Scientists joking on social media: An empirical analysis of #overlyhonestmethods

Molly J. Simis[1]

University of Wisconsin-Madison, Madison, United States

Haley Madden

University of Wisconsin-Madison, Madison, United States

David Lassen

University of Wisconsin-Madison, Madison, United States

Leona Yi-Fan Su

University of Wisconsin-Madison, Madison, United States

Dominique Brossard

University of Wisconsin-Madison, Madison, United States

Dietram A. Scheufele

University of Wisconsin-Madison, Madison, United States

Michael A. Xenos

University of Wisconsin-Madison, Madison, United States

Introduction

Science has long been described as an undertaking by those in the ivory tower, and as inaccessible to lay audiences. This conceptualization of scientists is problematic, because it is essential that publics trust scientists and accept well-grounded scientific theories. Additionally, the public is increasingly involved in science and scientific decision-making. Meaningful engagement with public audiences can help to humanize scientists, which researchers argue is desperately needed and would help build trust. Indeed, there is an ongoing discussion among scientists about the importance of trust among public audiences (e.g., Ravetz & Saltelli, 2015; Yarborough, 2014).

A recent brief in the high profile science journal *Nature* says that "[m]aintaining the public's trust in science calls for an urgent evaluation of its imperfections and vulnerabilities" (Ravetz & Saltelli, 2015). The authors call "[s]ocial-media channels" outlets that "are starting to teach the public more about new views of science," and argue that a more scientifically literate audience is potentially around the corner partly due to science's portrayal on social media (Ravetz & Saltelli, 2015). However, some of the ways that science is portrayed on social media are not necessarily going to heighten public trust in science. For example, the #overlyhonestmethods hashtag on the microblogging platform Twitter has been a place where scientists reportedly share "all the hilarious science not fit for publication" (Grenoble, 2013). Bringing methodological issues that are unfit for peer review to the public's attention in the content of humor is ethically problematic. In the present research, we delve into the #overlyhonestmethods phenomenon, exploring the characteristics of the posts and contributors, as well as the integral and complex role that humor plays in

[1] Corresponding author, email: molly.wilkinson@gmail.com

this previously private conversation among apparent scientists that is taking place in a public space. First, we will examine some background information.

Twitter has become a popular communication tool for many scientists, including Neil deGrasse Tyson, Richard Dawkins, and Sam Harris (You, 2014). More broadly, however, a Pew poll of AAAS scientists in 2014 found that 27% of respondents use social media regularly, and their communication efforts are directed primarily towards the public or equally to the public and other scientists (Rainie, Funk, & Anderson, 2015). The use of Twitter among college faculty is higher still at 35.2% (Liang, Yi-Fan Su, Yeo, Scheufele, Brossard, Xenos, Nealey, & Corley, 2014). Recently, science communication scholars found that regular Twitter use by scientists can increase a scientist's impact, potentially further encouraging scientists to use the social media platform (Liang et al, 2014).

Twitter is by its nature discursive. Users can carry out conversations (publicly or privately) or choose to passively receive information from a variety of sources. In addition, users can utilize hashtags to indicate their involvement with a particular topic, thus contributing to a much larger conversation. Despite the potential for increased engagement compared to other media, some researchers have argued that Twitter is largely used as a broadcasting platform, rather than a way for users to personally engage with individuals (Brenner, 2014; Finfgeld-Connett, 2015; Neiger, Thackeray, Burton, Giraud-Carrier, & Fagen, 2013). However, social media in general and Twitter in particular are still fundamentally different from broadcast media. Tweets are part of a discourse, and people posting do contribute to a large, evolving conversation. Additionally, interested parties can opt in to passively receive content from scientists, and parties who do not explicitly opt in may find that science worms its way into their feeds through retweets (re-postings of other users' tweets) or trending topics, which are indicated by hashtags.

Case study: #overlyhonestmethods

From this background, the development of the overlyhonestmethods hashtag can begin to take shape. Some scientists are increasingly trying to engage with the public and are taking to social media, like Twitter, to do so. Other scientists may simply be using Twitter to have conversations with other scientists in public. The emergence of new, direct channels through which scientists may interact with the public may have also affected the nature of public science commentary. Potentially unrestrained by traditional journalistic practices and goals, scientists who publicly comment on their disciplines may do so in previously unavailable ways.

On January 7th, 2013, a neuropharmacologist who goes by Dr. Leigh tweeted: "@dr_leigh: we did experiment 2 because we didn't know what the (redacted) to make of experiment 1 #overlyhonestmethods" and the hashtag overlyhonestmethods was born. Since then, the hashtag has been used almost 60,000 times by 33,713 Twitter users, with as many as 4,800 unique users posting on a single day. Originally used by @dr_leigh (who no longer uses this handle) to "blow off steam" by making jokes, the hashtag has gained a great deal of media attention (Gould, 2013). Publications writing about the hashtag include the *Huffington Post*, *The Guardian*, *Scientific American*, *The Telegraph*, and the *Boston Globe*.

The content distributed through #overlyhonestmethods tweets may therefore serve as a stark departure from earlier public writing about science or from scientists. Immediately apparent is its central use of humor, a stylistic element at most infrequently included in traditional science journalism. The tweet uses what appears to be exaggerated frustration to create a sense of shocking honesty. In this way it is consistent with most modern theories of humor that emphasize the central role of humor in communicating ideas and relationships that are unexpected or novel linguistically (Raskin, 2012), cognitively (Hurley, Dennette, & Adams, 2013), or socially (McGraw & Warren, 2010). At the same time, however, humor can be a strikingly imprecise and demanding form of communication, generally requiring audiences to simultaneously view a phrase through multiple, mutually exclusive frames (Attardo, 1994; Young, 2008; Martin, 2002). Yet because of these characteristics, humor is often an effective way to identify group boundaries, punish deviant behavior, and designate the limits of acceptable conversation in an inclusive manner that allows even marginalized participants to save face (Goldstein & McGhee, 1972; Fine & de Soucey, 2005). Thus, while a novel and seemingly consistent feature of #overlyhonestmethods content, the specific influence of humor in these conversations may vary considerably depending on the identity of the speaker(s) and audience.

Although we will complete an empirical thematic and humor analysis of this hashtag later in this paper, several scholars have examined it previously. While reviewing 75 tweets the overlyhonestmethods hashtag, Stemwedel (2013) lays out the themes she can find in the hashtag's use: the winding, bumpy path of science, disagreement and uncertainty, challenges in determining what steps of research were necessary, difficulty in controlling scientific studies, budgetary and physical limitations, human limitations, and competitive pressures. But there was a single overriding theme in all of these tweets: the methods described in the tweets would not be described in a peer-reviewed journal article. In this hashtag, the methods described juxtapose the polished, clean version of scientific events more common to scientific journals and popular mass media with the messy, complicated events that occur in everyday scientific practice.

However, little is known systematically about how the hashtag has been used. Several scholars have expressed opinions on the topic (e.g. Riesch, 2015; Stemwedel, 2013; Bezuidenhout, 2015), but we are aware of no rigorous empirical analysis of the hashtag. It is therefore unclear, for example, how extensively subsequent tweets mirrored @dr_leigh's original posts in using humor, venting specific frustrations, or talking about laboratory activities. The lack of a broad examination of these posts is not altogether surprising; adequate (to say nothing of comprehensive) compilations of tweets on a given topic over an extended time period have often been difficult to come by. In this paper, we overcome these challenges by leveraging a new dataset of every tweet using #overlyhonestmethods posted between January 7, 2013, and January 6, 2016 in two studies. These data allow us to examine the source and content of tweets using this hashtag in an unprecedentedly precise manner. In Study 1, we consider the following three research questions:

RQ 1.1: *Who is using #overlyhonestmethods?*

RQ 1.2: *What are the major characteristics of the tweets that include #overlyhonestmethods?*

RQ 1.3: *How do these tweets use humor to shape their messages about science?*

In Study 2, we more deeply examine the use of humor with this hashtag.

Humor and #overlyhonestmethods

In the sparse academic literature – and the plentiful popular press literature – one thing is clear: the overlyhonestmethods hashtag is meant to be funny. Its creator meant for it to be a humorous way to discuss the "messy" side of science (Gould, 2013). Injecting humor into a public, critical conversation of scientific methodology, however, may have multiple, not always sanguine effects.

By definition, humor is often ambiguous and cognitively demanding. Indeed, prevailing theories of humor generally suggest that interpreting a joke requires that readers hold multiple, competing, mutually exclusive interpretations of a single statement (Raskin, 2012; Young, 2008). Thus, while the presence of humor appears able to entice otherwise uninterested individuals to attend to both serious and comedic messages on a given subject (what Baum (2005) refers to as a "gateway" effect), it may do little to directly foster long term recall of specific facts. Therefore, humor on its own likely has little lasting effect on mass scientific literacy (e.g. Fisher, 1998). Individual differences in humor preferences and interpretation can lead different groups of people to perceive the underlying, serious elements of the same joke in distinctly different ways. At the same time, however, science-based joking may have a significant effect on public views of and trust in scientists. Kim and Vishak (2008), for example, find that many individuals use online processing for humorous messages, updating their opinions of actors and concepts based on the humor they see even though they are unable to later recall the details of the jokes that influenced them.

Humor may have a more uniform influence on content creators, incentivizing comments that would be difficult to make any other way. When successful, humor offers a nearly universally positive experience for all involved (Martin, 2002). The pleasures of conversational joking are as well-documented in the professional literature as in common experience (e.g., Norrick, 1993). In addition to making oneself and others laugh, however, telling a joke can enhance an author's social standing and thereby allow them to make sensitive or aggressive comments that would otherwise be considered in poor taste. One leading humor theorist, Peter McGraw, argues that, in fact, humor is most often found in moderate challenges to the status quo, what he and his colleagues call "benign violations of social norms" (McGraw & Warren, 2010). Such violations are significant enough to challenge an audience's sensibilities and values, but not severe enough to simply be off-putting. Instead, humor occurs in a "sweet spot" where conventions are challenged, but in a manifestly safe way (McGraw, Williams, & Warren, 2013). Humor is thus a prime

vehicle for communicating behaviors and ideas that deviate, at times significantly, from accepted group standards (Meyer, 2000). Given the right psychological distance (most often generated by temporal, social, or physical space), almost any error in judgment or behavior can become funny (McGraw, Warren, Williams, & Leonard, 2012). The combination of humor's innate pleasure, ingroup elevation, and social cover makes it an attractive communication style for many people, including scientists in the context of the overlyhonestmethods hashtag.

Most uses of humor in public discussions of science, however, remain unexamined. Many scientists contend that humor can often be found in common scientific practice (Kilbourne, 1996; Lewin, 1983; Mulkay & Gilbert, 1982). As Mulkay and Gilbert (1982) note, this often occurs in violations of shared processes: "Humour, then, is not a characteristic of events in themselves, but is an outcome of the ways in which participants portray and organize their versions of events in the course of social interaction" (p.605). As members of the same group with shared knowledge about how science "should" work, scientists are well positioned to joke with each other (if not the broader public) by playfully suggesting violations of expected methodologies. This kind of comedy may also be augmented through, for instance, a non-serious use of technical language or concepts, contrasting otherwise serious ideas with a playful setting or frame (see Bateson, 1953).

Tweets including the #overlyhonestmethods hashtag may therefore be using humor in a number of ways. As indicated above, humor can serve as an ingroup signal, a shorthand for group members based on their shared knowledge and history (Martineau, 1972). For example, most academics recognize a violation of "good science" in a tweet that describes selecting findings to report by cycling through results until something is statistically significant at $p<0.05$. By contrast, other tweets may simply use seemingly nonsensical language or describe events that even non-scientists can appreciate. Physical comedy, for example, often has broad appeal (consider the lasting popularity of the *Three Stooges*); an accident in a lab that causes minor or moderate injury may therefore strike both scientists and the mass public as funny.

Humor also frequently constitutes a form of backstage talk (Watson, 2011; Goffman, 1959). Because of the need for a shared set of knowledge or values, humor is often used as coded way to condemn behavior or individuals in a group while allowing all involved to save face (Zadjman, 1995). The ease of conversing via Twitter, however, appears to have drawn into the public some of these conversations that once occurred privately, making them accessible to outgroup members who do not share an author's background or ability to appreciate the context and subtext of a comment. The audience may then be required to fill in many implied details, raising the risk of miscommunication and fostering negative perceptions of science.

In this study, we use the aforementioned Benign Violation Theory of humor (McGraw & Warren, 2010). Rooted in extant theories of the importance of incongruity in humor, the Benign Violation Theory focuses on the type of humor that is likely the most consequential and common in #overlyhonestmethods tweets: violations of "good science." In Study 2, we therefore ask two questions in particular:
RQ 2.1: *What types of humor are employed in this social media conversation? In other words, what type and severity are the violations present in these tweets and from whose perspective are they written?*
RQ 2.2: *How accessible and inclusive is the language used? In other words, how likely is a non-scientist member of the public to get the jokes made in #overlyhonestmethods tweets? How likely is it that the public understands the larger, serious commentary that underlies these jokes?*

Before addressing the humor-related questions, we conduct Study 1 in which, as previously discussed, we assess the users of #overlyhonestmethods, the characteristics of the tweets included in the body of #overlyhonestmethods posts, and how these tweets use humor to shape their messages about science.

Study 1: Methods

Over a three year time period from the inception of the hashtag, there are 58,125 relevant Twitter posts incorporating #overlyhonestmethods. After filtering all Twitter posts from January 7, 2013, to January 6, 2016, that use the overlyhonestmethods hashtag, we sorted posts into major themes within a coding schema, which an intelligent algorithm then applied to the census of tweets. The coding schema was developed by the research team prior to involving machine learning.

To perform this coding sample extraction and subsequent analysis, we relied on a supervised machine learning-based content analysis tool provided by a commercial company Crimson Hexagon. This

tool is built on a combination of human- and computer-based coding that relies on human coders' ability to discern the underlying concepts embedded in texts and the computer's capacity to reliably and efficiently process a large amount of data. Through the use of Crimson Hexagon ForSight, researchers have the access to examine online content from millions of websites such as Facebook posts, full Twitter Firehose, and online news sites.

Researchers first specify the keywords, coding categories, and social media platform (i.e. Twitter in our study) that are related to the research of interest and assign time period for analysis. The software randomly selects posts that contain the keywords (in our case, #overlyhonestmethods) within the specific time range (January 7, 2013 to January 6, 2016) for further analysis. The detailed coding schema (explained below) is specified in advance so human coders can rely on it for further sample coding. During the coding process, human coders first analyze and categorize a sample of randomly selected Twitter posts. All coders must unanimously agree on the placement of each tweet into a particular code. Once every coding category is assigned the optimal number of tweets (20), the computational algorithm then extracts linguistic patterns from the categorized posts and systematically applies the learned patterns to analyze the population of textual data (Hopkins & King, 2010).

We created an exhaustive and mutually exclusive coding schema that includes eight themes: *personal reasons/interference*, *accidents*, *research constraints*, *literature*, *ignorance*, *statistics*, *scientific process*, *meta-commentary*, and *off-topic* (Table 1). *Personal reasons/interference* refers to posts that mention (internalized) reasons within a researcher's control that methods were not ideal. *Accidents* includes posts that mention accidents that the researcher incorporated into methodology. *Research constraints* refers to posts that mention (externalized) research related reasons why the methods presented in the tweet are too honest, e.g., small sample size. *Literature* refers to posts that mention the literature research and writing aspect of the scientific process. This refers to the actual writing, collecting of research, or citing. *Ignorance* refers to posts that admit ignorance in the scientific process, whether alluding to human forgetfulness or not knowing something about their research subject. *Statistics* includes posts that mention statistical analyses. *Scientific process* refers to posts that offer commentary on the scientific process, including mention of peer-review. Finally, *Meta-commentary* refers to posts that comment on the hilarity, utility, and general understanding of the #overlyhonestmethods hashtag itself. *Off-topic* tweets were those tweets in a foreign language or unrelated to the hashtag.

Study 1: Results and implications

Contribution to #overlyhonestmethods started with a single post on January 7, 2013, and spiked quickly (Figure 1). The number of unique Twitter authors peaked early; however, posts per author is steady, which indicates little to no spamming of the hashtag (Figure 2). Half of the top ten most prolific posters are scientists (Figure 3), while about 15% of the 50 most influential posters, as indicated by Klout scores, are scientists. Almost a third of the most influential posts were science media outlets.

The largest subgroup of posts, nearly one-quarter, fit into the *meta* category, which are users talking about the hashtag itself (Figure 4). Other than that, the *scientific process* and *research constraints* are the largest thematic categories, with 18% and 14% respectively. *Statistics* and *literature* are each themes that represent 11-12% of the total posts, and *personal interference*, *accidents*, and *ignorance* are all each 7-8% of the total.

These results help us understand the discussion occurring with this hashtag. Twitter users are not necessarily utilizing this hashtag to disclose mistakes they are making or ways they are trying to interfere with the scientific process. Instead, they are talking about the larger scientific process and how that may differ from the standard view of the scientific process. They are also describing external forces that affect their interactions with science, such as funding constraints and publication. To some extent, they are also discussing how their research is presented statistically.

It is clear, however, that this hashtag is presenting new and different information into the public sphere. Although it is beyond the scope of this study to determine the intended audience(s) for this hashtag is, one thing is clear: this conversation is taking place in public and is showing the public a side of science to which they may not otherwise have exposure.

Study 2: Methods

Because coding for humor with an intelligent algorithm is difficult, two coders manually coded a randomly selected 2 percent of tweets (n = 1167). Inter-rater reliability, measured by Scott's Pi, was at least

0.70 for each coded element. Consistent with the Benign Violation Theory of humor, we coded each tweet on three dimensions: type of violation, target of joke, and language.

We first identified the presence of three types of humorous violations of traditional or expected scientific practice. The three types of violations that we coded for are *minor, physical*, and *ethical*. Violations were considered *minor* if they appeared to pose no significant threat to either researchers' well-being or the project's results. These incidents were often characterized as accidents that were the unintended side-effects of otherwise good science and included minor memory lapses such as forgetting the names of spreadsheets. Note that here, as with the other violation codes we use, we examine only the language of each tweet and not the full set of consequences that may have occurred because of the incident described. In other words, seemingly minor violations were coded as such even if they may have led to events that warranted another code. We defined a *physical* violation as any language indicating events wherein an individual (including the author) was or could have been seriously harmed. Some tweets, for example, mention an individual in the lab passing out after inhaling toxic fumes. By contrast, rhetoric was coded as an *ethical* violation if it included a significant or perhaps intentional methodological breach that put the value of the related project's outcome in question. Language of this type includes fabricating data or running statistical models a large number of times until they provide a statistically significant result.

We next coded for the target of the humor in each tweet. Comedy often occurs at the expense of one or more individuals, groups, or institutions - the "butt" of a joke. We identified all tweets wherein the humor came at the expense of one or more of the following seven types of individuals: *tweet author*, *peer co-investigator*, *subordinate co-investigator*, *supervising co-investigator*, *journal reviewer/editor*, *university administration*, and *research subjects*. In practice, this meant that we identified the actor who was "blamed" for the violation at the heart of the humor. Often this was a scientist in a laboratory setting.

Many tweets were self-deprecatory (coded as *tweet author*), such as the following example: "pH was left at 5.486 instead of 5.5 bc only 6M NaOH was available & I didn't want to have to dilute that. #overlyhonestmethods." Other tweets implicitly targeted a *peer co-investigator*, as in: "RT @[redacted] We recorded in the hippocampus because it's big and we didn't know where else to look. #overlyhonestmethods." Language of this type indicates that at least some methodological decisions were made not because they would result in the most reliable and accurate outcome, but because they more convenient for one or more researchers involved. All tweets that included a joke made at the expense of an otherwise unidentified other are classified as *peer co-investigator*. Tweets were coded as targeting a *subordinate co-investigator* if they referred to undergraduate lab assistants or students. For example, one tweet mocked an undergraduate by presenting the following excerpt from an "actual student paper": "gender comparisons were made because a 2x2 factorial anova requires a 2nd variable." *Supervising co-investigator* codes were assigned to tweets that joked at the expense of a person in a supervisory position: "RT @[redacted] This method was used because my supervisor told me so when I started my PhD. #overlyhonestmethods." By contrast, the code *journal reviewer/editor* was applied to jokes that in some way targeted reviewers and editors involved in the peer review and publishing process. In this fashion, one tweet author claimed to "love to explicitly thank big names for their comments on [his or her] draft mainly to deter reviewers from commenting harshly #overlyhonestmethods," and by so doing poked fun at the reputation of reviewers. In the same manner, tweets coded as *university administration* or *research subjects* ridicule either non-research university faculty/staff or human research subjects. An example of *research subjects* as the butt of jokes is the following tweet: "RT @[redacted] More #overlyhonestmethods I've used: Late season gull chick data were unavailable because the bald eagle family on the island ate them."

Exclusivity of language was determined by three codes—*academic specific jargon*, *scientific jargon*, and *generally accessible language*—and by evaluating pronoun usage and sentence structure. Academic specific jargon includes language that would likely be foreign for readers who have not taken statistical or scientific methodology courses at a university. Jargon of this type includes mention of things like statistical tests, p-values, or methodologies that are widely known and used in academic research but are not recognizable to most lay audiences. Similarly, scientific jargon included language recognizable to science practitioners but not the broader public or even individuals in other areas of academia--for instance, western blots. Finally, we considered as generally accessible language anything that did not fall into one of these two "jargon" categories.

Noting the importance of psychological distance in creating humor, we also recorded two ways of pronoun and sentence structure used in each tweet. First, we identified all tweets that inserted a plural or non-first person pronoun (e.g., we, they, you) when a singular, first person pronoun could have been used. Language of this type is often an effective way of creating the psychological distance necessary to ensure that a violation of expected norms will be perceived as benign enough to be humorous (see, McGraw, Warren, Williams, & Leonard, 2012). Second, we also coded for language written in the passive voice. Non-active writing of this type is stereotypically common in academic writing and can be a barrier to a broader audience. If we could not identify a joke in the tweet, we did not analyze the tweet for humor. Similarly, all tweets that were off-topic were not coded for humor.

Study 2: Results and Implications

The results of Study 2 immediately suggest three things about scientists' dialogue in #overlyhonestmethods tweets. First, researchers rarely tweeted about potentially serious behavioral violations. More than three out of every four violations in our data were minor, not physical or ethical (Figure 5). This indicates that many scientists' goal in #overlyhonestmethods tweets was not solely to be comedic. We may expect individuals whose principal desire is to be funny online (i.e., writing to an audience likely at a significant spatial and temporal distance from themselves) to frequently draw on more serious behavioral violations or actions that are more broadly considered inherently funny (e.g., someone hurting themselves). Instead, most #overlyhonestmethods jokes were subtler and focused on minor infractions of the scientific method that may not be apparent to the lay public. In this manner, #overlyhonestmethods tweets appear to offer a public space in which practitioners can discuss the everyday practices that scientists engage in but that may not technically be part of the scientific method. Indeed, nearly 30 percent of all #overlyhonestmethods tweets we coded included no humor attempt (i.e., no norm violation) and often focused instead on promoting #overlyhonestmethods tweets to a broader audience of scientists. The tweets in our data therefore publicly highlight the unavoidable messiness of science done by human beings, no matter its variance with public beliefs or desires about how scientific research is conducted.

Still, fully 14 percent of the tweets we coded indicated some type of more serious violation of scientific ethics. While many of the specifics and the motivations behind the actions described in these tweets remains unclear, the fact that a non-negligible portion of our sample discusses this potentially troubling behavior should not be overlooked. The fact that so many authors were willing to share these behaviors suggests a perception that the behavior is relatively common and that the hashtag was primarily viewed by a sympathetic audience--that is to say, other scientists. It is unlikely that a scientist would want to publicly identify himself or herself as a methodological deviant and would instead post behavioral violations only if they were relatively sure the behavior was reasonably common and would not result in sanction. Many scientists have experienced pressure to obtain results, secure funding, or complete a degree or publication but previously may have been able to frankly discuss such pressures and their reaction to them in only private, face-to-face, small group settings. The #overlyhonestmethods tweets may therefore show a side of science that is rarely seen in public except after scandals (e.g., "We avoided experimenter bias by deciding what the results would be in advance of the experiment #overlyhonestmethods"). The potential loss of status and esteem from both the public and one's peers may be substantial in these tweets and raises questions about why scientists were willing to engage in communication so potentially fraught. What are the benefits that outweigh the potential downsides of such messages? Why do scientists feel comfortable sharing these unethical behaviors through social media instead of through other channels, such as with university administration?

Our second major takeaway from this study is that most scientists appear to have been more interested in commenting about laboratory procedure than criticizing figures or institutions of power. Overwhelmingly, the targets of jokes were the tweet author and peers (Figure 6). This suggests that this space has been used primarily to talk about "funny" things that happen to oneself or one's colleagues and not to vent about advisors, supervisors, or administrators. Instead, authors largely talk about themselves: what they are actively doing and how science really looks for them and their peers. University administration and research subjects were essentially absent from the sample.

Finally, the third major takeaway of our second study is that most #overlyhonestmethods tweets used accessible, if impersonal, language. Approximately ten percent of posts used *academic research*

jargon, and 15.7 percent of posts used *scientific jargon*, leaving about one in four tweets rhetorically inaccessible for the non-practitioner community. This suggests that humor is being used here as part of an exclusive, ingroup conversation for scientists to discuss their field and practice with a candor that accounts little for wider understanding. At the same time, however, more than 60 percent of #overlyhonestmethods tweets that include humor use broadly accessible language with relatively few references to technical scientific concepts or practices. A practical consequence of such language is to make #overlyhonestmethods tweets more understandable by members of the public, increasing the quantity of scientific content in public communication channels. Nevertheless, as we will explore below, the impact of scientific content like #overlyhonestmethods is unlikely to be consistently positive.

Scientific and academic jargon were somewhat more common in posts discussing ethical violations. This is likely due at least in part to the more complex nature of the violations depicted in these tweets. Yet it may also indicate a propensity among scientists to strive for greater specificity when discussing more serious violations. Comedy is inherently ambiguous and can leave its audience with multiple unintended interpretations that may include more worrisome meanings than an author intended. Scientists may have used more specific, jargon-infused language in order to stave off misunderstandings and incorrect extrapolations. The combination of foreign language and serious violations could be especially off-putting for many members of the public, however. Scientists who are speaking over the public's head on ethical matters could have serious implications for public trust in science--we are less likely to trust what we do not understand.

Similarly, most tweets used language that obfuscated responsibility for violations. Most, for example, used plural pronouns; fewer tweets included first person singular pronouns. "We" was the most common pronoun in our sample, appearing in 33.6 percent of total posts while "I" was used in 23.4 percent. It seems that most scientist-authors in our sample had little interest in incriminating their colleagues and therefore blurred lines of responsibility. Many authors also used passive voice to accomplish similar ends, with 34 percent of posts including language of this type. By shifting focus away from specific actors, content in passive voice makes it difficult to make direct causal attributions. This could be a play on scientific writing, recognizing the prevalence of passively voiced writing in scientific writing, creating a playful juxtaposition. It could also signal an unwillingness to connect specific other actors with the actions a given tweet makes light of.

Discussion

The patterns of public science writing we have identified connected to this hashtag raise several concerns and questions. The overlyhonestmethods hashtag was originally created and subsequently used to discuss the "messy" side of science (Gould, 2013) not typically available to the public. While it remains unclear if any of these tweets were composed explicitly for consumption by the mass public, their potential impact on public knowledge and trust of science is compelling. At the least, these tweets appear to spur a public conversation among scientists about the realities, intended and not, of their work. Our study found that the most of #overlyhonestmethods tweets were not intended to highlight methodologies that were intentionally being obscured by scientists in the typical academic publication process. Scientists using this hashtag may be trying to shed light on what science actually looks like, which is very different from the tidy, clean results we read about in peer-reviewed journals or that is discussed in mainstream media. Scientists likely have similar problems and challenges and discovered this is a space to discuss them.

One important outcome of this analysis is that evidence shows this hashtag is *not* necessarily meant to spark a public discussion (in so far as public discussion means discussion with the public, or lay audiences), as about 25% of posts used academic and/or scientific jargon. However, this conversation is nonetheless occurring in the public arena. This could have several implications for the public's trust in science. The public may value this honesty and see it as a humanization of scientists—that science is done by actual people, and the hashtag may then simply show more honest communication about the actual process of science and scientific knowledge production. Conversely, the public could also become worried about this honesty, since they were likely previously unaware of what goes on behind closed doors and might view the messy side of science as a reason not to trust scientists and science. The use of academic language and scientific jargon could also on its face turn audiences away from science. Although the conversation is occurring on Twitter in a more accessible space, if the language is not accessible to the lay

Twitter user, they could be turned away from both the tweet and science more generally. Essentially, they could feel left out of the joke, which may create feelings of insecurity, defensiveness, and negativity.

Conclusion

Contributors to the #overlyhonestmethods conversation are having an insiders' conversation in a public space. In this research, we explore the ethical implications of having such a conversation, while detailing exactly the specifics of the conversation and how humor is employed. Scientists, writers, science enthusiasts, media and educational institutions, and parody accounts all contributed to three years of an "overly honest" discussion of scientific methods. In this exploration of what makes these contributions "overly honest," and why and how tweets are funny, we find that posters are talking about the nitty gritty details of how science works (the *scientific process*) and how it doesn't (*research constraints*). Most benign violations that occur seem to be *minor*, rather than *ethical* or *physical*, which indicates posters are referencing subtle methodological problems. Given this knowledge and the fact that 75% of the language used is generally accessible, lay audiences are still likely on the outskirts of the conversation that contributors think they are having. Additionally, jokes are geared toward the tweet authors (presumably scientists) themselves and their peers, rather than punching up at, for example, University administration.

The #overlyhonestmethods conversation is one example of several ongoing and surely several emerging trends of scientists taking previously private conversations to the public sphere, all of which might be affecting public trust in science, regardless of contributors' intentions or enjoyment of the trend. Our analysis offers insight into what happens when social media phenomenon like #overlyhonestmethods puts science in a fishbowl—when science itself is in a position to potentially be the center of scrutiny in a way that public audiences do not often see it.

References

Attardo, S. (1994). *Linguistic Theories of Humor*. New York: Walter de Gruyter.

Bateson, G. (1953). The position of humor in human communication. In H. von Foerster (Ed.), *Cybermetics, Ninth Conference* (1-47). New York: Josiah Macey, Jr. Foundation.

Baum, M.A. (2005). *Soft News Goes to War: Public Opinion and American Foreign Policy in the New Media Age*. Princeton, NJ: Princeton University Press.

Bezuidenhout, L. (2015). Variations in scientific data production: What can we learn from #Overlyhonestmethods? *Science and Engineering Ethics, 21,* 1509-1523.

Brenner, J. (2014). Can Twitter survive in a Facebook world? The key is being different. *Pew Research Center*. Retrieved from http://pewrsr.ch/1l9BmE1.

Fine, G.A. & de Soucey, M. (2005). Joking cultures: Humor themes as social regulation in group life. *Humor - International Journal of Humor Research*, *18*(1): 1-22.

Finfgeld-Connett, D. (2015). Twitter and health science research. *Western Journal of Nursing Research*, *37*(10): 1269-1283. DOI: 10.1177/0193945914565056

Fisher, M. (1998). The effect of humor on learning in a planetarium. *Science Education*, *81*(6), 703-713.

Goffman, E. (1959). *The presentation of self in everyday life*. Garden City, N.J.: Anchor-Doubleday.

Goldstein, J.H. & McGhee, P.E. (1972). *The psychology of humor: Theoretical perspectives and empirical issues.* Academic Press: New York and London.

Gould, J. (2013). Special speaking to...Dr Leigh and #overlyhonestmethods. *Speaking of Science*. Retrieved from http://speakingofscience.juliegould.net/science-communication/special-speaking-to-dr-leigh-about-overlyhonestmethods/.

Grenoble, R. (Jan. 8, 2013). #OverlyHonestMethods captures all the hilarious science not fit for publication. *Huffington Post*. Retrieved from http://www.huffingtonpost.com/2013/01/08/overly-honest-methods-twitter_n_2435364.html.

Hopkins, D. J., & King, G. (2010). A method of automated nonparametric content analysis for social science. *American Journal of Political Science, 54*(1), 229-247.

Hurley, M.M., Dennett, D.C., & Adams, R.B. (2013). *Inside jokes: Using humor to reverse-engineer the mind*. Cambridge, MA: MIT Press, .

Kilbourne, E.D. (1996). Humor in science. *Proceedings of the American Philosophical Society, 140*(3), 338-349.

Kim, Y.M., and Vishak, J. (2008). "Just laugh! You don't need to remember: The effects of entertainment media on political information acquisition and information processing in political judgment. *Journal of Communication*, *58*(2), 338-360.

Lewin, R.A. (1983). Humor in the scientific literature. *BioScience, 3*(4), 266-268.

Liang, X., Yi-Fan Su, L., Yeo, S.K., Scheufele, D.A., Brossard, D., Xenos, M., Nealey, P. & Corley, E.A. (2014). Building buzz: (Scientists) communicating science in new media environments. J*ournalism & Mass Communication Quarterly, 91*(4), 772-791.

Martin, R.A. (2002). *The Psychology of Humor: An Integrative Approach*. Cambridge, MA: Academic Press.

Martineau, W. H. (1972). A model of the social functions of humor. *The psychology of humor: Theoretical perspectives and empirical issues*, 101-125.

McGraw, P. & Warren, C. (2010). Benign violations: Making immoral behavior funny. *Psychological Science, 21*(8), 1141-9.

McGraw, P., Warren, C., Williams, L.E., & Leonard B. (2012). Too close for comfort, or too far to care? Finding humor in distant tragedies and close mishaps. *Psychological Science*, 23(10), 1215-1223.

McGraw, P., Williams, L.E., & Warren, C. (2013). The rise and fall of humor: psychological distance modulates humorous response to tragedy." *Social Psychological & Personality Science, 5*(5), 566-72.

Medawar, P.B. (1964). Is the scientific paper fraudulent? Yes: It misrepresents scientific thought. *The Saturday Review*, 42-43.

Meyer, J.C. (2000). Humor as a double-edged sword: Four functions of humor in communication. *Communication Theory*, *10*(3), 310-331.

Mulkay, M. & Gilbert, G.N. (1982). Joking apart: Some recommendations concerning the analysis of scientific culture. *Social Studies of Science, 12*(4), 585-613.

National Science Board. (2014). Science and Engineering Indicators 2014. *National Science Foundation*. Retrieved from http://www.nsf.gov/statistics/seind14/index.cfm/chapter-7/c7h.htm.

Neiger, B. L., Thackeray, R., Burton, S. H., Giraud-Carrier, C. G., & Fagen, M. C. (2013). Evaluating social media's capacity to develop engaged audiences in health promotion settings: Use of Twitter metrics as a case study. *Health Promotion Practice, 14*, 157-162. DOI:10.1177/1524839912469378

Norrick, N.R. (1993). *Conversational Joking: Humor in Everyday Talk*. Bloomington, IN: Indiana University Press.

Rainie, L., Funk, C., & Anderson, M. (2015). How scientists engage the public. *Pew Research Center.* Retrieved from http://www.pewinternet.org/2015/02/15/how-scientists-engage-public/.

Raskin, V. (2012). *Semantic Mechanisms of Humor*. New York: Springer Science and Business Media.

Ravetz, J., & Saltelli, A. (2015). Policy: The future of public trust in science. *Nature, 524*(7564), 161-161. doi:10.1038/524161d

Riesch, H. (2015). Why did the proton cross the road? Humour and science communication. *Public Understanding of Science, 24*(7), 768-775.

Stemwedel, J.D. (2013). #overlyhonestmethods: Ethical implications when scientists joke with each other on public social media. In Jean Goodwin, Michael F. Dahlstrom & Susanna Priest (eds.), *Ethical issues in science communication: A theory-based approach* (pp. 287-298).

Sturgis, P. & Allum, N. (2004). Science in society: Re-evaluating the deficit model of public attitudes. *Public Understanding of Science, 13*(1), 55-74.

Watson, K. (2011). Gallows humor in medicine. *The Hastings Center Report, 41*(5), 37-45.

You, J. (2014). The top 50 science stars of Twitter. *Science Magazine.* Retrieved from__ http://www.sciencemag.org/news/2014/09/top-50-science-stars-twitter.

Yarborough, M. (2014). Openness in science is key to keeping public trust. *Nature, 515*(7527), 313-313. doi:10.1038/515313a

Young, D.G. (2008). The privileged role of the late-night joke: Exploring humor's role in disrupting argument scrutiny. *Media Psychology*, *11*(1), 119-142.

Zadjman, A. (1995). Humorous face threatening acts: Humor as strategy. *Journal of Pragmatics*, 23, 325-39.

Figures and Tables]

Tab. 1: Coding schema for Study 1. Note that "off-topic" includes any post in a language that is not English, as well as any non-nonsense posts (strings of letters, for example).

Code	Definition	Example
Personal reasons/interference	refers to posts that mention (internalized) reasons within a researcher's control that methods were not ideal	Participants in the 9:00 hour only received the control condition because my nails were wet. #overlyhonestmethods
Accidents	refers to posts that mention accidents that the researcher incorporated into methodology	Plasma samples were inverted three times, then rotated 90° and vibrated, in a manner similar to knocking the tray over. #overlyhonestmethods
Research constraints	refers to posts that mention (externalized) research related reasons why methods are being presented as too honest, e.g., small sample size	I use X different resampling methods and apply tests that produce colourful graphs, because my sample was too small #overlyhonestmethods
Literature	refers to posts that mention the literature research and writing aspect of the scientific process. This refers to the actual writing, collecting of research, or citing.	When I say "research indicates" I mean I read an abstract on pubmed because the full paper was behind a paywall #overlyhonestmethods
Ignorance	refers to posts that admit ignorance in the scientific process, whether alluding to human forgetfulness or not knowing something about their research subject	We don't know how the results were obtained. The postdoc who did all the work has since left to start a bakery. #overlyhonestmethods
Statistics	refers to posts that mention statistical analyses	"If at first you don't succeed, try and fail two more times so your failure is statistically significant." #overlyhonestmethods
Scientific process	refers to posts that offer commentary on the scientific process. Includes mention of peer-review	RT @[redacted] This is ALWAYS the case. RT @[redacted]: Coffee did more for this project than the Principal Investigator. #overlyhonestmethods

Meta	refers to posts that comment on the hilarity/utility/etc. of the #overlyhonestmethods hashtag itself	RT @[redacted]: I really think that #overlyhonestmethods is one of the most wonderful pieces of science communication I've seen in a while.

Fig. 1: The volume of #overlyhonestmethods posts over time, from early 2013 to early 2016

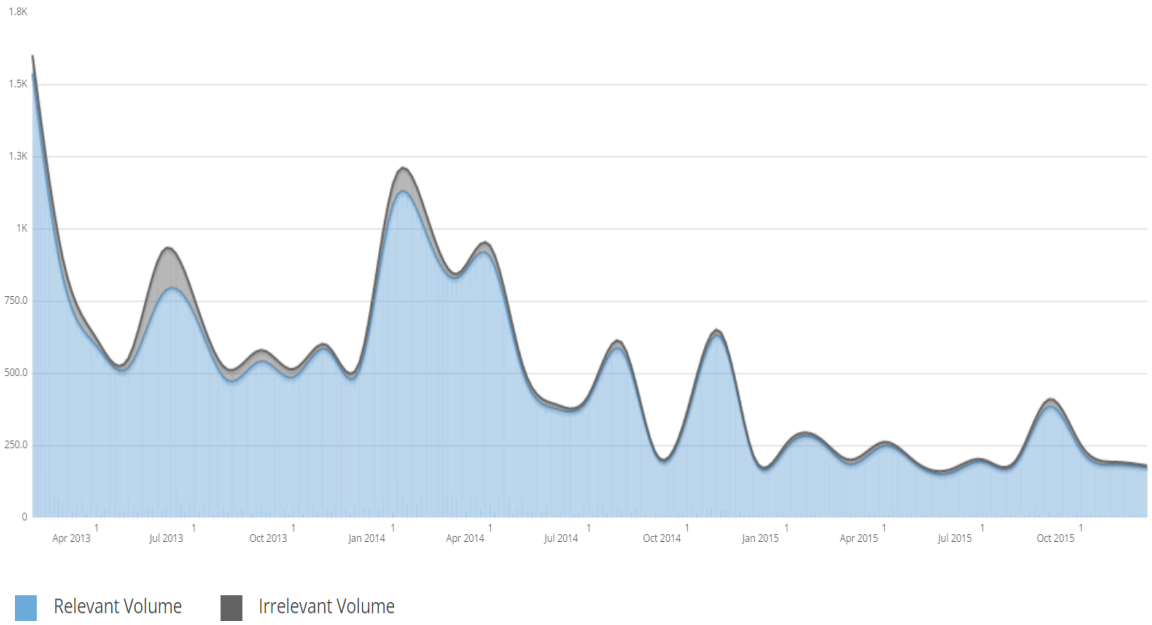

Fig. 2: The number of daily Twitter authors of #overlyhonestmethods posts and the average number of #overlyhonestmethods posts/author over a three year period

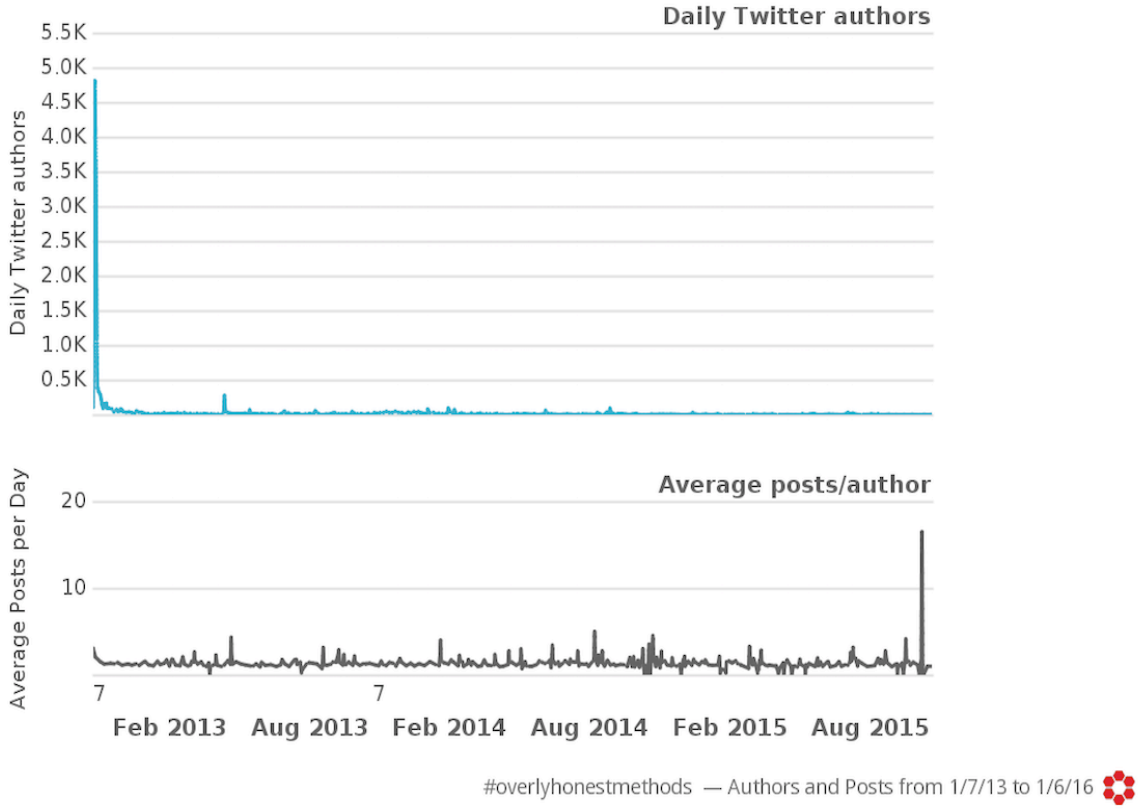

Fig. 3: The top 10 contributors to the #overlyhonestmethods conversation between January 7, 2013, and January 6, 2016

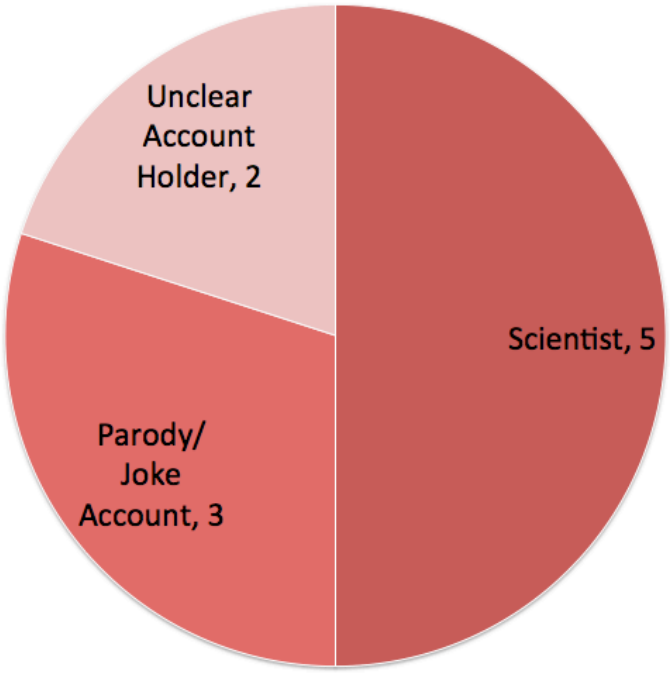

Fig. 4: Breakdown of themes in Study 1 analysis

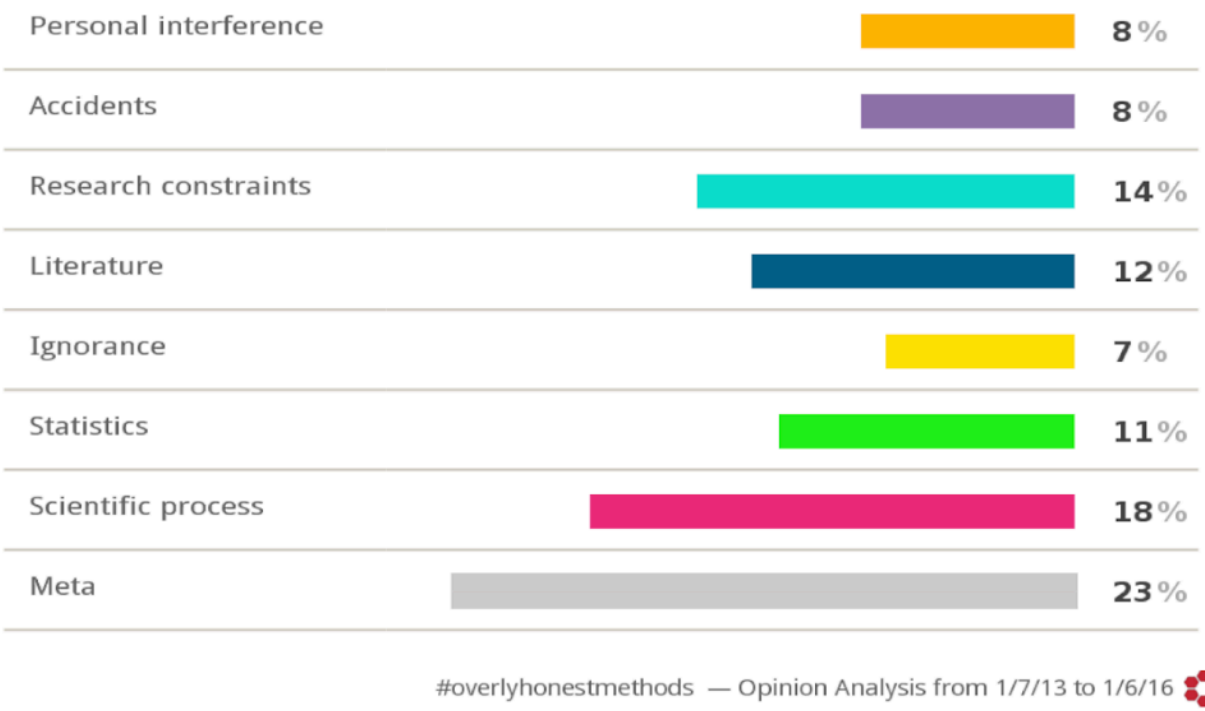

Fig. 5: Breakdown of violation type in sample of #overlyhonestmethods posts from Jan. 7, 2013-Jan. 6, 2016 (n = 1117)

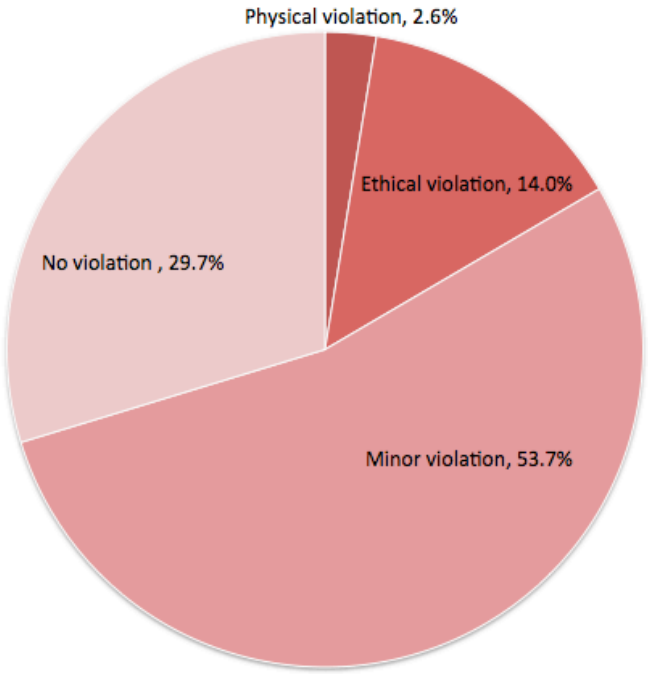

Fig. 6: Targets of jokes in sample of #overlyhonestmethods posts from Jan. 7, 2013-Jan. 6, 2016 (n = 1117)