LIWC (Linguistic Inquiry and Word Count) Categories in Engle Trial Closing Arguments

Within the main body of our article we provided an "open-vocabulary" analysis [1] of closing statements from Engle Progeny litigation (in Florida), using the 50,000 most common 1 to 10 grams within our dataset of defense vs. plaintiffs' closings. Here we extend this analysis by using Linguistic Inquiry and Word Count (LIWC) [2 3], a tool commonly used to investigate the traces left by beliefs, fears, styles of thought, and social relationships in our everyday language (for a review of such techniques, see [4]). One LIWC analysis has shown, for example, that as people grow older, they use fewer I-words (I, me, my, myself) and more we-words (we, us, our) and verbs that are more likely to be in the future tense. [5] Another found that lying was associated with "more negative emotion, more motion words (e.g. arrive, car, go), fewer exclusion words, and less first-person singular."[2] LIWC allows us to compare two sets of texts to identify which is more likely to use pronouns, articles, or auxiliary verbs, but also which is more likely to express particular types of emotion, cognition, or perception. LWIC contains a master dictionary of 6400 words, which belong to one or more of its 73 linguistic and/or psychometric categories. A category can be made up of a few to many hundreds of individual words. The LIWC category "anger," for example, includes words like "blame," "lie," "mad," or "fight."

Applying LIWC to our dataset allows us to investigate whether findings from our open-vocabulary analysis can be generalized--by looking more broadly into the sentiments attached to particular words and word strings. We showed in the main body of our article, for example, that plaintiffs highlight the "profit" tobacco companies make, and are thus far more likely to use that word. LIWC identifies a broader class of money-related terms--like "business," "marketing," or "sales," for example--which allows us to ask: does the industry avoid all money-related terms? Or just "profit" as we'd found? Or consider this example: We found that the terms "he" and "she" are highly distinctive for the defendants. LIWC allows us to test whether this finding can be generalized to *all* 3rd person singular pronouns: he, she, his, hers, him, her, himself, herself. Finally, by allowing us to investigate the use of verb *tenses* (instead of just individual verbs), LIWC reveals that cigarette industry lawyers consistently focus on the past, while plaintiffs focus on the present and the future--for reasons we shall explore in a moment.

Here, we review the five most distinctive LIWC categories for plaintiffs vs. defendants (see Table 1). Readers can investigate the scores produced for all 73 LIWC categories in further detail at www.tobacco-analytics.org/litigation.

Table 1. Maximally divergent LIWC categories by Mann-Whitney Rho scores

Defense attorneys will *often* use terms from the following LIWC categories						
	Defense	**Plaintiff**	**FS**	**MWR**	**Defense Terms**	**Plaintiff Terms**
Third person singular pronouns	96,185	61,103	0.39	0.12	he, she, his, her, him, himself, herself	*

Past focus	172,101	142,293	0.46	0.16	was, said, told, did, heard, had, asked, made	were, knew
Negations	63,850	53,549	0.46	0.21	no, not, never, none, nothing, neither, nor	without, nobody, cannot
Time	109,592	97,946	0.48	0.29	never, now, when, again, ever, before, after, time, during, first	*
Ingestion	50,556	40,648	0.45	0.29	smoking, smoke, smoked, smoker, digest, drinking, alcohol	cigarettes, cigarette, weight
Plaintiff attorneys will *often* use terms from the following LIWC categories						
	Defense	**Plaintiff**	**FS**	**MWR**	**Defense Terms**	**Plaintiff Terms**
Third person plural pronouns	27,796	59,792	0.69	0.96	*	they, their, them, themselves, theirs
First person plural pronouns	21,525	31,508	0.60	0.83	us, ourselves	We, our
Money	10,323	14,809	0.59	0.80	deposition	business, free, money, market, customers, selling, sales, sell, marketing
Present focus	244,868	277,582	0.54	0.78	now, has	is, are, do, know, be, say, go, believe
Future focus	22,628	28,488	0.56	0.78	wants, sometime	will, going, then, may, future, coming, ahead, hope,

P-value < 0.0001 for all Frequency and Mann-Whitney Rho scores.
Terms listed here as "Defense Terms" and "Plaintiff Terms" are the ten words that show the largest absolute difference in usage between these two adversarial parties. Terms that appear more often in defense closings are listed as "Defense Terms," while terms that

appear more often in closings by plaintiffs appear as "Plaintiff Terms." Some LIWC categories (e.g.,,first person plural pronouns) contain fewer than ten terms.
* None among the ten words with the largest absolute difference in usage appear more often in the plaintiff (or defense) side in this category.

Table 1 shows that a LIWC analysis reveals many of the same patterns we found in our open-vocabulary analysis. Defense attorneys use third person singular pronouns like "he," "she," "him," or "her" to focus attention on the individual smoker and what he or she may have heard, seen, or did (recall our section on "Pronoun Politics"). By contrast, plaintiffs use third person plural pronouns like "they" or "their" to focus attention on the industry's misdeeds. Plaintiffs' lawyers also use "we" and "our" to highlight the actions of industry executives: "Was it reprehensible? Is killing people for profit … while lying to them since the '50s saying, we want to be frank, we want to be frank, we want to be honest?" [6] Similarly, the frequency with which defense attorneys use negations ("no," "not," "never," "none") should not be surprising: juries in the Engle trials are asked to assign comparative fault, which means they can assign however much blame they want to the smoker, with the remainder put on the industry (figured as a percentage). Plaintiffs will often admit that the smoker is partly at fault, since that could still mean millions of dollars in a verdict. Defense attorneys will typically argue that cigarette makers bear no fault at all--which is why they are more to employ negations. They'll deny, for example, that a smoker was ever addicted to nicotine: "None of Mr. Ahrens' medical records ever revealed a diagnosis of addiction. They never revealed a diagnosis of nicotine dependence.... Nothing like that was in any of the medical records." [7]

A striking result emerging from this LIWC analysis is the difference in temporal focus between plaintiffs and defendants: industry lawyers focus on the past, while plaintiffs focus on the present and the future. This is partly a result of court procedure. The plaintiffs use future-oriented rhetoric because they deliver their closing argument first, before the defense takes its turn. Plaintiffs' lawyers thus try to preempt arguments they expect from their opponents: "They [the defense lawyers] are going to tell you that it was Frank Gafney's choice to quit smoking...."[8] By contrast, the defendants use more past-oriented rhetoric, as part of their strategy of re-narrating the smoker's life and all the decisions and choices that led him or her to smoke.

However, a more intriguing interpretation of this temporal asymmetry is that defendants are trying to establish a clear break between the industry's past misdeeds and its present state, while plaintiffs want to highlight continuities between Tobacco's past and present actions. Establishing such continuities can help to justify punitive damages. Juries can award punitives if they decide that punishment is warranted--but also to deter others from similar bad acts. Lawyers suing the industry will often use the present tense ("the historical present") when narrating past events: "This is from 1980 into the next decade…. [Brown & Williamson researchers] are bemoaning the fact that Philip Morris is getting more of the 14 year olds than they are, and they conclude, 'Hopefully our various planned activities that will be implemented this fall will aid in some way in reducing or correcting these trends.'" [9] The plaintiffs draw past, present, and future together: we will show you, the jury, the industry's past misconduct and you will have the opportunity to deliver justice: "Who on this earth with any sense of morality would make a cigarette that … they know causes cancer and kills people? … And that's why, like His Honor said, … you have a tremendous responsibility in this case to right this tremendous wrong." [10] By contrast, lawyers working for the industry try to create a separation between the industry's conduct past and present--and will often

disclaim past actions: "Well, you heard about what Reynolds is doing today. It is not the same company that Claude Teague worked for…. What Reynolds is doing today differs from some of the things you've read in some of the more egregious sounding documents." [11] This is linked to the industry's broader strategy of trivializing bad acts from their (purportedly distant) past: "We agree that these companies are not perfect. Did they hold on to some ideas too long? In 20/20 hindsight, probably. Did people have some stupid ideas in documents in the past? Probably." [12] Who are we to judge these distant times, when things were different?

It is important to realize certain limitations to a LIWC analysis in the courtroom context. LIWC classifies "weight" as an "ingestion" term, for example, along with terms like "cigarettes" and "smoking." But by actually reading the text, we find that "weight" most often appears in the phrase "greater weight of the evidence," referencing the standard plaintiffs must satisfy to prove their case. LIWC also classifies use of the term "honor" as a positive emotion, when it is actually part of the verbal ritual of court procedure ("Objection, your honor."). Defendants raise about three times as many objections as plaintiffs (2088 to 748), often to move for a mistrial--since cigarette makers count mistrials as wins. However, because those defense objections occur in the plaintiffs' closings, they are counted in our analysis as appearing on the plaintiffs' side. (During pre-processing, we stripped out the sidebars that follow objections but not the raising of an objection itself.) Because of these fine details of courtroom rhetoric, we have found LIWC to be most useful when analyzing part-of-speech categories but not so much for affect or cognition categories, that rely on the meaning of individual words

REFERENCES

1. Schwartz HA, Eichstaedt JC, Kern ML, et al. Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one* 2013;8(9):e73791.
2. Tausczik YR, Pennebaker JW. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology* 2010;29(1):24-54.
3. Pennebaker JW, Boyd RL, Jordan K, et al. The development and psychometric properties of LIWC2015, 2015.
4. Gonçalves P, Araújo M, Benevenuto F, et al. Comparing and combining sentiment analysis methods. *Proceedings of the First ACM Conference on Online Social Networks* 2013:27-38.
5. Pennebaker JW, Stone LD. Words of wisdom: Language use over the life span. *Journal of Personality and Social Psychology* 2003;85(2):291.
6. Closing phase 1 - Schlesinger, Hammer, Reilly, St Peter Griffith in Boatright vs Philip Morris, et al. November 10 2014. https://www.industrydocuments.ucsf.edu/docs/rlhn0226 (accessed 2019 April 20).
7. Ahrens vs RJ Reynolds Tobacco Company et al. Transcript of jury trial, volume 32. Closing statements. February 10 2016. https://www.industrydocumentslibrary.ucsf.edu/tobacco/docs/kglw0225 (accessed 2018 August 17).
8. Closing arguments by Gustafson, Lewis, Sales, Maguire, Bassett in Gafney. September 19 2013. https://www.industrydocuments.ucsf.edu/docs/snhn0226 (accessed 2019 April 21).

9. Engle progeny cases Tobacco litigation. Pertains to Martin, case no. 2007-ca-2520. Volume xxiv; phase 1 closing statements. May 29 2009. https://www.industrydocuments.ucsf.edu/docs/xrkw0225 (accessed 2019 April 21).
10. Closing statements by Shevin, Domnick and Belasic in Baker. September 19 2012. https://www.industrydocumentslibrary.ucsf.edu/tobacco/docs/klhn0226 (accessed 2018 October 01).
11. Engle progeny cases Tobacco litigation: Julie a Reese vs RJ Reynolds Tobacco Company, et al; volume 23. May 19 2011. https://www.industrydocuments.ucsf.edu/docs/hppk0191 (accessed 2019 April 21).
12. Closing argument by Woods in Sury. November 22 2011. https://www.industrydocuments.ucsf.edu/docs/qthn0226 (accessed 2019 April 21).