Comparisons of Various Text Complexity Equations

Hal Burdick
New Technologies, MetaMetrics, Inc.

Several comparisons of the Lexile Theory to other readability formulas have been researched since its inception 25 years ago. The first study was conducted in the late 1980s where the passages from five reading tests were analyzed with ten different readability formulas. Table 1 shows these correlations.

Table 1. *Correlations between Empircal & Theoretical Item Difficulties for Ten Readability Formulas.*

Ten Readability Equations	Lexile Native Items	SRA Passage Items	Battery Sentence Items	Mastery Cloze Items	Peabody Picture Items
Lexile	.90	.92	.85	.74	.94
Flesch	.85	.94	.85	.70	.85
ARI	.85	.93	.85	.71	.85
FOG	.85	.92	.73	.75	.85
Powers	.82	.93	.83	.65	.74
Holquist	.81	.91	.81	.84	.86
Flesch-1	.79	.92	.81	.61	.69
Flesch-2	.75	.87	.70	.52	.71
Coleman	.74	.87	.75	.75	.83
Dale-Chall	.76	.92	.82	.73	.67

Adapted from Wright & Stenner, 1999

Readability formulas in general were strong predictors of the difficulty of reading comprehension test items of various item types with correlations ranging from .52 to .94. The Lexile theory tended to exceed the correlations of the other formulas.

A more recent study in 2005 looked at both the correlations of readability formulas with empirical item difficulties and the correlations of readability formulas with each other. The results for the 29 items investigated are displayed in Table 2.

Table 2. *Correlation Matrix of Empirical Item Difficulties and Eight Different Readability Formulas.*

	Empirical	Lexile	Dale Chall	Flesch	FOG	Powers	SMOG	Forcast	Spache
Empirical	1	0.91	0.83	0.86	0.86	0.86	0.84	0.65	0.87
Lexile	.	1	0.81	0.90	0.89	0.88	0.84	0.67	0.89
Dale Chall	.	.	1	0.87	0.88	0.88	0.88	0.67	0.93
Flesch	.	.	.	1	0.99	0.98	0.94	0.73	0.97
FOG	.	.	.	.	1	0.97	0.95	0.68	0.96
Powers	.	.	.	.	.	1	0.94	0.82	0.95
SMOG	.	.	.	.	.	.	1	0.67	0.91
Forcast	.	.	.	.	.	.	.	1	0.71
Spache	.	.	.	.	.	.	.	.	1
Mean	712	747	5.9	5.7	7.9	5.0	7.5	8.8	3.6
SD	368	433	1.7	4.4	4.8	1.2	3.9	1.3	1.6

The Lexile Theory again showed the highest predictive power on empirical item difficulties at .91. Not surprisingly, the various readability formulas were highly correlated with a range of .65 to .99.

References

Wright, B. D. & Stenner, A. J. (1999). One Fish, Two Fish: Rasch Measures Reading Best. *Popular Measurement* (1), 34-38.