Framing the Ethical Boundaries of Humor

Abstract: Humor is unlike other forms of communication because its content is not meant literally. Like acts of play, humor is not intended to be taken at face value. As a consequence, the assumptions and rules that govern normal conversation do not apply. Humor therefore depends upon both the speaker and the audience fully understanding that what was communicated should be treated in this unique way. The play frame refers to this shared understanding about the nature of the communication. Analyzing whether a communication falls within the play frame may help us better understand not only whether the communication can be deemed non-serious or serious, but also whether it can be treated uniquely as merely the speaker’s attempt at play or as a typical instance of literal speech for which the speaker can be deemed ethically responsible.

Keywords: humor, play frame, ethics, moral responsibility

In her introduction to a collection of the works of New Yorker humorist S.J. Perelman, Dorothy Parker wrote: “I had thought, on starting this composition, that I should define what humor means to me. However, every time I tried to, I had to go and lie down with a cold wet cloth on my head” (Parker 1958).

Indeed, trying to figure out what humor means can be challenging (and sometimes even painful) because it is such a unique type of expression. While a joke may convey informational content, it is fundamentally unlike other forms of communication because that content is not intended to be taken literally—like acts of play, humor is not meant to be taken at face value. As a consequence, the assumptions and rules that govern normal conversation simply do not apply, and it stands outside of the way we ordinarily navigate and understand the world. Humor therefore depends on both the speaker and the audience fully understanding that what was said should be treated in this unique way.

The play frame refers to this shared understanding about the nature of the communication. It is a reflection of the overall context of what was actually conveyed as well as the dynamic interaction between the speaker and the audience. Analyzing whether a communication falls within the play frame may therefore help us better understand not only whether the communication can be deemed non-serious or serious, but also whether it can be treated uniquely as merely the
speaker’s attempt at play or as a typical instance of literal speech for which the speaker can be deemed responsible. As such, the play frame analysis provides another way to sketch the ethical boundaries between when a person may be held accountable for telling a joke that, if taken at face value, would be deemed an immoral act, and when their behavior may be deemed appropriate (or at least not inappropriate). In fact, even broader debates about the ethical limits of humor may ultimately be better understood as debates about the existence and extent of the play frame.

1 The Concept of Framing

Generally speaking, “framing” refers to the way that people categorize the world around them. It functions like a scaffolding of logical understanding—a perspective that brings with it a set of already formed logical relations, assumptions, and expectations that can be applied like a heuristic to a current situation. It helps people categorize and organize their history of experience in order to construct new meaning and save mental effort. The concept was first applied to humor research by Gregory Bateson, and research on framing has since developed into two approaches: cognitive (which considers knowledge frames), and interactional (which considers interactive frames) (Yus 2016, 81–2; Dynel 2011, 219–20; Bateson 1953).

The concept of knowledge frames is derived from cognitive science, where frames or schemas have been used to describe the way memories and experience are used as frameworks against which new information is compared and synthesized (Wu 2003, 6–10). Knowledge frames are cognitive phenomena that serve as “data structures” that embody a routine or stereotyped situation, bringing with it a system of logical rules that help with processing the new experience. This concept is also widely used in fields like artificial intelligence, cognitive psychology, and linguistic semantics (Dynel 2011, 219; Minsky 1974).

Interactive frames, on the other hand, focus on the dynamic nature of framing as “an interactive event oriented towards a particular goal and centered on rules and expectations but negotiated and co-constructed by interacting parties” (Dynel 2011, 219). This recognizes that because of the dynamic nature of communication, frames can change constantly, even multiple times within the same conversation. It is used in fields such as anthropology, sociology, linguistic anthropology, and sociolinguistics (Dynel 2011, 219). Thus, unlike cognitive frames which focus on the “mental construction of the situation,” the interactive approach recognizes how meaning is created within and through the process of interactive communication—a process that is guided by a shared understanding
As Francisco Yus helpfully summarized in the context of jokes, “cognitive frames deal with what is going on in the joke and interactive frames deal with what is going on in telling the joke” (Yus 2016, 82).

Most humor research has generally focused, explicitly or implicitly, on the interactive nature of framing (Dynel 2011, 220). This is likely due to the dynamic nature of joke telling, the most commonly thought of form of humor, which often directly depends on the ongoing interactive interpretation of the frame between the speaker and audience. At the same time, however, the shared understanding of the frame is still very much dependent upon the overall context as well as the participants’ past experience, which unavoidably implicates cognitive framing as well.

2 The Play Frame and Humor

Humor employs a very particular type of framing—the play frame. This embodies the understanding that the content that is being communicated is actually not intended at face value, and therefore is not meant to be interpreted seriously or literally. As William Fry discussed in the early 1960s (although the idea traces back to Gregory Bateson in the 1950s), this type of behavior occurs in animal play, where playful activity is usually very similar to other forms of behavior such as fighting but takes place under a shared understanding that it is not intended in the same way. “Somewhere in the process of play is the metacommunication that this particular episode is play, not a fight, etc. ... Play, then, is behavior which depends on the mutual recognition (through metacommunication—internal and external) that behavior (play) does not mean the same thing as does that behavior (fighting, etc.) which play represents” (Fry 2010, 125–6).

In other words, in animal play an action like a dog nipping at another is accompanied by some type of signal that the bite is not meant as an attempt at genuine aggression. Anthropological linguist John J. Gumperz referred to these metacommunicative signals as a “contextualization cues” (Yus 2016, 82; Gumperz 1992). When such a cue is properly understood, the recipient of the bite knows to respond differently than she would had she been bitten by a dog who was being truly aggressive. Thus, the metacommunication of the play frame means that the typical rules that govern behavior are suspended and other rules apply, and the participants must therefore alter the way they process information and behave.

The metacommunication that establishes the play frame can be made in a number of ways. In the context of the play nip between dogs, it could be accom-
panied by a physical gesture (a “play bow” or tail-wagging), verbally (the high-pitched tone of a playful bark), or it can be communicated through the overall context (encountering another dog at a dog park that is already associated with non-aggressive playing). It could even be embodied in the communication itself (an extremely gentle bite), or in some feature of the actor (a Chihuahua puppy who nips a Clydesdale). Through her actions and the metacommunicative cues, the initial communicator also defines the relative size of the play frame, which governs the range of behavior that would be appropriate within it. (Responding to a barely perceptible nibble with a vicious chomp would be interpreted as outside of the established frame of play and could be treated as a true act of aggression—and subject the inappropriate aggressor to the non-playful consequence of a vicious chomp in return.)

Just as importantly, because the play frame embodies a shared understanding, the recipient of the communication must also accept the play frame in some way. In the context of play, this can be done by responding in kind or with a non-threatening bark or gesture, but may also be through acquiescence. Conversely, the recipient can also indicate her rejection and refusal to enter the play frame. This can be because she is uninterested in taking part in the activity, but may also be if she thinks the communication itself was outside of the conceivable scope of the play frame (such as if the initial invitation to play was a chomp rather than a nibble).

As it applies to humor, the play frame has been described in various ways—as the humorous frame, the non-serious frame, the jocular frame, the non-serious key, the humorous key, and the humorous mode (Dynel 2018, 391). It embodies the same foundational notion as in animal play that there is a shared understanding that what is taking place is not serious. When this understanding is established, the communication “coincides with activity that necessitates at least partial detachment from serious and goal-oriented behavior”—and this ultimately facilitates “a state of mind which is conducive to amusement in response to humorous stimuli” (Dynel 2018, 391). The concept of framing therefore helps describe this “process of shifting from one type of interactional logic to another” (Kuipers 2008, 378; Goffman 1974).

Moreover, “play” in the context of humor is not limited to the most basic sense of playing (as in playing a game with someone) but it can also constitute a creative endeavor. For example, Steven Gimbel likens the play frame to an aesthetic stage, where what transpires is more akin to artistic expression than serious conversation (Gimbel 2017, 39). This is rooted in his definition of play as using something in a different way than it was intended—like playing with one’s food (or more specifically, like pretending a banana is a phone) (Gimbel 2017, 41). Needless to say, in such examples the creativity—and hence the
humor—depends upon establishing a conceptual space within which the underlying communications are neither meant nor received literally.

2.1 Inside the Play Frame

It is one thing to say that a communication is not intended to be taken seriously, but practically speaking, what does this really mean? How does that change the dynamic of the communication and how meaning is constructed?

Being within the play frame means that the intentions behind the communication are different from what we would typically expect, and because of that, the typical rules we use to interpret communications are no longer applicable. These rules are often described in H.P. Grice’s familiar Cooperative Principle and his four conversational maxims. In normal conversation it is assumed that speakers abide by these maxims, and that we can all rely on this collective behavior in order to understand and be understood by others (which is the essence of Grice’s cooperative principle). Humor, however, seems to violate every one of these maxims. It has even been contended that all jokes violate at least one of these maxims (Krikmann 2006, 45; Attardo 1993, 541).¹ For example:

1. **Quantity:** Be as informative as necessary, no more and no less.

   Example of violation: “Excuse me, do you know what time it is?”—“Yes” (Attardo 1993, 541).

2. **Quality:** Say what is true; do not say what is known to be false or what is not supported by evidence.

   Example of violation: “Why did the Vice President fly to Panama?”—“Because the fighting is over” (Attardo 1993, 542).

3. **Relation:** Say what is relevant.

   Example of violation: “How many surrealists does it take to screw in a lightbulb?”—“Fish!” (Attardo 1993, 541).

¹ There are, of course, those who disagree with this and contend that the maxims are not violated by jokes. For example, Haruhiko Yamaguchi drew on Dan Sperber and Deirdre Wilson’s “mention theory” (which was offered to explain irony), asserting that humorous utterances were not direct propositions but were essentially echoes of references or quotations that the speaker wants to ridicule. In other words, the speaker was not responsible for the violations of the maxims made in the joking statements because they had delegated the responsibility for them “inside the joke” to aspects of the narrative. However, as Avro Krikmann noted, while Yamaguchi’s theory may apply to violations that occur in the direct speech of joke characters, it still fails to account for the maxims that are violated by “narrative reports” and indirect speech (Krikmann 2006, 45–6).
4. **Manner**: Be clear, brief and orderly; avoid ambiguity and obscurity.


In normal (non-humor) speech, violating these maxims has consequences. For example, according to the maxim of quality, we should not knowingly make false statements. This means that people assume a speaker believes what they say and can back up their factual assertions with some form of evidence. They can therefore be called to account if they knowingly lie or make unsupported statements—and repeated violations can earn a speaker a poor reputation and lose them the trust of others in their community. (This also shows how these rules have implications beyond just how to construct meaning, but also in determining when a normal communication may be potentially unethical.)

Humor, however, is a unique form of communication because it does not convey meaning in the same way and cannot be evaluated in accordance with the traditional conversational maxims. Just as the rules of chess do not apply to Calvinball, Grice’s principles do not apply to humor. In other words, when a communication falls inside the play frame, the assumptions on which normal day-to-day conversation depend are out the window. This also means that communications recognized to be within the play frame will not bear the same consequences for failing to abide by these rules and assumptions—which is perhaps the most significant implication of the play frame.

An example of this can be seen in an insult reportedly leveled by William Faulkner against Ernest Hemingway:

He has never been known to use a word that might send a reader to the dictionary.

To the extent this could be interpreted literally, it asserts a fact that Faulkner obviously could never know to be true which would violate the conversational maxim of quality. However, because it is such an obvious exaggeration—which in this case provides the metacommunicative cue that it is made within the play frame—the typical expectation that the statement must abide by that maxim is not applicable or appropriate. (In fact, it is even generally assumed that the requirement of truthfulness is suspended within the play frame [Dynel 2018, 398]). Had Faulkner said something that could only be interpreted

---

2 For more on Calvinball, see: https://calvinandhobbes.fandom.com/wiki/Calvinball
3 Hemingway allegedly responded, “Does he really think big emotions come from big words?” The blog Quote Investigator cites sources that support the substance of this exchange, although in slightly different wording (O’Toole 2016).
as a serious assertion (such as “Hemingway failed out of college”) it would be viewed differently, and the maxim would apply. Faulkner could then be asked to give the evidence he had for such a statement—or face the consequence of being labeled a liar. Because his jibe about the dictionary is clearly a humorous exaggeration, it is not held to the maxim, and it would not be appropriate to ask him to back it up or subject him to negative consequences for violating the maxim (no matter how many people actually consulted a dictionary while reading Hemingway).

2.2 Establishing the Play Frame

Given that humor is such a unique form of communication, it becomes especially important to clearly establish that it is humor. So, what does this mean in practical terms? How do these metacommunicative cues work? In short, how exactly is the play frame established?

The most obvious way to establish the play frame is to do so explicitly in communication. This can be done by using the familiar kind of setup that precedes jokes, such as: “Have you heard the one about...” or “Stop me if you’ve heard this one before...” The content of the cue establishes an intention to create the play frame, which the audience will understand by filtering it through their experience of how that cue typically works. This, along with non-verbal physical gestures, operate as distinctly cognitive framing cues because they are not dependent on the dynamic nature of communication. (Indeed, no matter how seriously a statement is intended, prefacing it with the cue, “So a funny thing happened to me on the way to the office...” will lead the audience to believe that the speaker does not sincerely mean it. The cue alone communicates the framing.)

The play frame can also be established much more subtly through context, which can be done in a number of ways. One way is through the context of the communication itself, where its very nature implies that it does not involve purely serious content. This happens most notably with humor that relies on incongruity. For example, no matter where the following statements are made or how they are delivered, they would not be taken at face value:

I like a woman with a head on her shoulders. I hate necks. (Steve Martin)

Sometimes my grandma guilt trips me about doing drugs. That’s when I know I’m super high, cause she’s dead. (Kate Quigley)

The larger and more dynamic context in which humorous statements are made—the interactive frame—can also serve as the metacommunicative cue. Most obvi-
ously, this could involve the physical setting, such as a comedy club where people go specifically to listen to jokes. It could also have to do with the behavior of the communicator herself: She may have already told a number of jokes or may be speaking authoritatively about something that is clearly unknowable. It may even be separate and apart from anything she says or does and dependent upon her very identity. Take the following line:

> Online dating is tough. Every time I meet someone new, they end up in jail.

Read at face value, this statement is not humorous at all, and seems like a serious assertion (albeit one that omits some important details). It is only when learning that the speaker, Saffron Herndon, was eleven years old does it becomes clear that it is meant as a joke. Her age serves as the cue that it was made within the play frame.

As stated above, however, the play frame is a shared understanding, so establishing it is not solely a matter of the communicator’s intention—it must also be accepted by the audience. This can also be done in many ways. A typical reaction to humor is laughter, which generally indicates it was accepted as intended and succeeded at its intended purpose, although a “failed” joke that receives no laughter may still remain within the play frame.⁴ The context can also manifest consent to the frame, such as voluntarily going to a comedy club to listen to comedians.

Additionally, not all play frames are the same size; the extent of the play frame likewise depends on the shared understanding. A racy joke told to an old college roommate at a bar may not be received in the same manner when told to families gathered at a church social. In that latter context, the audience may not consent to the same kind of aesthetic stage that the communicator desires. If the communication falls outside of the play frame for any reason, the communicator may not be able to claim immunity from the normal conversational maxims—and may face consequences.

---

⁴ This is generally true, unless the response involves what Michael Billig referred to as “un-laughter” which serves as a rejection of the frame (Billig 2005). This is not the same as rejecting a joke for simply being a bad joke. An unclever dad joke that elicits a groan and an eye-roll still falls squarely within the play frame, because it is still a statement that is clearly intended not to be taken at face value and is understood as such.
2.3 Ambiguity and Mixed Intentions

All humor involves some degree of ambiguity. This is because the metacommunicative cue that all humor carries—the signal that communicator does not actually intend what she is communicating—renders the entire communication a paradox (Marsh 2018, 447). As Fry (expanding upon Bateson) explains, the communication itself is not to be taken literally, while at the same time, the cue that establishes this understanding (which can be a part of the communication itself) is meant to be taken literally. This becomes similar to the Epimenides paradox “I am lying”: “If the situation is unreal, so is the cue-message, and the situation becomes real. If it becomes real, so does the cue-message, and the situation is unreal. A circular paradox is quickly apparent” (Fry 2010, 143).

In other words, ambiguity is woven into the very fabric of humor. In fact, ambiguity may be the fundamental thing about humor that makes it humor. Humor conveys meaning not by directly communicating it, but by directly communicating what it does not mean. The humorous meaning is therefore indirect, which leaves it up to the listener to decipher the ambiguity and figure out the non-serious message—and even to decide whether to accept it as humor and how to respond. This dynamic process of dealing with this ambiguity is what allows humor—and perhaps all forms of play—to exist.

5 Of course, humor is not unique simply because it communicates something other than a literal meaning. There are other forms of nonliteral communication, including things like metaphor, idioms, hyperbole (or understatement), rhetorical questions, antiphrasis, verbal irony, and sarcasm (Kreuz 2019). What sets humor apart from these other forms of nonliteral communication is that humor serves a distinctly different purpose. While these other devices serve the goal of conveying some form of information, the primary purpose of humor is not strictly about communicating but about playing. Thus, even when humor employs one or more of these specific devices, it uses them differently to serve that purpose. For example, when a person asks a rhetorical question, they do not expect or desire an answer, but intend to make a point—to convey information. When a person looks to the sky and asks, “Why me?” or responds to a question by asking, “Is the Pope Catholic?” they are expressing frustration or confusion, or they are saying they think something is obvious. With humor, a question may also be posed without any regard for the answer, but it is not used as a way to convey information. When Steven Wright asks, “How do you tell you’re out of invisible ink?” he is not posing the question to make a point—he’s just simply playing.

6 If communication could only convey serious, literal meanings, it is hard to envision any kind of humor. Things like puns and incongruity-based humor in general would certainly never get off the ground, but even superiority-based humor like ridicule would not be possible. (Faulkner’s slight of Hemingway, for example, is potentially humorous not for what it literally said, but for subtly conveying the non-literal message that Hemingway did not use—and perhaps was not capable of using—big words.)
This ambiguity can, of course, often lead to confusion over the speaker’s actual intentions, which can lead to confusion over whether the communication truly falls within the play frame. For example, humor can consist of simply stating the exact opposite of the intended message, which is the essence of satire and sarcasm. While this kind of humor is usually obvious to most people, the capacity to decode such communications can vary, as illustrated by the character of Sheldon from the show *The Big Bang Theory*, who has a longstanding difficulty discerning sarcasm:

Leonard: Hey, Penny. How’s work?
Penny: Great! I hope I’m a waitress at the Cheesecake Factory for my whole life!
Sheldon: Was that sarcasm?
Penny: Nooo.
Sheldon: Was *that* sarcasm?
Penny: Yes.
Sheldon: Was *that* sarcasm?

Here, the speaker clearly intended the sarcastic comment as a form of humor. This contributes to the humor of the situation, because the obviousness of her attempt at humor makes it even more surprising and incongruous that the listener, a genius-level theoretical physicist, is sincerely confused and unable to discern her non-serious intent.

Humor, however, does not need be driven exclusively by a non-serious intent; sometimes a speaker may have multiple intentions, both non-serious and serious. Although communications made within the play frame are by definition intended to be taken as something other than a serious and literal statement, they may still be conveyed with the intention that they also be taken seriously and literally *at some level*.

---

7 The Faulkner insult towards Hemingway may have been intended to be taken seriously *as an insult*. Nonetheless, this is the wrong kind of “taking seriously”; the statement was still not intended to be taken seriously as a literal fact. Rather, it intended to convey something different—in that case, a sense of superiority by using an exaggeration as a way to ridicule. An example of an insult that relies on incongruity as much as superiority can be seen in one of Oscar Wilde’s oft-repeated quips: He has no enemies, but is intensely disliked by his friends. As with the Faulkner jibe, if this were leveled against another person, it may be meant seriously *as an insult*, but it would not be meant as a serious statement to be taken literally. Indeed, it certainly suggests the person is unlikable or perhaps has only superficial relationships, but the fact that it would seem impossible by definition to be disliked by one’s friends makes it impossible to take at face value. (This quote is taken from *The Picture of Dorian Gray*. The original passage reads: “Two of the
as jokes—as non-serious statements designed to amuse, ridicule, or express cleverness—which are also meant, at least partially, to convey a serious message. Such a joke is what Gimbel refers to as an “impure joke” which contains these multiple intentions, in contrast with “pure jokes” that are intended solely as aesthetic acts to amuse, ridicule, or express cleverness (Gimbel 2017, 111–2). This kind of mixed-intentioned humor is often associated with comedians who speak about social issues. For example:

One explanation for the gender wage gap is that men typically gravitate towards higher paying jobs, like doctor, engineer, CEO... While women tend to gravitate towards lower paying jobs, like female doctor, female engineer and female CEO. (Jeremy McClellan)

This statement is still a joke; it is an observation that is not intended at face value, but as an attempt at humor. (It is obviously not meant to convey that “female doctor,” “female engineer,” and “female CEO” are specific types of jobs that women tend to pursue.) But at the same time, one underlying point that it literally expresses—that there is a gender wage gap—does appear to be intended seriously. The statement’s status as a joke does not change just because it carries this more serious message in addition to the non-serious message.⁸

When humor employs mixed motivations like this, the non-serious message of the joke can function like a Trojan Horse to deliver the serious message almost covertly. As self-proclaimed social justice comedian Negin Farsad observed about social justice humor, “It’s funny but sneaky. Like you could be hearing an interesting treatise on income inequality that’s encased in a really sophisticated poop joke” (Farsad 2016).⁹

On the other end of the spectrum, humor with mixed motivations can also present the serious, literal message overtly:

people he had never seen before, and the others consisted of Ernest Harrowden, one of those middle-aged mediocrities so common in London clubs who have no enemies, but are thoroughly disliked by their friends” [Wilde 1983, 166]).

⁸ This is the case with play in general. For example, a person can instigate a play fight with someone who they may also desire to challenge to a real (non-play) fight. As long as the instigator’s actions remain consistent with play, the fact that they also possess an additional more serious intention does not automatically convert the activity of play into an actual fight. As long as the non-serious intention is communicated and accepted (i.e., the play frame is established), the presence of additional serious intentions, by themselves, do not change that activity. ⁹ In her 2016 TED Talk titled, “A highly sophisticated taxonomy of haters,” Farsad also offers an explanation why this kind of social justice humor works: “Because, first off, it makes you laugh. And when you’re laughing, you enter into a state of openness. And in that moment of openness, a good social justice comedian can stick in a whole bunch of information, and if they’re really skilled, a rectal exam” (Farsad 2016).
Racism isn’t born, folks, it’s taught. I have a two-year-old son. You know what he hates? Naps. End of list. (Dennis Leary)

One in four Americans is skeptical about climate change. Who gives a shit? That doesn’t matter. You don’t need people’s opinions on a fact. You might as well have a poll asking which number is bigger, five or fifteen? Or do owls exist? Or are there hats? (John Oliver)

There is no question what the serious messages are in these examples—they are explicitly stated up front. Nonetheless, they are also clearly jokes told within the play frame, as they are fundamentally non-serious statements that come from clever, incongruous analogies (and are presented by comedians).

The fact that an instance of humor may bear multiple intentions, both serious and non-serious, therefore does not by itself remove it from the play frame and render it subject to the normal rules of communication. However, the fact that a communication can (and often does) involve multiple intentions does mean that it can sometimes be difficult for a listener to discern whether it truly is intended as humor. This raises important questions about how we can assess whether a communication legitimately belongs within the play frame or should be viewed as falling outside of it.

3. The Play Frame and Ethics

Like all forms of behavior, communications can be subject to ethical constraints. Generally speaking, a communication runs afoul of these constraints when its content violates some accepted standard of decorum or conduct. Perhaps the single most important effect of the play frame is that by establishing that a joke (or any kind of humor communication) was intended and received as a non-serious act of play, the speaker can avoid accountability for a communication whose literal content would violate applicable norms. In short, if the communication falls within the play frame, it can be insulated from ethical scrutiny.

After all, if the speaker was “just joking,” then they did not actually intend that problematic literal content of the communication—and if the communication was “just a joke” and was therefore an attempt at play rather than an attempt to convey information, then others should not be harmed by it. Conversely, if the communication falls outside the play frame, the speaker cannot claim immunity from the typical rules of conversation. In fact, they can be viewed as having truly meant the literal content of the communication and held accountable for it.

Practically speaking, the two most obvious ways that a person can violate such ethical norms by telling a joke are: (1) when they tell a joke that actually
causes harm—particularly if they knew or had reason to know that such harm could result (like making a rape joke knowing there were rape survivors in the audience who would find the joke traumatic); and (2) when they actually intended the underlying content of the joke literally (i.e., if they actually believed and intended to convey the offensive content, like a racist making a racist joke). In both instances, the communications represent unethical acts that may fall outside the protection of the play frame. However, as explained below, in both instances the analysis can also be complicated.

The first type of unethical communication can be said to involve a joke that “crosses the line.” The fact that the communication resulted in harm suggests that it was not interpreted by the audience as a purely aesthetic act or as an attempt at play, but as a serious communication with literal meaning. As such, the joke would fall outside of the play frame as construed by the audience. Had the audience construed the play frame more broadly and thus interpreted the communication as a non-serious artistic expression, it would not have had the same power to harm, and the audience would have understood it was “just a joke.” The fact that it caused harm, however, shows that the joke teller either miscommunicated the bounds of the play frame or exceeded it, or both. (Indeed, knowingly causing harm can itself be a violation of an ethical norm.) Moreover, even if the joke teller did not sincerely intend to harm anyone in the audience by the communication, it could be claimed that they should have known—based on the applicable ethical norms—that the content was outside of the play frame as reasonably conceived by the audience.

Reaching this conclusion is rarely this easy though, because the analysis is rarely as straightforward. Establishing the contours of the shared understanding of the play frame is often difficult, especially when it involves potentially subtle metacommunicative cues, and there can often be a disconnect between the communicator’s conception of the play frame and how the audience views it. Individual members of any audience may also have different baseline conceptions of how far the play frame can extend: Some may view it so broadly that virtually no joke could ever be deemed unethical, while some may construe it unreasonably narrow. Similarly, jokes may have very different impacts on different audience members—some people may be particularly susceptible to being offended while some may be offended at virtually nothing. In other words, determining whether a joke crosses the line is especially difficult because the speaker and the audience draw their own lines.

The second type of unethical communication can be said to involve a serious statement that the joke teller intends literally but only veils as humor to avoid being held accountable for it. It could also include a joke with mixed intentions, where the joke teller may at some level still intend it to be taken at least partially
as a non-serious communication akin to an aesthetic act. In these cases, it is because the joke teller does truly intend for the offensive content to be taken at face value that the humor (and the play frame itself) may be seen as a pretext. The joke teller could therefore be deemed responsible for the literal content of the communication as if they made the offensive statement seriously outside of the play frame.

In this case, the analysis is both more streamlined and more complex. It is more streamlined because rather than considering the dynamic process of how the joke teller and the audience constructs their shared understanding of the existence and extent of the play frame, the analysis here concentrates only on the joke teller and their underlying motivation. Specifically, the focus is on whether their underlying intent was simply to play or was to actually convey the problematic informational content, and thus, whether their attempt to construct the play frame was sincere and appropriate. The analysis is also more complex because of the inherent ambiguity of humor and its potential to carry multiple motivations, which can make discerning the speaker’s true intent—and when it can override any accompanying non-serious intent—extremely difficult.

This is especially true for contemporary comedians who develop on-stage personalities that are similar but exaggerated versions of themselves. For instance, Sarah Silverman’s early standup personality was something of an extremely superficial New York Jewish female, who said things like:

I don’t care if you think I’m racist. I just want you to think I’m thin.

I dated a guy who was half-black, but he dumped me because I’m such a loser. Wow, I shouldn’t say things like that, I’m such a pessimist... He’s actually half-white.

Without knowing her personal background, it seems possible that many who heard these lines and who thought her stage persona was real could have viewed her as a shallow bigot who meant what she said. In this case, however, the fact these were so incongruous and so pithy makes it relatively clear they were carefully crafted jokes.

Sometimes, however, it may be much harder to view offensive comments as attempts at humor. A few years ago, comedian Shane Gillis was hired as a cast member for Saturday Night Live, but was quickly fired after it was revealed that he used racist, sexist, and homophobic slurs in his podcast (Otterson and Schneider 2019). Unlike Silverman’s joke, at least some these comments did not come off as crafted one-liners, but as more spontaneous insults. The public statement issued by SNL about Gillis’s termination noted, “We were not aware of his prior remarks that have surfaced over the past few days. The language he used is offensive, hurtful and unacceptable.” Gillis responded that “I’m a come-
A comedian who pushes boundaries. I sometimes miss. ... My intention is never to hurt anyone but I am trying to be the best comedian I can be and sometimes that requires risks” (Drury and Porter 2019).

Some people accepted Gillis’s explanation, including presidential candidate Andrew Yang which was notable because some of Gillis’s racial comments were targeted at Asian Americans (Stevens 2019). Gillis’s comments were also made on a podcast which is a relatively unrestricted forum where anything goes and where people typically know what they are getting—i.e., it is where the play frame can be extremely wide. Nonetheless, Lorne Michaels and others at SNL concluded that his comments were still out of bounds. Their concern was not simply about the subject matter; after all, SNL skits also routinely make jokes about race, sex, and sexual orientation. Given their public statement (and their reference to Gillis’s “remarks” rather than his “jokes”), it seems that they felt his comments may have been motivated by serious, rather than non-serious intent—or at least could be perceived as such—and were therefore not protected as falling within the play frame. Still, this was something on which many people reasonably disagreed—some accepted Gillis’s broad conception of the play frame while others rejected it—which shows how difficult this analysis can be.

3.1 Manipulating the Play Frame

As can be seen, anything said inside the play frame is at least presumptively construed as a non-serious or aesthetic act that the speaker did not intend literally. Even if the content is deemed unethical, a speaker can make the claim that the audience should be no more offended at them than they would be at a work of art in a museum—and if they are still offended, they are being unreasonable. Moreover, because the very nature of humor is ambiguity, a joke teller can claim the protection of the play frame broadly, as they can defend even the most seemingly offensive and harmful comment by claiming they were “just joking.” The play frame can therefore serve as a form of blanket immunity for speakers of controversial, offensive, or otherwise unethical statements.

This can be seen in Gillis’s explanation for his past remarks. He claimed that he pushes the boundaries of propriety in his humor, and that sometimes he “misses” which seems to admit his attempts at jokes are sometimes offensive, perhaps even to an inappropriate degree. But at the same time, he also emphasized that he never meant to hurt anyone, suggesting his comments should still be protected as non-serious statements falling within the play frame. This effectively distances himself from the offensive comments—they were not things he meant or believed but just jokes that happened to “miss.” In this way, he
seems to take responsibility for them, but only as bad or failed acts of artistic expression, not as potentially offensive or hurtful comments.

Explaining away potentially offensive comments by using the play frame in this way is not just a defensive maneuver—it can also operate offensively by accusing those offended or harmed as having “no sense of humor.” Although humor was seen negatively throughout most of history, this has dramatically changed; having a sense of humor is now widely accepted to be an extremely positive, if not essential trait.¹ It is viewed as a reflection of a well-adjusted personality and often considered the most sought-out quality in a mate (M. Smith 2009, 157–8). It is also associated with being able to laugh at oneself and the capacity for critical self-assessment, and implies a strong character that is not easily wounded by others’ comments. Conversely, being “humorless” is associated with distinctly negative psychological traits, such as egotism, extremism, inflexibility, and even mental illness (M. Smith 2009, 158). Being accused of lacking a sense of humor is therefore more than just an insult for not having a beneficial personality trait, but actually cuts to the heart of someone’s character, painting them as fundamentally deficient as a person.¹¹

Consequently, when a speaker who is accused of making an offensive statement relies on the play frame by claiming they were “just joking,” and therefore implies (or explicitly states) that anyone who was offended “can’t take a joke,” she effectively shifts the burden in the debate. In fact, this even saddles the audience member with a significant burden and one that becomes personal. The issue is no longer about whether the joke teller can prove that her comment was not inappropriately offensive but becomes about whether the audience member can prove two things: that the play frame was exceeded and that any harm he suffered was not inappropriate—in other words, he must establish that he, personally, is not humorless. In this way, invoking the play frame can be used as both a shield and a sword.

The play frame can also provide the speaker with deniability (or “decommitment”) and “paves the way for the possibility of a relatively easy retraction of a truthful meaning, if necessary” (Dynel 2018, 401). This can be a very positive

---

¹ This early view of humor as a vice largely stems from the popularity of superiority theory as the first theory of humor—the basic idea that humor is essentially about ridicule and comes from feeling better than someone else. The later development of incongruity theory and the notion that humor can arise from deciphering incongruous situations led to an understanding that humor can be a distinctly cognitive endeavor and thus, can even be a virtue (Morreall 2009, 11).

¹¹ Indeed, to be labeled as having no sense of humor suggests that one is “literally an incomplete person” and some people doubt that there could even exist a human being without any sense of humor (Wickberg 1998, 85, 88–9; M. Smith 2009, 158).
thing—in fact, this is the basis of the romantic view of the court jester who “speaks truth to power” by couching otherwise unsayable criticisms of the king as harmless jokes. It also allows comics to discuss taboo subjects that could not be discussed in normal conversation. For example, the comedian Lenny Bruce was infamous for pushing social boundaries in order to make larger points. In one skit, he began by asking the crowd, “By the way, are there any n*ggers here tonight?” (Emblidge 1977, 104). He then repeatedly used that word and a host of other ethnic slurs to describe people in the audience. As he said the intensely offensive words over and over and over again, his increasingly exaggerated characterizations sounded more and more absurd. He then concluded:

The point? That the word’s suppression gives it the power, the violence, the viciousness. If President Kennedy got on television and said, “Tonight I’d like to introduce the n*ggers in my cabinet”, and he yelled ‘n*gger’ at every n*gger he saw, ‘boogyboogeyboogeyboogeyboogey, n*gger’ till n*gger didn’t mean anything anymore, till n*gger lost its meaning—you’d never make any four-year-old n*gger cry when he came home from school. (Emblidge 1977, 104—5)

Because he was in the play frame (although clearly pushing its boundaries), Bruce was able to say what was typically unsayable, and he was arguably able to make his point much more strongly than had he just given the point. He was able to show the audience what he meant, not just tell them. Because his words were ultimately understood as a joke, he was also able to blunt the harm that the members of the audience would normally suffer from being subjected to such racial epithets, as they knew he did not intend the comments literally. In fact, it became clear that his intent was not to use the words to cause harm, but to do the exact opposite—to use the words to undermine their potential to harm.

However, it is this positive aspect of humor and the play frame—its ambiguity of meaning and distance from literal intent—that also leaves it susceptible to being invoked disingenuously and used to communicate harmful anti-social information or even just something that would be otherwise inappropriate to say. This often happens when someone makes a serious statement and then, in the

---

12 Bruce would, of course, object to modifying the n-word as was done in this quotation of Bruce’s speech. Indeed, that was his entire point—that such an approach actually contributes to the problem it is meant to address. Nonetheless, that position remains controversial, and may also not translate in the same way from a standup monologue in a comedy club to an academic paper. This footnote and the act of censoring the word above is not meant to assert a position in this debate, but to show recognition of the issue and the underlying concern.
face of criticism, tries to distance himself from it by claiming after the fact that he was “just joking.” For example, during a speech to law enforcement, President Donald Trump told the police officers present that when they assisted handcuffed suspects into their squad cars they should not worry about protecting the suspect’s head from bumping against the car (Berman 2017a). Many in the audience laughed or applauded, but when numerous police departments and civil rights groups issued statements condemning the comments, the White House and the president’s allies explained that he was just joking (Berman 2017b). It still remains unclear whether he truly meant the comments or was joking—which is the point.

Another notable example occurred during the 2016 presidential campaign when Trump declared: “Russia, if you’re listening, I hope you’re able to find the 30,000 emails that are missing” (Dynel, Brock, and de Jongste 2016). The next day he claimed he was just being sarcastic, and his supporters explained to media outlets that it was nothing more than a joke. Additionally, in his written responses to Special Counsel Mueller during the subsequent investigation into Russian interference in the election, Trump similarly characterized the comment as having been made “in jest and sarcastically, as was apparent to any objective observer” (Dowd 2019). It was, however, later revealed that the day after Trump made the original statement, Russian hackers actually made their first attempt at breaking into computer servers at Hillary Clinton’s personal offices (Schmidt 2018; Mazzetti and Benner 2018).

This kind of example shows how the play frame can be used not simply to avoid accountability, but how it can also allow a person to engage in a type of coded communication. Assuming for the sake of argument that Trump seriously intended to ask Russian hackers to find Clinton’s emails, the fact that Russian hackers actually did try suggests he was able to get his message across, while still being able to claim that he had no such intent. Of course, it cannot be known if the president truly intended to ask the Russians to engage in those ac-

---

13 As reported in the Washington Post:

“When you guys put somebody in the car and you’re protecting their head, you know, the way you put their hand over?” Trump said, miming the physical motion of an officer shielding a suspect’s head to keep it from bumping against the squad car. “Like, don’t hit their head, and they just killed somebody—don’t hit their head,” Trump continued. “I said, you can take the hand away, okay?” (Berman 2017a)

14 For example, the twitter account Blue Lives Matter posted a video of Trump delivering the comments and wrote: “Trump didn’t tell police to go out & brutalize people as the media would have you believe. It was a joke. 😂” https://twitter.com/bluelivesmtr/status/891085581152866306
activities—and that is exactly why using the play frame in this way can be so problematic. Just introducing the potential that a speaker was “just kidding” is enough to insert ambiguity of intent into the communication—and this makes it difficult for the audience to ever know what the speaker’s true intentions are. It becomes forever steeped in uncertainty: Even if they later deny they were kidding, the intent behind that denial becomes suspect.

Humor can therefore provide a vehicle for covertly communicating antisocial information in plain sight. When a controversial statement that could not otherwise be said (such as a statement supporting police brutality) is made and then later described as just a joke, those in the audience who are receptive to such a controversial remark will still hear the initial statement—and they will often know that this kind of remark could not actually be said in public, so they understand that characterizing it as joke was not actually true but was merely a practical necessity to avoid criticism by the “others” in the audience who do not share that view. This is not a case of surreptitiously conveying a controversial message, like a message laden with secret hidden meaning, but a case of saying the controversial message clearly and explicitly, while still claiming that it was not really said. Humor is uniquely able to do this.

It is also important to note that members of the audience have the power to unfairly manipulate the play frame as well. They do this by claiming that the play frame was much narrower than it was, and then try to hold the speaker responsible for intending statements that were truly made without serious intent. This is what happened when Al Franken first ran for U.S. Senate in 2008. As he described in his book, *Al Franken, Giant of the Senate*, his political opponents took his earlier satirical works and ran them through the “DeHumorizer™” which stripped away “anything that made it clear I was engaging in irony (or hyperbole, or ambiguity, or any number of other comic devices)” and presented them to voters as statements to be taken at face value. “Without their comedic context, those words often weren’t funny anymore. And in fact, they could appear downright offensive” (Franken 2017, 79).15 Franken won the race, but only by 312 votes, in one of the closest elections in Senate history.

---

15 In one instance, the Minnesota Republican Party released a letter signed by six prominent female Republicans responding to a satire he wrote for *Playboy* eight years earlier titled “Porn-o-Rama!” about visiting a sex institute where he had sex with robots. In the letter, they claimed his piece demeaned and degraded women “as thoroughly and disrespectfully as any article we have ever seen.” They even claimed he only attempted to defend his writing as satire, and implied he personally had the desire “to have sex with women or robots that look like women simply to give yourself a good time” (Jaffe 2008).
3.3 Social Hierarchies and the Boundaries of Humor

The play frame may be universal to humor, but it does not occur in a vacuum; the existence and extent of the play frame is influenced by existing social hierarchies and power structures. Because it depends on a shared understanding between the joke teller and the audience, when a joke teller is from an empowered group (such as white males) and tells disparaging jokes that target marginalized groups (such as people of color or women), those in the audience from those marginalized groups may feel pressure to go along and accept the play frame even when they find the jokes inappropriately offensive. Ultimately, this means that the play frame may artificially become much wider when it is used by those at the top of the social hierarchy and when it is used to protect jokes that disparage those in less powerful social groups.

There are a number of reasons for this pressure on those in marginalized groups to accept the play frame even when it is used to target them. As a general observation, subordinates tend to laugh at the jokes of their superiors more often than the superiors laugh at their subordinates’ jokes (M. Smith 2009, 163; Coser 1960). There seem to be obvious reasons for this based on the power relations—people typically have incentives to please those with power over them, and disincentives to rankle them. Moreover, the action of a person without power rejecting the play frame offered by a member of the powerful majority can be seen as an act of voluntary “othering” oneself from that dominant group. This can be especially true when the disparaging humor overtly flouts social convention, because those who laugh at such humor become akin to partners in crime sharing a guilty pleasure, which promotes a special solidarity (A. Smith 2009, 160). If a target of the humor laughed, it would provide cover for others to laugh because they could claim that they were not harming anyone; however, if the target did not laugh, it could be interpreted as a judgment that the others were wrong for laughing, which could compound existing divisions between the groups. Thus, for a minority to reject the play frame under these circumstances would be to reinforce their outsider status.¹⁶

This is exacerbated by the potential criticism that not laughing at a joke indicates a lack of a sense of humor. As discussed above, such a criticism enables a speaker to use the play frame not just defensively but offensively—and this

¹⁶ Those in marginalized groups are often left in a “double bind.” For women faced with sexist jokes, for example, it has been noted that they “are left with two options—laugh at the joke or express dismay at the joke’s content ... if she laughs, she is complicit in her own group’s humiliation. If she does not laugh then she is a ‘spoiled sport,’ someone with no sense of humor ....” (Pérez and Greene 2016, 3; Bemiller and Schneider 2010, 463).
kind of accusation can be especially potent when directed at a member of a marginalized group who already faces being labeled as an outsider. Just as sharing a similar sense of humor brings people together by accentuating their common interests, not sharing a sense of humor accentuates their differences. Moreover, being humorless is often seen as not just a personality defect but as even being less human—and dehumanization serves as a psychological justification for treating other people differently.¹

Ultimately, the pressure that can be brought to bear on members of marginalized groups can be significant, particularly when those doing the pressuring wield the power of being in socially dominant hierarchies. This can be seen in the reaction to comments made by comedian Daniel Tosh at Laugh Factory in Los Angeles in 2012. According to audience accounts, after Tosh discussed the merits of rape jokes, an audience member yelled, “Actually, rape jokes are never funny!” Tosh, a white male who was known for his deliberately controversial and offensive humor, responded, “Wouldn't it be funny if that girl got raped by like, five guys right now? Like right now? What if a bunch of guys just raped her?” (Holm 2016, 110). The incident resulted in both moral condemnation of Tosh for making the remarks, as well as fervent defense of him and the freedom to joke about anything. As Nicholas Holm described the reaction, it became a familiar debate with one side “de-emphasizing the comic nature of the material and interpreting it as a directly offensive and hurtful statement” and the other side “invok[ing] a language of comic distance and performance that positioned the remarks as playful rather than aggressive” (Holm 2016, 111).

The debate, however, was not simply an intellectual dispute about freedom of speech but was very largely influenced by the power relations associated with gender. After a televised debate about the incident and the propriety of rape jokes that took place between comedian Jim Norton and feminist blogger Lindy West, West received a great deal of criticism for her position that comedy should not be free from constraints.¹⁸ The criticism she received was apparently not just from internet trolls, but from other comics and even co-workers, and the most extreme criticisms were less about challenging her ideological positions and more about challenging her status. In other words, she was seen as less entitled to question the extent of the play frame established by Tosh because she

¹ In fact, some jokes can be told to specifically targeted individuals or groups with the goal of having them reject the play frame to reinforce their outsider status. Those targets need not explicitly protest to reject the frame but could exhibit their rejection even unintentionally by simply not laughing (or responding with “unlaughter”). That reaction can then be seen as a reason to blame them for their own outsider-ness (M. Smith 2009, 164).
¹⁸ This debate can be viewed here: https://youtu.be/uplqyxHsBmg.
was a woman who was overly sensitive. As she reported, internet commenters also wrote extremely aggressive comments, seemingly designed to silence her—things like, “I’m just bitter because I’m too fat to get raped and suggesting that the debate would have been better if it had just been Jim raping me” (Travers 2013). Although those comments were outliers and obviously do not represent the majority of people who hold that general position, they still reveal the existence of the power dynamic as an influence on how to construe the play frame.

It is worth considering if Tosh’s comments would have been received differently if they were told by a female. Jokes disparaging a certain group are seen differently when the joke teller is a member of the targeted group, seemingly because the joke teller’s identity emphasizes the non-seriousness nature of the joke—the perception is that the joke teller could not mean the jokes literally because she would be disparaging herself as well.¹ This makes it more acceptable for a woman to tell sexist jokes, because sexist jokes told by a man can be much more easily seen as statements intended seriously. Thus, had Tosh been female, the comments would have been at least somewhat less controversial.

Additionally, had Tosh been female the comments themselves may not have had the same power or potential to harm. A woman is uniquely situated to tell sexist jokes without being seen as intending them literally; therefore, even if this still promotes a greater acceptance of rape jokes, it does not necessarily promote accepting them as seriously intended statements. In contrast, a male who tells sexist jokes is more likely to be seen as meaning them literally, which may pave the way for acceptance of not only sexist jokes but acceptance of the underlying serious intent behind them—which can result in an increased degree of misogynistic behavior.²

¹ In fact, when people poke fun at their own group it can also be empowering—to turn the critical focus back on the forces that marginalize the group in the first place. As an example, Key & Peele have a skit called Auction Block, which takes place in an early American slave auction: https://youtu.be/zB7MichlL1k (In this, two slaves become jilted when no one bids on them, and increasingly try to promote themselves to the buyers.) As one of the head writers explained, it was important that the slaves themselves controlled the narrative of the joke, so they were actually the ones in power in terms of the joke (Jung 2015).

² There have been a number of psychological studies that have found a correlation between exposure to sexist humor and an increase in tolerance for discrimination against women, as well as a connection between sexist humor and things like sexual harassment, discrimination, tolerance for sexual violence, and even rape proclivity. Moreover, these behaviors are typically present and increased to a greater degree in those who already harbor strong sexist attitudes (Pérez and Greene 2016, 3; Ford 2000; Ford et al. 2008; Ford, Wentzel, and Lorion 2001; Romero-Sánchez et al. 2010; Ryan and Kanjorski 1998; Thomae and Viki 2013).
In terms of offensive humor generally, this does of course present a familiar debate. Should a comedian be able to joke about anything? Or are there some things that should never be joked about? Although this is a normative debate couched in terms of the limits humor, the real debate may be a quasi-normative one about how far the play frame extends. Importantly, much of the debate over the Tosh comments did often discuss the context of the larger social setting—rape jokes being made by a male on stage, directed at a lone woman in the audience—and the harm that such a dynamic can have. However, analyzing this in terms of the play frame places a greater emphasis on the more dynamic nature of humor as a negotiated understanding between the joke teller and her audience. Humor is an inherently ambiguous form of communication, and its underlying intent and meaning will always be influenced by many practical realities—the content, the form of the communication, the communicator, the audience, the space in which the communication was made, and so on. Offensive humor pushes boundaries, so it is important to consider the nature of those boundaries at least as much as the humor itself.

To be clear, the play frame analysis does not simply reveal where the hard-and-fast ethical boundaries are. Indeed, most ethical issues—especially the ones that become the subject of debate—are far from black and white, and there are often sound reasons and sincere beliefs supporting both sides. What the analysis does is help provide an additional perspective and a different approach to understanding how these ethical boundaries may actually be drawn in pencil rather than ink. This can provide us with additional important considerations to weigh in the process of sketching out these boundaries—a process that itself may be inherently dynamic and collaborative.

Finally, it is worth noting how the nature of these ethical boundaries relate to humor’s ability to drive social change. Humor works by breaking rules—after all, by its very nature, it violates the social conventions of conversational maxims. It is also because of this and its ability to “say the unsayable” that it is able to flout all kinds of social conventions. This subversive character is what allows humor to promote positive social change by challenging problematic but long accepted norms, and by conveying controversial ideas that may one day gain acceptance. However, when social progress occurs and the once-controversial ideas become normalized, they then become targets for humor. As things like sexism and racism and homophobia come to be more universally understood as wrong, it becomes more subversive—and thus more potentially humorous—to promote them. Given the ambiguity inherent in humor, such antisocial and potentially unethical messages may be meant not only as jokes but literally, while still being protected as falling within the play frame.
This leads to the paradoxical conclusion that the more successful humor is at driving social change, the more humor can also be used to undermine it. While the paradoxical nature of humor may be the source of its power, whether it is a power for positive or negative ends depends not only on how it is intended, but on how it is construed. The boundaries we draw around humor will therefore always depend on a shared understanding of the play frame, because the window of the play frame opens from both sides.

References


