

Eleonore Neufeld and Elise Woodard, 'On subtweeting'

BENJ HELLIE

BENJ.HELLIE@UTORONTO.CA

BENJ.CA

APA-Pacific Division, 12 April 2022

1 Person-evaluative tweeting: two dimensions

Valence *insulting* := $V < 0$... *friendly* := $V > 0$

Directness *inexplicit* := $D < 0$... *de-@-ed* := $D = 0$... *@-ing* := $D > 0$

2 What is 'subtweeting'?

the NO-@ analysis $D \leq 0$

A tweet that mentions a Twitter member **without using their actual username**. Usually employed for negative or insulting tweets; the person you're mentioning won't see the subtweet in their Twitter timeline as it doesn't contain the @ symbol that every Twitter username has.¹

the BEHIND-THE-BACK analysis $D \leq 0, V < 0$

Basically, it's **talking about someone behind their back** but sort of in their face on Twitter!²

Subtweeting is a way to **publicly talk about someone behind his back**.³

[Talk 'behind someone's back' has a clear negative valence—so not merely 'talk unawares']

the BLIND-ITEM analysis $D < 0, V < 0$

The art of subtweeting isn't a particularly difficult one. It involves posting something on social media (usually Twitter) that's most likely **negative**, most likely about someone you know, but **always indirect**. **No names are mentioned, no handles are included**, and the post is **passive-aggressive, rarely specific, but hints that someone has offended** the original poster.⁴

['Most likely negative' is compatible with $V \geq 0$, but the remaining discussion undermines any likelihood of nonnegative valence]

3 'Essential features of subtweets'

N&W sign on with the BLIND-ITEM analysis, contending in particular that $D < 0$ is an 'essential feature[] of subtweets'; but

↪ Both NO-@ and BEHIND-THE-BACK analyses are compatible with $D \geq 0$

↪ N&W's substantiation for the BLIND-ITEM analysis tends instead to undermine it. Two examples:

- (i) The Merriam Webster dictionary defines [the subtweet] as 'a usually mocking or critical tweet that alludes to another Twitter user without including a link to the user's account and **often** without directly mentioning the user's name'. [] **By definition**, the target of a subtweet is not mentioned by name.⁵

What happens only seldom is not yet ruled out by definition

¹Current leading definition on Urban Dictionary: my boldface emphasis; compare also a 22 March 2013 'Branch' discussion, initiated by Hillary Reinsberg, in which subtweeting is legislated as carrying this more expansive meaning: web.archive.org/web/20130322071844/http://branch.com/b/defining-the-subtweet

²From the initial definition on Urban Dictionary, pulled from knowyourmeme.com/memes/subtweet: my boldface emphasis

³www.businessinsider.com/subtweeting-explained-2012-12: my boldface emphasis

⁴sova.pitt.edu/social-media-guide-the-art-and-harm-of-subtweeting: my boldface emphasis

⁵N&W, section 2: my boldfaced emphasis

- (ii) Indeed, one popular article argued that subtweets that include obviously uniquely identifying descriptions are not subtweets[]. At the very least, they are not *good* subtweets.⁶

The article instead addresses only the status of the weaker claim, that a subtweet does not ‘name a person directly’; on one reading, its point is only that this makes for a *bad* subtweet⁷

⁶N&W, fn5; discussing www.gq.com/story/dear-everyone-that-was-not-a-subtweet

⁷An indication: ‘In order for it to be a **real** subtweet, the target should remain nameless, so that people replying can go ‘Oooooooooooh, SHADE!’’. But I guess SOME people just don’t understand how to do a drive-by properly now, do they?’: my boldface emphasis; compare Austin 1962 on ‘real’

4 What’s the problem with subtweeting?

Well, on which analysis?

4.1 NO-@ subtweeting: a discourse-structure complaint

When $D \leq 0$,

[I]t intentionally breaks down the @-tagging, conversational framework that Twitter is built on, whereby you’re very much able to track everything people are saying about you⁸

⁸www.buzzfeed.com/hillaryreinsberg/what-is-a-subtweet

Here, V does not matter: inaccurate praise would also countervail, and thus to some extent corrupt or undermine, the community presumption that discussion of target T will be ‘transparent’ to T

↪ This community presumption is interesting, and deserves further exploration: clearly, pre-internet, no such thing could be realized, at least not in a publication environment of any size, and (to my very limited knowledge) was not realized prior to Twitter, either—are there distinctive values attained through the @-tagging framework?

Certain prior phenomena of discourse structure—assertion (Stalnaker 1978) and subject-matter (Roberts 2012)—are well-studied; but less is known about the interaction of *subject-matter* with *participation* (when a conversant *is* the subject-matter): e.g., a person’s distinctive doxastic/epistemic position in regard to themselves may bear interestingly on the structure of conversation directed at the aggregation of distributed information/knowledge⁹

⁹The *command* literature treats a different sort of conversation about its participants: Lewis 1979, Portner 2004

4.2 BEHIND-THE-BACK subtweeting: beyond libel?

When source S ’s $D \leq 0$ tweet about target T also has $V < 0$, the result is a statement about T which is *insulting* and *published* (and may thus be *defamatory* to T), and which also countervails the transparency presumption

↪ N&W contrast *subtweeting* with various ‘in-person’ defamatory speech acts, correctly observing that ‘the sheer number of expected readers can be far greater[; and that] the potential consumers include potential future readers’ (section 2): still, the contrast is that tweets are publications, and is not *eo ipso* about subtweets

↪ Are there important ethical differences between BEHIND-THE-BACK subtweeting and print-media defamatory publication? E.g., perhaps violating the transparency presumption brings an additional increment of unfairness to the target, or cowardliness by the source

N&W discuss BEHIND-THE-BACK subtweets qua ‘supertweets’, contending that

to attempt to respond to the tweet or defend oneself is to risk appearing self-centered, self-serious, or overly defensive, if not worse (section 3)

↪ This is a general challenge for seeking redress against defamation, and appears to accrue alike to print-media publication and to $V < 0$, $D > 0$ tweets

4.3 BLIND-ITEM subtweeting: N&W's complaints

4.3.1 Inspecific and deniable?

When S 's tweet about T of $V < 0$ also has $D < 0$ and is thus 'inspecific' or 'indefinite' or—most accurately—*inexplicit* that T is targeted, N&W contend that this endows S with 'deniability' that they have insulted T

↪ But $D < 0$ is not *requisite* for 'deniability': the all-purpose denial, 'it was a joke', invariably *can* be given

↪ And defamation tort law recognizes a 'doctrine of innuendo':

To render the defamatory statement actionable, it is not necessary that the false charge be made in a direct, open and positive manner. A mere insinuation is as actionable as a positive assertion if it is false and malicious and the meaning is plain. Statements therefore may be either defamatory on their face, or defamatory by way of innuendo. Innuendo is extrinsic evidence used to prove a statement's defamatory nature. It includes the aid of inducements, colloquialisms, and explanatory circumstances.¹⁰

If case law here encodes the relevant 'folk morality', the latter apparently does not in fact endow blanket deniability just from inexplicitness

¹⁰Fountain v First Reliance Bank, 730 SE 2d 305 - SC: Supreme Court 2012: internal citations omitted

4.3.2 Downstream 'ethical harms'?

N&W locate (section 3) four distinctive 'ethical harms' in BLIND-ITEM subtweeting; following their labeling, these are:¹¹

Normative cover 'the deniability and inspecificity of subtweets gives their authors a publicly acceptable way to perform actions that would otherwise be clearly unacceptable'

↪ Even aside from the 'doctrine of innuendo' worry, science casts doubt on the 'public acceptability' claim:

readers formed consistently bad impressions of people who subtweet: they're less likely to want to befriend them, less likely to think they're socially competent and less likely to think they shared any personal similarities¹²

↪ In what sense is this a 'harm'? Of course, when S defames T , S harms T ; but the harm in defamation is the public recognition that negative features have been attributed to T : if a statement is hard to recognize as a defamation of T , would this not instead *diminish* the harm to T ?

↪ If someone burglarizes my house, the burglary is a harm; is it an additional harm to me if they burglarize skillfully enough to evade punishment? This seems at odds with the perhaps commonsensical view that the offender's punishment *redresses* the harm: the offender's evasion of punishment *deprives the victim of redress*, but does not contribute to the victim an additional increment of harm

Gaslighting 'in virtue of [the inexplicitness of subtweets], even the most plausible target often second-guesses their fit, wondering whether their belief results from paranoia, social anxiety, or self-centeredness[;] in instances where the author of the subtweet is confronted, they can exploit the possible deniability of subtweets, which can exacerbate gaslighting experience'

↪ Wow, that would require a pretty thick layer of shade!¹³

Self-defense 'To defend oneself, or call out the tweet, would require first outing oneself as the target and centering oneself in the message. However, this risks appearing paranoid at best and narcissistic at worst'

¹¹N&W's central, 'Anastasia v Prof Warren' example did not strike me as very true-to-life: by the time one is invited to give a 'keynote address' in pretty much any forum, one's views are sufficiently well-developed that there can be no such thing as a 'devastating objection' (or else one is sufficiently boneheaded that one would not notice such a thing if it came in!); or one is good enough at debating that no objection could plausibly be made to stick; or one is interested enough in ideas that an incisive objection would be a valued opportunity to revisit fundamentals and 'do philosophy in public'—and one's skin is sufficiently thick that one would care little at all about some grad student's complaint.

¹²www.washingtonpost.com/news/the-intersect/wp/2016/06/06/study-confirms-what-you-always-knew-people-who-subtweet-are-terrible/, reporting work by Edwards and Harris

¹³I worry that no such layer is found in the 'Anastasia/Warren' example: given that Warren was in Anastasia's audience, Warren's subtweet manifestly singles out Anastasia determinately.

↔ Analogues of the second and third complaints about the **Normative cover** ‘harm’ seem applicable: to the extent that one is not ‘centered’, this would seem to instead *diminish* the defamation harm; and the problem this locates is more for *interference with redress*, rather than an additional increment of harm

Complicity ‘the participating public—often unknowingly—[] may facilitate, enable, or condone the harm by engaging positively with subtweets[;] it will often result in engaging with content they would often not consent to engage with, were they aware of the respective target’

↔ The thought perhaps is that subtweeting is a kind of ‘hyper-libel’, luring the audience into accelerating the libel well beyond standard, print-media libel.

But the ‘hyper-libel’ worry would initially seem to be a problem with *easy online publication*, and not, *eo ipso*, with *subtweeting*—what is that aspect of @BenjHellie has vices that makes it significantly less liable as a lure for approbation from the general, ignorant public than Benj Hellie has vices, or A certain north-country philosopher who brags about his cute dog has vices?

References

- Austin, John Langshaw, 1962. *Sense and Sensibilia*. Oxford: Oxford University Press.
- Lewis, David, 1979. ‘A problem about permission’. In *Essays in Honour of Jaakko Hintikka on the Occasion of His Fiftieth Birthday on January 12, 1979*, 163–175. Dordrecht: Reidel. Reprinted in [Lewis 2000](#).
- Lewis, David, 2000. *Papers in Ethics and Social Philosophy*. Cambridge: Cambridge University Press.
- Portner, Paul, 2004. ‘The semantics of imperatives within a theory of clause types’. *Proceedings of SALT*, 14.
- Roberts, Craige, 2012. ‘Information structure in discourse: Towards an integrated formal theory of pragmatics’. *Semantics and Pragmatics*, 5:1–69.
- Stalnaker, Robert C., 1978. ‘Assertion’. *Syntax and Semantics*, 9:315–32. Reprinted in [Stalnaker 1999](#).
- Stalnaker, Robert C., 1999. *Context and Content*. Oxford: Oxford University Press.