

The Evolution of Substance Use Coverage in the Philadelphia Inquirer

Layla Bouzoubaa,¹ Ramtin Ehsani,¹ Preetha Chatterjee,¹ Rezvaneh Rezapour¹

¹ College of Computing and Informatics, Drexel University, Philadelphia, USA
{lb3338, re378, preethac, shadi.rezapour}@drexel.edu

Abstract

The media’s representation of illicit substance use can lead to harmful stereotypes and stigmatization for individuals struggling with addiction, ultimately influencing public perception, policy, and public health outcomes. To explore how the discourse and coverage of illicit drug use changed over time, this study analyzes 157,476 articles published in the Philadelphia Inquirer over a decade. Specifically, the study focuses on articles that mentioned at least one commonly abused substance, resulting in a sample of 3,903 articles. Our analysis shows that cannabis and narcotics are the most frequently discussed classes of drugs. Hallucinogenic drugs are portrayed more positively than other categories, whereas narcotics are portrayed the most negatively. Our research aims to highlight the need for accurate and inclusive portrayals of substance use and addiction in the media.

Introduction

The portrayal of illicit substance use in the media has long been a topic of concern, given the potential impact it can have on public perception and policy. Previous studies highlighted the harmful effects of stigmatizing coverage on individuals struggling with substance use disorder (SUD) and the potential for biased and inaccurate media coverage to perpetuate harmful stereotypes and stigma (Johnston, O’malley, and Bachman 2004; Caburnay et al. 2003; Hughes, Lancaster, and Spicer 2011). Moreover, media coverage often prioritizes criminal justice actions over critical health concerns associated with substance use (Hayden Griffin et al. 2023). This bias has been further observed in the case of designer drugs such as “bath salts,” which were portrayed as an “epidemic” with harmful effects while neglecting relevant clinical research and mental health issues associated with them (Swalve and DeFoster 2016). As a result, health legislation was inadequately informed, and underlying issues related to substance use are not addressed.

Given the potential impact of media coverage, it is essential to carefully examine how the media frames discussions about illicit drug use and how that framing changes over time. In this study, we examine news articles from the Philadelphia Inquirer to understand changes in media coverage of illicit substance use over time. Philadelphia was cho-

sen because it has a high rate of overdose deaths (40 per 100K) and is home to a significant open-air drug market (Niamatullah, Auchincloss, and Livengood 2022). Analyzing local media coverage of SUD can reveal insights on its portrayal, identify reporting gaps, and inform policymaking and public education. This type of study can also track the policy process and evolving relationships between actors, like advocacy groups or healthcare professionals, over time (Paalman 1997). More specifically, we analyze the portrayal of nine main drug classes, including *stimulants*, *narcotics*, *cannabis*, *hallucinogens*, *depressants*, *designer drugs*, *drugs of concern*, *treatment*, and *miscellaneous* over a period of ten years and across 157,476 published articles. The study then focuses on approximately 3,903 articles that mention at least one commonly abused substance and applies various natural language processing (NLP) techniques to extract drug-related themes over time. We aim to answer the following research questions:

- RQ1: How do the occurrences of various drug classes change over time, and in what contexts are different substances co-mentioned?

We use a list of commonly abused drugs, map them to drug classes, extract articles with at least one mention of drugs, and use TF-IDF to find salient words in the subsets of articles discussing co-occurring drug classes.

- RQ2: How has media coverage of illicit substance use evolved over time?

Dynamic topic modeling and sentiment analysis are employed to capture the shift of prevalent themes and tone with respect to drug classes, over time.

Our analysis indicates a shift in the discussion of drug-related topics, with cannabis being the most frequently discussed drug class, followed by narcotics. Articles on narcotics and treatment drugs frequently co-occur, focusing primarily on criminality, legislature, and overdose. News articles discussing cannabis, hallucinogens, and treatment drugs tend to be positive while those about narcotics, stimulants, and depressants are negative. The insights gained from this work can inform better public health messaging, affect drug policy decisions made by lawmakers, and enable medical professionals to create more targeted patient education and awareness programs while promoting the principles of harm reduction.

Methodology

Data Collection We utilize *ProQuest* for data collection and accessing news articles. *ProQuest* is an online platform consisting of thousands of databases that provide access to a diverse set of publications, including journals and newspapers. In order to collect data, it was necessary to log in to the *ProQuest* database using our individualized credentials, after which we were able to extract/scrape the data. We collect a total of 157,476 news articles from the Philadelphia Inquirer, covering the period from January 1, 2013, to December 31, 2022. It’s important to note that our dataset may not represent the complete set of articles published by the Philadelphia Inquirer, as *ProQuest*’s database may not be comprehensive. Our data comprises the complete text of news articles, alongside several metadata attributes such as date, author, title, links, and subject keywords.¹

Drug Name Extraction We utilize the categorization provided by the *National Institute on Drug Abuse* (NIDA), which comprises 28 drug categories, each with its corresponding commercial names. To minimize irrelevant articles, we manually verified drug names and removed ambiguous terms with multiple meanings, e.g., the term “pot” could refer to “flowering pot” or “cooking pot” in addition to the drug “marijuana”. We then assign each drug to one of the nine classes established by the Drug Enforcement Administration (DEA) (DEA 2020). These classes include:

- *Cannabis (C)*: marijuana is a mind-altering (psychoactive) drug, produced by the cannabis sativa plant. “Cannabis” and “marijuana” are often used interchangeably, resulting in the appearance of cannabis as both the drug class and the drug name.
- *Depressants (D)*: known to induce sleep, relieve anxiety and muscle spasms, and prevent seizures, e.g., barbiturates, and sedative-hypnotic substances like GHB.
- *Designer Drugs (DD)*: produced illicitly with a slightly altered chemical structure to mimic the pharmacological effects of controlled substances, e.g., synthetic marijuana or synthetic cathinones.
- *Drugs of Concern (DC)*: unregulated drugs that can be harmful if abused, e.g., kratom and xylazine.
- *Hallucinogens (H)*: derived from plants and fungi and renowned for their capacity to modify human perception and mood, e.g., LSD, mushrooms, and ecstasy.
- *Narcotics (N)*: refers to opium, opium derivatives, and their semi-synthetic substitutes. “Opioid” is a more current and precise term to describe these drugs, e.g., heroin, OxyContin, codeine, morphine, and fentanyl.
- *Stimulants (S)*: drugs accelerating body’s functions, e.g., methamphetamine, cocaine, and amphetamines.
- *Treatment (T)*: substances aiding the treatment of opioid addiction, e.g., methadone, Suboxone, and naloxone.
- *Miscellaneous (M)*: substances that can be abused but don’t belong to any classes, e.g., steroids.

¹The meta-data of news articles, list of drugs, and code can be found on https://github.com/social-nlp-lab/drugs_in_inquirer

	C	D	DD	DC	H	N	S	T	Total
2013	152	23	3	2	38	142	108	16	484
2014	150	10	0	1	53	117	64	21	416
2015	132	6	0	0	35	117	49	30	369
2016	152	12	0	0	44	132	63	43	446
2017	203	12	0	0	37	223	53	78	606
2018	226	14	8	6	43	228	73	98	696
2019	234	10	4	2	55	148	90	48	591
2020	119	7	0	0	18	69	48	31	292
2021	124	7	0	0	40	86	41	41	339
2022	145	5	1	0	31	99	38	43	362
Total	1637	106	16	11	394	1361	627	449	

Table 1: Number of drug class occurrences per year, with bold values indicating the most frequent drug class per year. Drug classes include C (Cannabis), D (Depressants), DD (Designer Drugs), DC (Drugs of Concern), H (Hallucinogens), N (Narcotics), S (Stimulants), and T (Treatment).

Using our mapped list of drug names, we identify news articles that mention any of these drugs in either their title or body. Subsequently, we group these articles into clusters based on the respective drug classes. If an article mentions drugs from multiple classes, we assign it to all relevant drug classes. The number of articles per drug class and year is shown in Table 1. Out of a total of 157,476 articles analyzed, 3,903 referenced at least one drug class within their text.

In addition to assessing the frequency of drugs, we also examine co-occurrences of drugs in the same articles each year. To better understand these co-occurrences and the surrounding contexts, we extract the most significant words based on their TF-IDF scores from a subset of articles where drugs most frequently co-occurred.

Topic Modeling To investigate the evolution of drug-related articles, we employ dynamic topic modeling (Blei and Lafferty 2006) to understand the distribution of topics over time. We use BERTopic as it can preserve the semantic structure of texts (Grootendorst 2022). Since BERTopic operates on a bag-of-words representation of the text (Grootendorst 2022), we preprocess the text by converting it to lowercase, removing URLs, extending contractions, and eliminating stopwords. This minimizes noise and dimensionality, allowing our models to recognize important patterns and topics. It is important to note that BERT has a token limit of 512, while the median length of news articles in our data is 749 tokens with a standard deviation of 573. News articles’ opening paragraphs contain the most important information, therefore, we assume the first 512 tokens of news articles in our data represent the key points.

Tone Detection VADER (Hutto and Gilbert 2014), a validated rule-based model for sentiment analysis, is used to analyze shifts in the tone of news articles that center on substance use. We calculate the average sentiment scores of articles in six-month periods using VADER’s compound score, allowing us to monitor changes in sentiment over time.

Results

How do the occurrences of various drug classes change over time, and in what contexts are different substances co-mentioned? The results show a significant shift in the

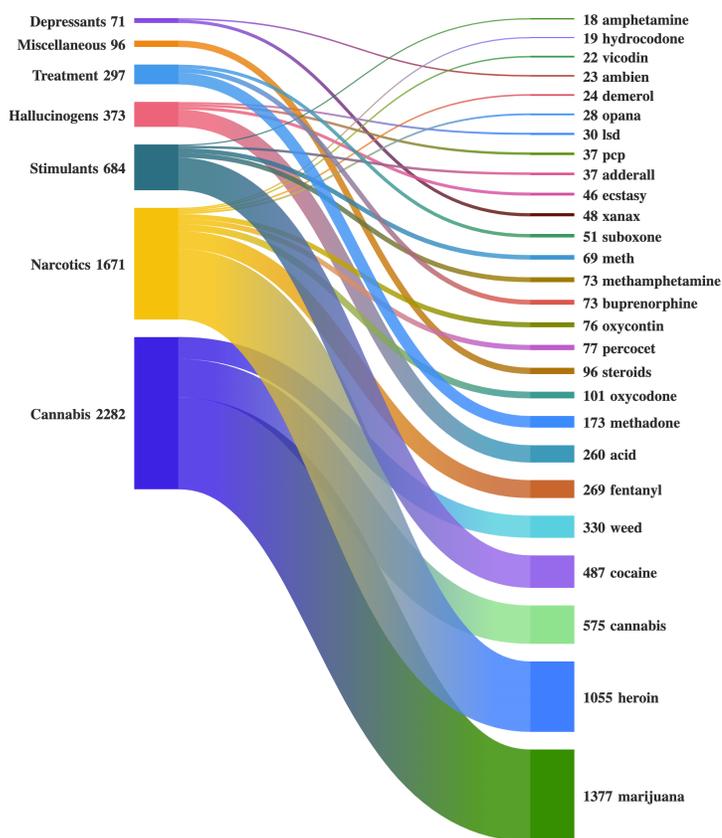


Figure 1: Number of substances per drug class in our dataset.

discussion of illicit substance use over the past decade. As shown in Table 1, cannabis is the most frequently discussed drug class, followed by narcotics. The annual count of news articles indicates that 2018 saw the highest frequency of drug-related mentions in news coverage. In 2018, several significant events occurred, including a former player for the Philadelphia Eagles admitting to recreational cannabis use during his career (Hermann 2019), Philadelphia Mayor Jim Kenney advocating for the legalization of recreational cannabis use (Maas 2018), and the City of Philadelphia hosting its first-ever Cannabis Opportunity Conference (Brown 2018). The frequency of cannabis in the news doesn't necessarily imply that it is used more than other substances, as it could be affected by the prevailing social and political environment. Our findings suggest that the majority of articles mentioning cannabis are centered around the legality of its use. Narcotics (i.e., opioids) are the second most discussed drug class, and also the most prevalent in 2017 and 2018, indicating the severity of discussions and possibly the opioid crisis during those years. Philadelphia experienced some of the highest rates of unintentional overdose rates, and the synthetic opiate fentanyl became increasingly fatal (DOH 2022). As shown in Table 1, *designer drugs* and *drugs of concern* are mentioned in only 16 and 11 articles, respectively. Due to the sparse data on these topics in our corpus, we exclude them from further analysis.

Further analysis of drug counts per class shows that

	Drug 1	Drug 2	N		Drug 1	Drug 2	N
2013	heroin	methadone	52	2018	fentanyl	cocaine	270
	heroin	cocaine	45		heroin	cocaine	194
	barbiturates	cannabis	16		heroin	marijuana	127
2014	heroin	naloxone	42	2019	heroin	marijuana	115
	marijuana	cocaine	36		heroin	naloxone	77
	heroin	cocaine	34		heroin	cocaine	71
2015	heroin	marijuana	40	2020	heroin	methadone	198
	heroin	methadone	19		heroin	buprenorphine	81
	heroin	meth	11		fentanyl	cocaine	53
2016	heroin	marijuana	51	2021	buprenorphine	cannabis	108
	heroin	cocaine	45		heroin	cannabis	78
	heroin	naloxone	44		heroin	marijuana	41
2017	heroin	methadone	157	2022	buprenorphine	cannabis	351
	heroin	cocaine	90		buprenorphine	marijuana	153
	heroin	naloxone	80		fentanyl	cocaine	114

Table 2: The most frequent co-occurred drugs per year.

cannabis is primarily associated with marijuana (Figure 1). Despite the controversy surrounding the use of the term “marijuana” (Halperin 2018), there has been a consistent and steady usage of this word in news articles over the years. In contrast, depressant drugs like Xanax and Ambien are more frequently mentioned together, likely due to their widespread use for anxiety and sleep disorders and their potential for abuse and addiction. Similarly, cocaine and acid are the most frequently mentioned drugs in the stimulant and hallucinogen drug classes, respectively, indicating their enduring popularity and widespread use. It is worth noting that our temporal analysis of drug mentions uncovers patterns in drug discussion over time, which can be influenced by a variety of factors, such as changes in drug policies and cultural shifts in attitudes towards various elements of drug use, like policy, treatment, and research.

The co-occurrence analysis (Table 2) revealed that certain substances are frequently discussed together, indicating a connection between their use, among other possible associations. Heroin was once the most discussed substance in Philadelphia until it was surpassed by fentanyl in 2018. This is in line with reports of adulteration and an increase in unintentional overdose deaths containing fentanyl in Philadelphia (Drug Enforcement Administration 2021). Stimulants, like cocaine, and cannabis are commonly observed to co-occur with narcotics, and there is growing concern among the public and policymakers over an increasing number of unintentional overdose deaths involving stimulants. Marijuana and heroin are frequently mentioned together in discussions around legalization and concerns about adulteration from fentanyl. Our analysis also shows that narcotics and treatment drugs, including methadone, naloxone, and buprenorphine, are the second most commonly co-located drugs in articles. These articles primarily focus on criminality, legislature, and overdose, as determined by our TF-IDF analysis. There is growing concern over the co-occurrence of stimulants, like cocaine, with narcotics, like heroin, due to the rising number of fatalities associated with their use (Johnson et al. 2021). The Philadelphia Inquirer has been particularly active in discussing the “opioid epidemic,” with a focus on treatment. The use of treatment drugs like methadone and naloxone is seen as a positive and proactive response to drug addiction (National Academies of Sciences et al. 2019). Buprenorphine, another popular substance used to treat opioid use disorder, gained attention in 2022 after

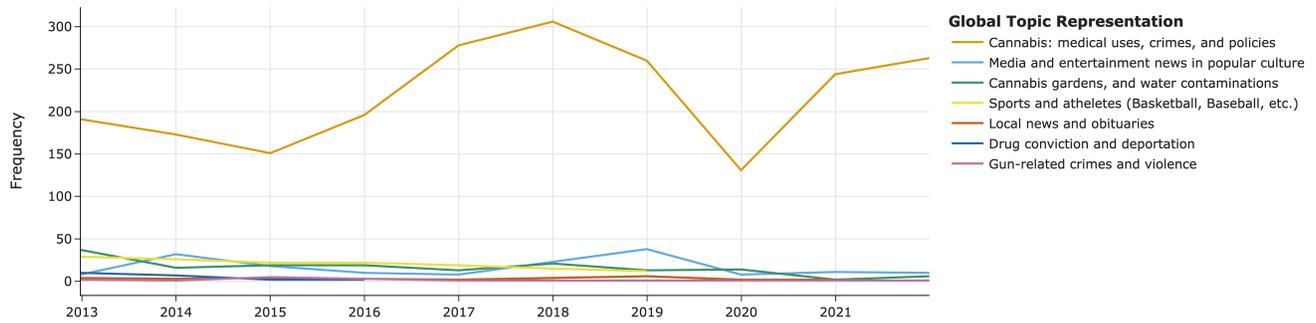


Figure 2: News topic evolution over time, with each line representing a cluster of related topics.

President Biden eliminated the restriction on the number of prescriptions providers could issue per month, making it more accessible compared to methadone (Pew 2021).

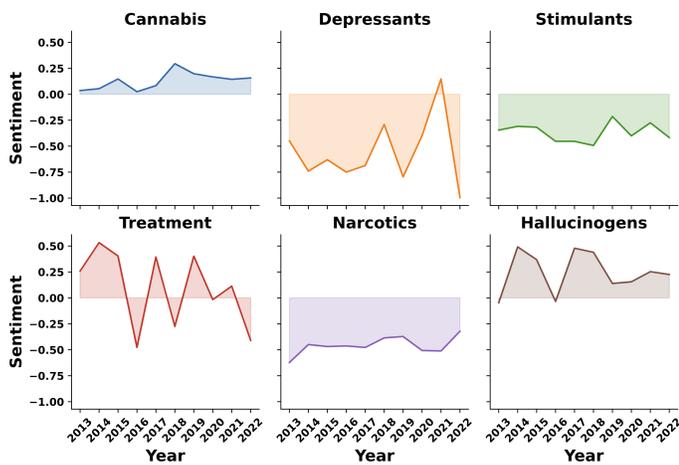


Figure 3: Drug class sentiment scores over the past decade.

How has media coverage of illicit substance use evolved over time? Utilizing dynamic topic modeling with BERTopic, we were able to generate a diverse set of topics covering various themes. The review of the initial results revealed a number of overlapping topics. Therefore, we iteratively refined the model by decreasing the number of topics until we achieved a high-quality output. Figure 2 presents temporal changes of topic clusters in our data. The result shows that cannabis-related topics are the most prevalent in our dataset, with a specific focus on legislation. We also observe a significant number of news articles discussing drug use in films and television series, indicating the possible role of popular media in shaping public perceptions and attitudes toward drugs and social issues (Rezapour and Diesner 2017; Diesner, Rezapour, and Jiang 2016).

The sentiment analysis of news articles demonstrates that the tone of news articles varies across drug classes and evolves over time (Figure 3). News articles with hallucinogenic drugs tend to have a more positive tone compared to other classes. The positive tone may be attributed to recent studies on the use of psychedelic drugs for the treatment of mental health disorders such as treatment-resistant de-

pression and post-traumatic stress disorder. For instance, we identified an article published in the Philadelphia Inquirer discussing the use of LSD and shrooms to treat PTSD (Ao 2020). This article reflects the positive sentiment toward the therapeutic potential of hallucinogenic drugs. Furthermore, the majority of news articles mentioning cannabis, hallucinogens, and treatment drugs exhibit positive sentiments, whereas news articles about narcotics, stimulants, and depressants depict negative sentiments. The sentiment scores assigned to each article range from -1 (most negative) to +1 (most positive), and neutral sentiments are rarely observed in our dataset

Overall, our analysis indicates a shift in the discussion surrounding drug-related topics, especially hallucinogens, towards a more positive view of medicated-assisted treatment and novel applications of these substances. This emphasizes the importance of reporting drug-related incidents using harm reduction principles to encourage safer and managed use. Further research is necessary to fully understand the context and implications of these findings. It is also important to acknowledge that our analysis only captures a subset of the wider discussion surrounding these drugs and that additional studies may reveal further insights.

Conclusion and Future Work

Over the past decade, there has been an evolution in how the media portrays illicit substance use, as our study has revealed. We analyzed news articles published in the Philadelphia Inquirer between 2013 and 2022 and found that cannabis was the most frequently discussed drug class, followed by narcotics. News articles about hallucinogenic drugs tend to have a more positive tone compared to other categories of drugs, while articles on narcotics were the most negative. By examining changes in the tone and frequency of drug-related discussions over time, we can gain a better understanding of how societal attitudes towards drugs and drug policies are evolving, and how this may impact public health and well-being. It is important to note that our study focused solely on news articles published in the Philadelphia Inquirer, and therefore our findings are specific to this source and its coverage of substance use within the Philadelphia and United States. Further research is needed to explore how the portrayal of substance use in the media varies across different regions and cultures.

References

- Ao, B. 2020. Can magic mushrooms and LSD treat depression and anxiety? Scientists are optimistic. *Philadelphia Inquirer*, 17.
- Blei, D. M.; and Lafferty, J. D. 2006. Dynamic topic models. In *Proceedings of the 23rd international conference on Machine learning*, 113–120.
- Brown, C. 2018. Over 700 Gather for Philadelphia’s First Cannabis Opportunity Conference hosted by Senator Street & DACO.
- Caburnay, C. A.; Kreuter, M. W.; Luke, D. A.; Logan, R. A.; Jacobsen, H. A.; Reddy, V. C.; Vempaty, A. R.; and Zayed, H. R. 2003. The news on health behavior: coverage of diet, activity, and tobacco in local newspapers. *Health education & behavior*, 30(6): 709–722.
- DEA. 2020. Drugs of Abuse, a DEA Resource guide (2020 edition).
- Diesner, J.; Rezapour, R.; and Jiang, M. 2016. Assessing public awareness of social justice documentary films based on news coverage versus social media. *IConference 2016 Proceedings*.
- DOH, P. 2022. Unintentional drug overdose fatalities in Philadelphia, 2021.
- Drug Enforcement Administration. 2021. Operation Engage Philadelphia. <https://www.dea.gov/engage/operation-engage-philadelphia>.
- Grootendorst, M. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*.
- Halperin, A. 2018. Marijuana: is it time to stop using a word with racist roots? *The Guardian*.
- Hayden Griffin, O.; Leban, L.; Singh, D.; Webb, M.; and Narayanan, S. 2023. Deconstructing Differential Drug Coverage within a Malaysian Media Source. *Journal of Psychoactive Drugs*, 0(0): 1–7. Publisher: Taylor & Francis .eprint: <https://doi.org/10.1080/02791072.2023.2186287>.
- Hermann, A. 2019. Former Eagle Chris Long, now retired, says he regularly used marijuana during NFL career.
- Hughes, C. E.; Lancaster, K.; and Spicer, B. 2011. How do Australian news media depict illicit drug issues? An analysis of print media reporting across and between illicit drugs, 2003–2008. *International Journal of Drug Policy*, 22(4): 285–291.
- Hutto, C.; and Gilbert, E. 2014. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1): 216–225. Number: 1.
- Johnson, J.; Pizzicato, L.; Johnson, C.; and Viner, K. 2021. Increasing presence of xylazine in heroin and/or fentanyl deaths, Philadelphia, Pennsylvania, 2010–2019. *Injury Prevention*, 27(4): 395–398. Publisher: BMJ Publishing Group Ltd Section: Brief report.
- Johnston, L.; O’malley, P. M.; and Bachman, J. G. 2004. Monitoring the future: national survey results on drug use, 1975–2003.
- Maas, D. 2018. Mayor Kenney pushes marijuana legalization, again.
- National Academies of Sciences, E.; Division, H. a. M.; Policy, B. o. H. S.; Disorder, C. o. M.-A. T. f. O. U.; Mancher, M.; and Leshner, A. I. 2019. *The Effectiveness of Medication-Based Treatment for Opioid Use Disorder*. National Academies Press (US). Publication Title: Medications for Opioid Use Disorder Save Lives.
- Niamatullah, S.; Auchincloss, A.; and Livengood, K. 2022. Drug Overdose Deaths in Big Cities 2022.
- Paalman, M. 1997. How to do (or not to do)... media analysis for policy making. *Health policy and planning*, 12(1): 86–91.
- Pew. 2021. Buprenorphine Treatment for Opioid Use Disorder in Philadelphia.
- Rezapour, R.; and Diesner, J. 2017. Classification and detection of micro-level impact of issue-focused documentary films based on reviews. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 1419–1431.
- Swalve, N.; and DeFoster, R. 2016. Framing the Danger of Designer Drugs: Mass Media, Bath Salts, and the “Miami Zombie Attack”. *Contemporary Drug Problems*, 43(2): 103–121. Publisher: SAGE Publications Inc.