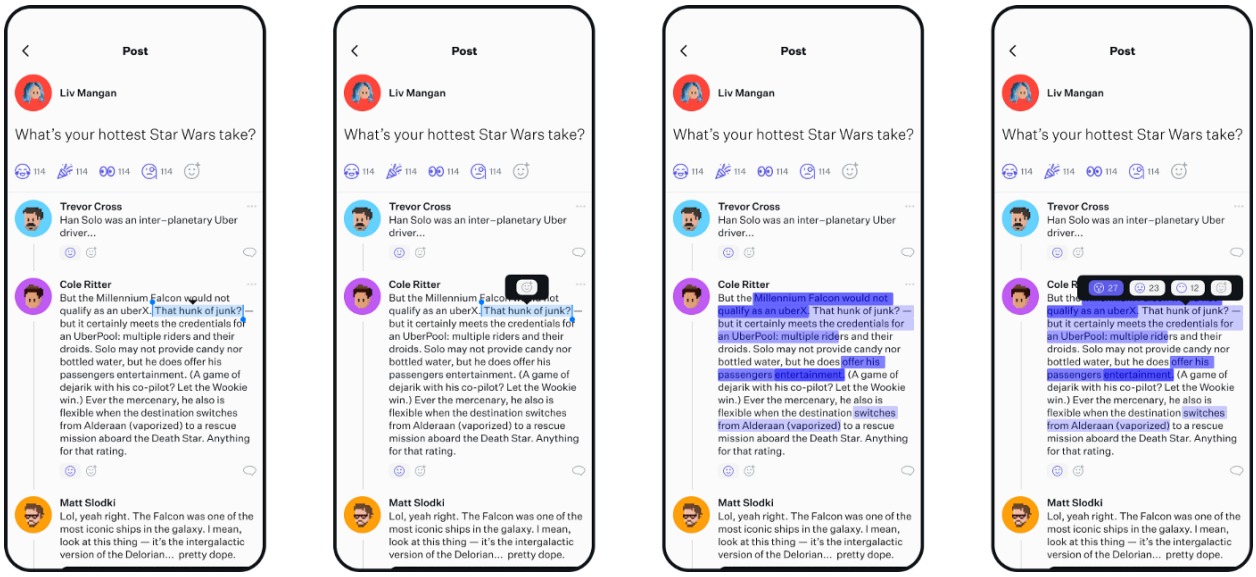

Figure 20: Highlight user flow

Foreground is the progression of the “Nth Degree Responses" intervention from above. This feature would allow users to respond to the minutiae of posts online. The core of the idea is to combat a wholesale reaction to a longer poster. What if someone says something I simultaneously agree and disagree with? Responding with a “Like”, doesn’t necessarily do justice to my visceral and complex reaction. Foreground gives the user not only the ability to Like certain aspects of posts, but they can also respond with the emoji of their choosing, adding a level of richness that simply is not there otherwise.

After a certain threshold of responses has been crossed, the data collected will start to bubble up to the surface of the interface. Highly responded sections will begin to evidence themselves, allowing individuals to find and interact with the most valuable parts of the post.

Emojis need some extra love.

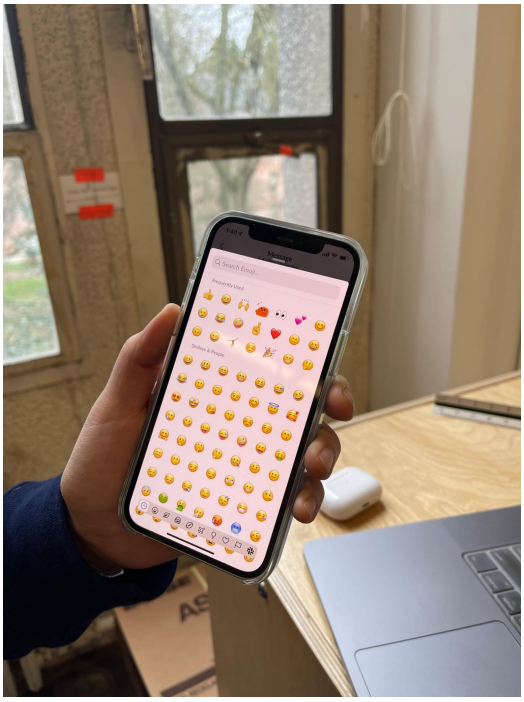

Figure 21: Image from emoji interview

I always knew that I wanted to do something about the way in which people use emojis. The motivating idea behind these interventions is that, from anecdotal experience, the way we currently use emojis seems to lack meaning and nuance. After thinking about this idea in more detail, I became motivated to create a set of emojis that allows users to show contradictory emotions and thoughts in a single symbol.

To better understand if this was a viable project, I interviewed a number of individuals about their experience using emojis. People of different ages coming from different backgrounds all gave me interesting, but fairly unusable insights into designing emojis. More than anything, it became clear that emojis are cultural artifacts that get their meaning from how they might be used by the wider public. No one who I interviewed indicated that they wished there would be more, or better, emojis. The only usable idea I received was to make emojis “cooler” — make them look like they were made for “cool adults to use”. Not necessarily the insight I had been expecting.

Conclusion

Building Convo was one of the more challenging and life–enriching projects I've worked on. Combining my interests and passions while working with and speaking to incredibly smart and knowledgeable experts was an experience that has made me a stronger designer and a more informed designer.

As far as where Convo will go moving forward, I hope to test it if possible. It would be fascinating to understand if people would actually use and benefit from the framework. Getting the idea of Convo into the psyche of the wider public in general and social media designers and decision makers in particular might have a positive impact on the way in which we talk to each other online.

Bibliography

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In Groups, leadership and men; research in human relations (pp. 177–190). Carnegie Press.

Asker, D., & Dinas, E. (2017). Do Online Media Polarize? Evidence from the Comments' Section (SSRN Scholarly Paper ID 3018223). Social Science Research Network. https://doi.org/10.2139/ssrn.3018223

Baughan, A., Petelka, J., Yoo, C. J., Lo, J., Wang, S., Paramasivam, A., Zhou, A., & Hiniker, A. (2021). Someone Is Wrong on the Internet: Having Hard Conversations in Online Spaces. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW1), 1–22. https://doi.org/10.1145/3449230

Boxell, L., Gentzkow, M., & Shapiro, J. (2020). Cross-Country Trends in Affective Polarization (No. w26669; p. w26669). National Bureau of Economic Research. https://doi.org/10.3386/w26669

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. Proceedings of the National Academy of Sciences, 114(28), 7313–7318. https://doi.org/10.1073/pnas.1618923114

Budak, C., Garrett, R. K., Resnick, P., & Kamin, J. (2017). Threading is Sticky: How Threaded Conversations Promote Comment System User Retention. Proceedings of the ACM on Human-Computer Interaction, 1(CSCW), 1–20. https://doi.org/10.1145/3134662

Cheng, J., Bernstein, M., Danescu-Niculescu-Mizil, C., & Leskovec, J. (2017). Anyone Can Become a Troll: Causes of Trolling Behavior in Online Discussions. CSCW : Proceedings of the Conference on Computer-Supported Cooperative Work. Conference on Computer-Supported Cooperative Work, 2017, 1217–1230. https://doi.org/10.1145/2998181.2998213

Digital civility at lowest level in 4 years, new Microsoft research shows. (2020, February 11). Microsoft On the Issues. https://blogs.microsoft.com/on-the-issues/2020/02/10/digital-civility-lowest/

Facebook ad revenue 2009-2020. (n.d.). Statista. Retrieved June 9, 2022, from https://www.statista.com/statistics/271258/facebooks-advertising-revenue-worldwide/

Facebook MAU worldwide 2022. (n.d.). Statista. Retrieved June 8, 2022, from https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/

Gerrard, Y. (2018). Beyond the hashtag: Circumventing content moderation on social media. New Media & Society, 20(12), 4492–4511. https://doi.org/10.1177/1461444818776611

Grygiel, J., & Brown, N. (2019). Are social media companies motivated to be good corporate citizens? Examination of the connection between corporate social responsibility and social media safety. Telecommunications Policy, 43(5), 445–460. https://doi.org/10.1016/j.telpol.2018.12.003

Haidt, J. (2022, April 11). Why the Past 10 Years of American Life Have Been Uniquely Stupid. The Atlantic. https://www.theatlantic.com/magazine/archive/2022/05/social-media-democracy-trust-babel/629369/

Hare, C., & Poole, K. T. (2014). The Polarization of Contemporary American Politics. Polity, 46(3), 411–429. https://doi.org/10.1057/pol.2014.10

How to Talk to People Who Disagree with You Politically—Center for Media Engagement. (n.d.). Retrieved October 19, 2021, from https://mediaengagement.org/research/divided-communities/

Humprecht, E., Hellmueller, L., & Lischka, J. A. (2020). Hostile Emotions in News Comments: A Cross-National Analysis of Facebook Discussions. Social Media + Society, 6(1), 2056305120912481. https://doi.org/10.1177/2056305120912481

Kantrowitz, A. (n.d.). Man Who Built The Retweet: “We Handed A Loaded Weapon To 4-Year-Olds.” BuzzFeed News. Retrieved June 8, 2022, from https://www.buzzfeednews.com/article/alexkantrowitz/how-the-retweet-ruined-the-internet

Mierlo, T. van. (2014). The 1% Rule in Four Digital Health Social Networks: An Observational Study. Journal of Medical Internet Research, 16(2), e2966. https://doi.org/10.2196/jmir.2966

Munn, L. (2020). Angry by design: Toxic communication and technical architectures. Humanities and Social Sciences Communications, 7(1), 1–11. https://doi.org/10.1057/s41599-020-00550-7

OHCHR | Report: Online hate increasing against minorities, says expert. (n.d.). Retrieved December 17, 2021, from https://www.ohchr.org/EN/NewsEvents/Pages/sr-minorities-report.aspx

Rainie, L., Kiesler, S., Kang, R., & Madden, M. (2013, September 5). Part 1: The Quest for Anonymity Online. Pew Research Center: Internet, Science & Tech. https://www.pewresearch.org/internet/2013/09/05/part-1-the-quest-for-anonymity-online/

Rathje, S., Van Bavel, J. J., & van der Linden, S. (2021). Out-group animosity drives engagement on social media. Proceedings of the National Academy of Sciences, 118(26), e2024292118. https://doi.org/10.1073/pnas.2024292118

Roberts, S. T. (2017). Content moderation. https://escholarship.org/uc/item/7371c1hf

Schwartz, A. (n.d.). Father of virtual reality: Facebook and Google are dangerous “behavior-modification empires” resulting from a tragic mistake. Business Insider. Retrieved June 8, 2022, from https://www.businessinsider.com/facebook-google-mark-zuckerberg-behavior-modification-empires-2018-4

Seering, J., Fang, T., Damasco, L., Chen, M. “Cherie,” Sun, L., & Kaufman, G. (2019). Designing User Interface Elements to Improve the Quality and Civility of Discourse in Online Commenting Behaviors. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1–14). Association for Computing Machinery. https://doi.org/10.1145/3290605.3300836

Stroud, N. J., & Scacco, J. M. (n.d.). Restructuring Comment Sections. Engaging News Project, 12.

Suler, J. (2004). The Online Disinhibition Effect. CyberPsychology & Behavior, 7(3), 321–326. https://doi.org/10.1089/1094931041291295

Thaler, R., & Sunstien, C. (2008). Nudge: Improving Decisions About Health, Wealth, and Happiness. Yale University Press.

The cooperative human. (2018). Nature Human Behaviour, 2(7), 427–428. https://doi.org/10.1038/s41562-018-0389-1

Toxic Online Comments: How They Happen, and How To Stop Them. (n.d.). OpenWeb. Retrieved November 17, 2021, from https://www.openweb.com/blog/toxic-online-comments-how-they-happen-and-how-to-stop-them

Vogels, E. a. (2021, January 13). The State of Online Harassment. Pew Research Center: Internet, Science & Tech. https://www.pewresearch.org/internet/2021/01/13/the-state-of-online-harassment/

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. Science, 359(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Wang, Y., Norcie, G., Komanduri, S., Acquisti, A., Leon, P. G., & Cranor, L. F. (2011). "I regretted the minute I pressed share": A qualitative study of regrets on Facebook. Proceedings of the Seventh Symposium on Usable Privacy and Security - SOUPS '11, 1. https://doi.org/10.1145/2078827.2078841

Xu, C., Ryan, S., Prybutok, V., & Wen, C. (2012). It is not for fun: An examination of social network site usage. Information & Management, 49(5), 210–217. https://doi.org/10.1016/j.im.2012.05.001