



Selected Papers of #AoIR2017:
The 18th Annual Conference of the
Association of Internet Researchers
Tartu, Estonia / 18-21 October 2017

‘IF YOU’RE GOING THROUGH SOMETHING DIFFICULT, WE’D LIKE TO HELP’: THE LIMITATIONS OF HASHTAG LOGICS IN PRO-EATING DISORDER CONTENT MODERATION.

Ysabel Gerrard
University of Sheffield

Paper rationale

Social media companies loudly encourage their users to ‘share’ content about themselves. But they are much quieter about how they moderate problematic posts and why they choose to do so (see Gillespie, 2015). Platforms often make decisions about moderation when they face public pressures, such as accusations that they host pro-eating disorder (pro-ED) content. This is precisely what happened in February 2012, when a *Huffington Post* writer published a widely read exposé on the ‘secret world’ of Tumblr’s thinspiration blogs (Gregoire, 2012). By May of 2012, Tumblr - along with Instagram and Pinterest - had publically announced its plans to minimize the spread of pro-ED content. To do this, the three platforms issue public service announcements (PSAs) when users search for troubling terms, like #proana and #thinspiration, and Instagram also blocks results for certain hashtag searches.

Using pro-ED content as a case study, this paper addresses the problems of hashtag logics in decisions about, and discussions of social media content moderation. It explores how: (1) pro-ED content can still be found without using the hashtag as a search mechanism; (2) users learn to recognize and signal each other as pro-ED in the absence of hashtags, and given the vulnerability of their posts to user-driven forms of moderation like flagging (see Crawford and Gillespie, 2016), and (3) platforms’ recommendation systems suggest pro-ED content to users *despite* policing. This makes hashtag moderation a minor, perhaps even ineffective intervention, if indeed platforms should intervene at all.

Instagram, Pinterest and Tumblr have joined a longer debate about mediated depictions of disordered and typically young, female, and white bodies (see Bordo, 2003), following the standards set by older social networking sites (SNSs) like MySpace and Xanga, and Gerrard, Y. (2017, October 18-21). *‘If you’re going through something difficult, we’d like to help’: the limitations of hashtag logics in pro-eating disorder content moderation*. Paper presented at AoIR 2017: The 18th Annual Conference of the Association of Internet Researchers. Tartu, Estonia: AoIR. Retrieved from <http://spir.aoir.org>

Web hosts like Yahoo! (see Dias, 2003). But what made the 2012 iteration so unique was the way the platforms chose to intervene: through hashtag moderation. Hashtags are an appealing point of intervention for various reasons. They are convenient tools for aggregating relevant content between users outside of each other's follower/followee networks (see Schmidt, 2014), they circumvent the difficulties of algorithmically tagging visual imagery to categorize it as 'pro-ED', and they help commercial content moderators to interpret posts within the 'seconds' (Roberts, 2017) they have to decide whether it should stay or go.

But hashtag moderation has produced a partial way of understanding how pro-ED content circulates on platforms. These limitations are reflected in popular media discussions, as the press use the hashtag as shorthand for a larger and more complex issue. For example, a *BuzzFeed* writer argues that images of 'bony torsos and skeletal limbs' are easily findable on social media 'thanks to popular hashtags like #anorexia, #ana, and #thinspiration' (Lee, 2016). The hashtag also plays an oversized role in recent research on eating disorders and social media. It is used as a tool for data collection (see Ging and Garvey, 2017), and for understanding how users evade Instagram's hashtag ban (see Chancellor et al, 2016).

Research methods

This paper explores how pro-ED content circulates *beyond the hashtag*. It responds to recent scholarly calls for more methodological approaches to obtaining untagged content (see Bruns et al, 2016; D'heer et al, 2017). It asks how we locate pro-ED content without relying on in-platform hashtag searches; a 'methodologically difficult' (Bruns et al, 2016, p.21) question that required some innovation.

I began my research by creating new accounts on Instagram, Pinterest and Tumblr to avoid algorithmic bias. I took a platform-specific approach to sourcing content, influenced in part by Light et al's (2016) app walkthrough method. Instagram, for example, has several levels of content moderation (see Suzor, 2016). Although it does not return results for certain hashtag searches, the platform does not ban searches for users with pro-ED terms in their usernames or profile bios. Similarly, Tumblr allows its users to 'follow' certain terms like 'proana' and 'promia', despite issuing PSAs on tag searches. It then presents users with relevant content through their algorithmically organized dashboards. Pinterest allows users to save images to their 'pin boards'. After creating some of my own boards, the platform recognized me as a 'pro-ED user' and began recommending such content to me via email and as I browsed the site. In short, it algorithmically presented me with the very same content it tries to moderate.

Using these platform-specific techniques, I collected 2162 posts from 272 Instagram, Pinterest and Tumblr accounts. In this paper, I discuss both the posts and the methods used to obtain them, highlighting the limitations of hashtag logics in understanding how pro-ED content circulates on social media.

References

Bordo, S. (2003). *Unbearable weight: feminism, Western culture, and the body*. 10th anniversary edition. Berkeley: University of California Press.

Bruns, A. and Burgess, J. (2015). The use of Twitter hashtags in the formation of ad-hoc publics. In N. Rambukkana. (Ed.), *Hashtag Publics: The Power and Politics of Discursive Networks* (pp. 13-28). New York: Peter Lang.

Bruns, A., Moon, B., Paul, A. and Münch, F. (2016). Towards a typology of hashtag publics: a large-scale comparative study of user engagement across trending topics. *Communication Research & Practice*. 2(1). 20-46.

Chancellor, S., Pater, J.A., Clear, T., Gilbert, E., and De Choudhury, M. (2016). #thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities. *CSCW '16 Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. [Online]. Available at: http://www.munmund.net/pubs/cscw16_thyghgapp.pdf [Accessed: 6th December 2016].

Crawford, K. and Gillespie, T. (2016). What is a flag for? Social media reporting tools and the vocabulary of complaint. *New Media and Society*. 18(3). 410-428.

Dias, K. (2003). The ana sanctuary: women's pro-anorexia narratives in cyberspace. *Journal of International Women's Studies*. 4(2). 31-45.

D'heer, E., Vandersmissen, B., De Neve, W., Verdegem, P. and Van de Walle, R. (2017). What are we missing? An empirical exploration in the structural biases of hashtag-based sampling on Twitter. *First Monday*. 22(2). DOI: <http://dx.doi.org/10.5210/fm.v22i2.6353>.

Gillespie, T. (2015). Platforms intervene. *Social Media and Society*. 1-2. DOI: 10.1177/2056305115580479.

Ging, D. and Garvey, S. (2017). 'Written in these scars are the stories I can't explain': a content analysis of pro-ana and thinspiration image sharing on Instagram. *New Media and Society*. 1-20. DOI: 10.1177/1461444816687288.

Gregoire, C. (2012, 9th February). THE HUNGER BLOGS: a secret world of teenage 'thinspiration'. *Huffington Post*. [Online]. Available at: http://www.huffingtonpost.com/2012/02/08/thinspiration-blogs_n_1264459.html [Accessed: 12th July 2017].

Lee, S.M. (2016, 14th April). Why eating disorders are so hard for Instagram and Tumblr to combat. *BuzzFeed*. [Online]. Available at: https://www.buzzfeed.com/stephaniemlee/why-eating-disorders-are-so-hard-for-instagram-and-tumblr-to?utm_term=.qj0vjPzPK#.gayq5EmEa [Accessed: 13th August 2017].

Light, B., Burgess, J. and Duguay, S. (2016). The walkthrough method: an approach to the study of apps. *New Media and Society*. 1-20. DOI: 10.1177/1461444816675438

Roberts, S.T. (2017, 8th March). Social media's silent filter. *The Atlantic*. [Online]. Available at: <https://www.theatlantic.com/technology/archive/2017/03/commercial-content-moderation/518796/> [Accessed: 28th September 2017].

Schmidt, J. (2014). Twitter and the rise of personal publics. In K. Weller, A. Bruns, J. Burgess, M. Mahrt and C. Puschmann. (Eds.), *Twitter and Society* (pp. 3-14). New York: Peter Lang.

Suzor, N. (2016, 17 September). How does Instagram censor hashtags? *Medium*. [Online]. Available at: <https://digitalsocialcontract.net/how-does-instagram-censor-hashtags-c7f38872d1fd> [Accessed: 29th July 2017].