

Research Note: A Modest Experiment

James L. Horton

Recently, in a moment of curiosity, I decided to look at press releases appearing on PR Newswire. I had two questions: 1. How many releases were newsworthy and should have run on a newswire? 2. How well written were the releases?

A modest experiment was in order. I chose an arbitrary starting point to begin the examination and took every other release posted. I chose 20 releases randomly while excluding brief announcements of events, dividends or other short notes. I used a standard text analyzer to look at releases for basic information – number of words, paragraphs, average words per sentence, etc.

What I learned won't surprise most PR practitioners. Only 50 percent of the releases had news value beyond narrow audiences. It was questionable whether they ever should have run on PR Newswire. The practitioners could have easily sent the releases directly to key audiences and posted them on their organizations' web sites where they would be picked up eventually by search engines. Secondly, most of the releases were written in news style but varied in quality. Some were short and to the point. Others were long and self-laudatory. One wonders whether the releases were sent out for consideration by news media or to bolster the self-image of those who ordered them to be written. One wonders also whether PR practitioners tried to fight off sending some of these releases or dutifully bowed their heads over their keyboards. If sending inane releases was solely to get them into online databases, there was possibly merit in spending the money to transmit them. If not, they were a waste of time and talent.

The second question formed the basis for most of the modest experiment. I looked at common writing statistics to see how the releases stacked up. I scored each release individually then ranked them against the cumulative averages and standard deviations for all 20 releases. Doing this provided a third answer as well. That is whether common writing statistics make much sense and whether PR practitioners should pay attention to them.

Here then is the summary of results ranked by readability index with a low number more readable than a high one:

	No of words	No. Distinct Words	Lexical Density Ratio distinct/no	Av words - Sentence	Ave words - Para	Ave Char/word	No. sentences	No. Para	Ave Sent/para	Longest word- char	Longest Sentence	Readability index (lower is better)
Bnro Text 11	554	223	0.403	10.65	110.8	4.36	52	5	10.4	12	48	4.4203
Sgr Text 5	198	131	0.662	13.2	66	4.4	15	3	5	11	28	5.8892
EPA - Text 19	210	145	0.690	11.05	52.5	4.77	19	4	4.75	14	48	6.5697
Freddie - Text 18	286	159	0.556	12.43	57.2	4.99	23	5	4.6	13	39	8.3045
Power - Text 8	305	168	0.551	12.71	61	5.58	24	5	4.8	14	36	11.2075
Premium - Text 20	469	237	0.505	15.63	117.25	5.49	30	4	7.5	15	33	12.2465
Gypsum - Text 1	287	170	0.592	17.94	95.67	5.28	16	3	5.33	15	40	12.4016
Gas - Text 4	728	356	0.489	18.67	121.33	5.24	39	6	6.5	15	54	12.5855
Oakland - Text 13	795	346	0.435	17.28	88.33	5.51	46	9	5.11	17	62	13.1489
US - Text 16	538	235	0.437	15.37	134.5	5.81	35	4	8.75	14	51	13.6227
Assn - Text 2	309	172	0.557	20.6	44.14	5.28	15	7	2.14	13	59	13.7461
Stewart - Text 6	294	186	0.633	16.33	147	5.76	18	2	9	15	52	13.8752
Mdivt - Text 7	1516	546	0.360	18.72	116.62	5.51	81	13	6.23	17	57	13.8827
Xtch - Text 12	787	378	0.480	20.18	78.7	5.49	39	10	3.9	14	46	14.5378
EEO- Text 3	377	209	0.554	18.85	62.83	5.64	20	6	3.33	16	37	14.5434
Enrcm - Text 14	2309	764	0.331	20.8	85.52	5.71	111	27	4.11	16	48	15.8581
NASA - Text 17	520	247	0.475	23.64	65	5.47	22	8	2.75	15	38	16.1392
Lions - Text 15	1151	534	0.464	25.02	164.03	5.46	46	7	6.57	15	65	16.7792
Consult - Text 9	364	184	0.505	28	91	5.6	13	4	3.25	15	54	18.94
Magna - Text 10	718	319	0.444	35.9	143.6	5.53	20	5	4	15	101	22.589
Total	12715	5709	10.123817	372.97	1903.02	106.88	684	137	108.02	291	996	261.2871
Average	636	285	0.506	18.65	95.15	5.34	34.2	6.9	5.40	14.55	49.80	13.06
Standard Deviation	515.6	164.3	0.095	6.1	35.1	0.4	24.6	5.4	2.2	1.5	15.7	4.3

Note that some of the statistics are in a narrow band and others are wide. The number of words per release was within a standard deviation of more than 1100 words to barely 100. The ratio of different words to total words in lexical density was in a ratio of one new word for every two with a variation of ten percent. Average words per sentence ranged from 12 to 24.

Average characters per word were five without much variation. Average sentences per paragraph ranged from three to seven. The readability index ranged from a low of four to a high of 22.

The next step was to look at the releases themselves and to evaluate them – admittedly subjectively – for their writing style. Among the four most readable, based on the statistics:

- Text 11 was dense and difficult to understand since it was reporting the results of a mine testing program.
- Text 5 was fluffy and self-congratulatory emoting from an entertainment group that was not written in news style but was easily understandable.
- Text 19 was an official notice of a settlement over an environmental case written in straightforward news style that was easily understandable.
- Text 18 was a personnel announcement written in over-hyped style of the entertainment industry but easy to digest.

On the other end of the spectrum, the least readable releases, I found the following:

- Text 10 was a financial release filled with jargon.
- Text 9 was an inconsequential announcement written in run-on sentences.
- Text 15 was an acquisition announcement from the entertainment industry (again) that ran well beyond any news value.
- Text 17 was a release from the National Aeronautic and Space Administration (NASA) that was easily understandable and interesting.

It is clear that readability statistics have little or nothing to do with the intrinsic value of text itself. Good text can be dense and difficult to decipher or easy to read. The only thing that statistics can do is to limn basic parameters for how text was written. Reflecting on the statistics, if there were news value in each of the releases, chances are that the four ranked as most readable

would be indeed be that way except for the first one. The four ranked as the least readable belonged there except for the NASA announcement.

The final result of this modest experiment is that one should use readability statistics with care. They are not that meaningful. But, chances are you knew that anyway.

#