Short Report

Psychological Science
2018, Vol. 29(7) 1178–1184

DOI: 10.1177/0956797618759465
www.psychologicalscience.org/PS
SAGE

Are Atypical Things More Popular?

Jonah Berger[1] and Grant Packard[2]
[1]The Wharton School, University of Pennsylvania, and [2]Lazaridis School of Business and Economics, Wilfrid Laurier University

Abstract
Why do some cultural items become popular? Although some researchers have argued that success is random, we suggest that how similar items are to each other plays an important role. Using natural language processing of thousands of songs, we examined the relationship between lyrical differentiation (i.e., atypicality) and song popularity. Results indicated that the more different a song's lyrics are from its genre, the more popular it becomes. This relationship is weaker in genres where lyrics matter less (e.g., dance) or where differentiation matters less (e.g., pop) and occurs for lyrical topics but not style. The results shed light on cultural dynamics, why things become popular, and the psychological foundations of culture more broadly.

Keywords
popularity, natural language processing, cultural success, psychological foundations of culture, music

Received 8/28/17; Revision accepted 1/18/18

Why do some things catch on? Academics and popular writers alike have long been interested in cultural dynamics, or why some songs, movies, and other cultural items become popular (Lieberson, 2000; Simonton, 1980). Some songs become hits while others fail, and some movies become blockbusters while others do not. What makes some cultural items successful?

One possibility is that popularity is random. Even domain experts have difficulty predicting success (Bielby & Bielby, 1994), and researchers have argued that popularity is driven by chance patterns of social influence (Salganik, Dodds, & Watts, 2006). These perspectives suggest that success has little to do with features of cultural items themselves.

Building on research on the psychological foundations of culture, however, we suggest that the similarity between items may help shape success. Research on cross-cultural psychology has demonstrated the influence of culture on individual-level psychological processes (Markus & Kitayama, 1991). But the reverse is also true; psychological processes shape the norms, practices, and items that make up culture (Akpinar & Berger, 2015; Berger & Milkman, 2012; Kashima, 2008; Norenzayan, Atran, Faulkner, & Schaller, 2006; Schaller & Crandall, 2004). In particular, people have a drive for stimulation (Zuckerman, 1979), and novelty can increase evaluation (Berlyne, 1970). This suggests that similarity between concepts should shape cultural success. The things people have experienced should determine how novel a cultural item seems. Taken to a collective level, cultural items that are more atypical, or differentiated from other cultural items, may be liked more and, consequently, become more popular.

Unfortunately, empirically testing such propositions has been constrained by the ability to quantify differences among cultural items at scale. To address this issue, we used textual analysis, measuring lyrical differentiation across thousands of songs. Importantly, the differentiation we examined is bounded, not infinite. A country song with lyrics that sound like death-metal lyrics would be different from most country songs but would also be unlikely to be classified as country in the first place. Thus, we examined whether among songs classified as belonging to a given cultural category (i.e., genre), those whose lyrics are more atypical are more successful. Such songs are different from the prototype but not so different as to be outside the

Corresponding Author:
Jonah Berger, The Wharton School, University of Pennsylvania, 700 Jon M. Huntsman Hall, 3730 Walnut St., Philadelphia, PA 19104
E-mail: jberger@wharton.upenn.edu

genre, similar to Boyer's (1994) notion of minimally counterintuitive concepts.

Method

First, we collected data on song popularity. To focus on individuals' preferences, we used Billboard's digital download rankings (www.billboard.com/biz), which capture more than 90% of major paid song services (e.g., Apple iTunes and Google Play). We focused on this measure of popularity, rather than, say, radio airplay, because it is more likely to be driven by individual preferences than by a small number of institutionalized actors (e.g., DJs). We sampled the ranking data once every 3 months over a 3-year period (2014–2016) for each of seven major genres (Christian, country, dance, pop, rap, rock, and rhythm and blues). Comprehensive data were unavailable for the alternative genre. We obtained all songs that appeared in each genre ranking and their position in that genre's chart (1–50). We reverse-coded song ranks so that positive coefficients described a positive relationship with song success. This resulted in a data set of 4,200 song rankings and 1,879 unique songs. We captured artist name and whether the song appeared on the Billboard radio airplay lists for the same periods as covariates.

Second, we acquired the complete lyrics for each of these songs at SongLyrics.com. Third, we used latent Dirichlet allocation (LDA; Blei, 2012) to determine the main themes discussed across songs. This approach takes texts (e.g., song lyrics) and by measuring word co-occurrence within and across texts, determines the latent topics or themes that make up those texts and the words that make up each topic (see Table 1; for more detail on the methodological approach and potential shortcomings, see the Supplemental Material available online). Aggregating across all songs within a genre provides that genre's average topic composition (see Fig. 1). Country songs, for example, feature a lot of lyrics about girls and cars (39%) and less about body movement (2%).

Table 1. Topics and Demonstrative Topic Words

Topic	Example topic words
Anger and violence	bad, dead, hate, kill, slay
Body movement	body, bounce, clap, jump, shake
Dance moves	bop, dab, mash, nae, twerk
Family	American, boy, daddy, mamma, whoa
Fiery love	burn, feel, fire, heart, love
Girls and cars	car, drive, girl, kiss, road
Positivity	feel, like, mmm, oh, yeah
Spiritual	believe, grace, lord, one, soul
Street cred	ass, bitch, dope, rich, street
Uncertain love	ain't, can't, love, need, never

Note: Longer lists of high-probability words by topic are presented in Table S2 in the Supplemental Material available online.

Fourth, we calculated how lyrically different each song is from its genre. For each topic, we took the absolute value difference between that song's lyrical topic composition and the genre mean. Then, we aggregated these differences across topics using Ireland and Pennebaker's (2010) language style matching equation. We inverted the resultant value to describe differentiation (rather than matching).

Finally, we used ordinary least squares regression to examine the relationship between lyrical differentiation and song performance. An analysis treating rank dependent measures as continuous was appropriate given the large number and fixed range of ranks.

Results

Results indicated that the more differentiated a song's lyrics were from its genre, the more popular that song was, $b = 6.45$, $t(4198) = 3.23$, $p = .001$ (see Fig. 2 and Table 2, Model 1). A 16% increase in lyrical differentiation, for example, was associated with a one-position improvement in chart ranking. Results were the same using an ordinal logistic specification (proportional odds), estimate = 0.77, $t(4198) = 3.22$, $p = .001$, or log transformation of the rank dependent measure, $b = 0.18$, $t(4198) = 3.42$, $p < .001$. Results were also the same using alternate methods of calculating lyrical differentiation, such as squared (rather than absolute value) differences, $b = 43.08$, $t(4198) = 2.83$, $p = .005$, and Jensen-Shannon divergence, $b = 155.14$, $t(4198) = 3.10$, $p = .002$.

Robustness checks

We included numerous covariates in the model to assess the stability of the main results and rule out alternate explanations. Even when we controlled for a range of factors, including radio airplay, artist, time, and the topics themselves, the effect of lyrical differentiation remained significant, $b = 8.01$, $t(4175) = 3.27$, $p = .001$ (see Table 2, Model 2). We also considered other factors pertaining to lyrical content. One might wonder whether things such as number of words, language complexity, or other major linguistic features not captured by our LDA approach could explain the results. To account for these factors, we ran a model adding word count, six-letter words (a proxy for language complexity), and baskets of words empirically linked to social or psychological constructs (cognitive processing, emotion, sociality, perception, motivation, time,

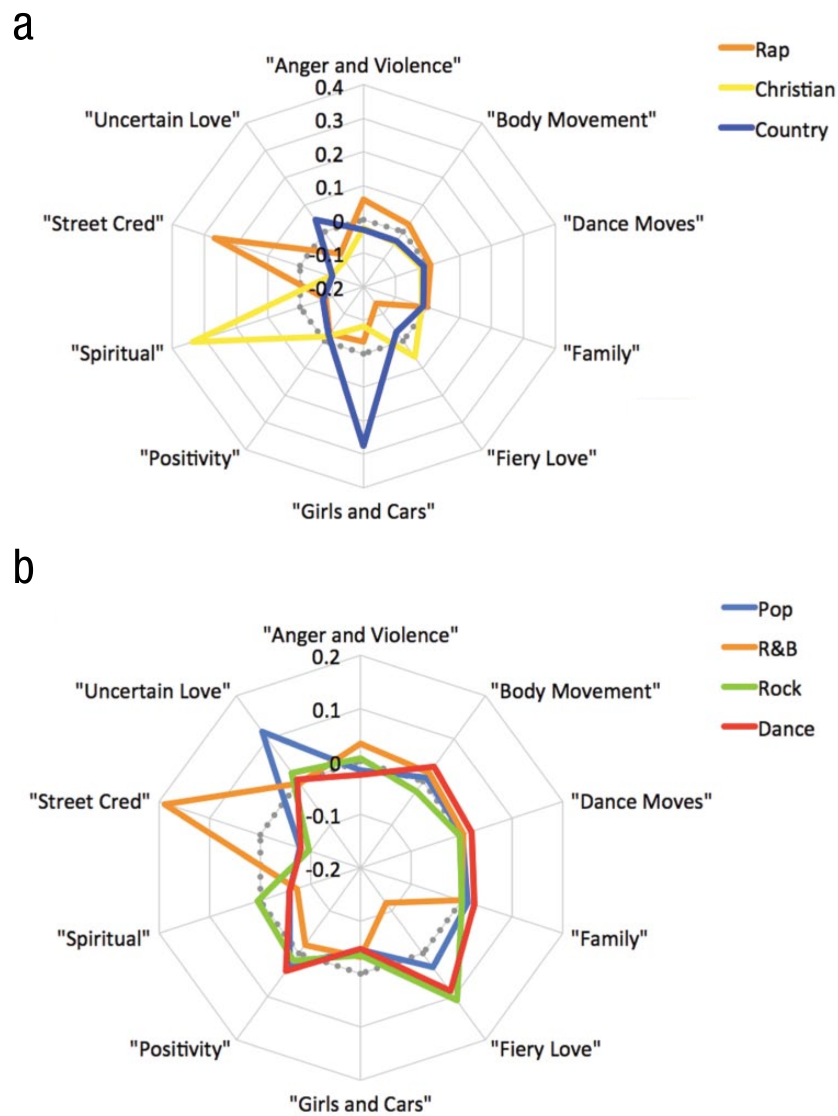

Fig. 1. Relative use of lyrical topics by genre. Values indicate the difference between a genre's topic use and the average topic use across all songs, with zero (gray dotted line) indicating no difference. Country songs, for example, contain 27% more lyrics about girls and cars than the average song. Panel (a) depicts genres with more extreme variation in topic use, and panel (b) depicts genres with less extreme variation in topic use.

relativity, and formality) from the Linguistic Inquiry and Word Count (LIWC) dictionaries (Pennebaker, Boyd, Jordan, & Blackburn, 2015). However, even after these factors were included, lyrical differentiation remained significant, $b = 8.38$, $t(4164) = 3.38$, $p < .001$ (see Table 2, Model 3).

We also examined whether song success is driven by similarity to other genres. For example, one could argue that atypical songs are more popular because they are more similar to other genres and thus appeal to a broader set of people, rather than because they are different from their own genre. We tested this alternative in two ways. First, we created a variable capturing a song's average lyrical similarity (i.e., the inverse of differentiation) to the genres to which it was not assigned. When this new variable was included as a predictor of song success in the main model, it was nonsignificant, $b = 1.30$, $t(4197) = 0.31$, $p = .76$, whereas lyrical differentiation from a song's own genre remained significant, $b = 7.42$, $t(4197) = 1.99$, $p < .05$. Second, we created separate variables capturing a song's similarity to each of the six genres in which it was not assigned and included these as simultaneous predictors of song success alongside lyrical differentiation (from a song's

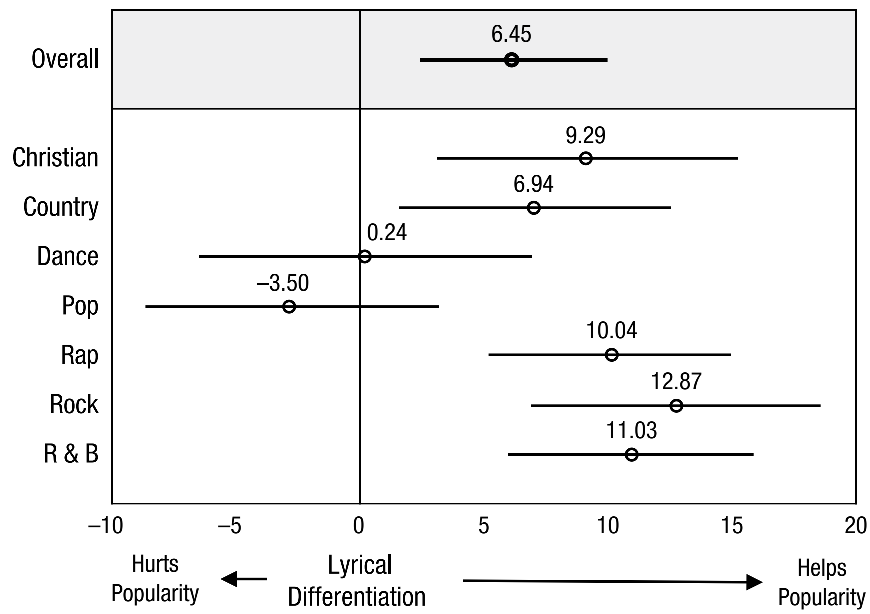

Fig. 2. Impact of lyrical differentiation across genres. Unstandardized coefficient estimates and 95% confidence intervals from the main ordinary least squares regression are shown for all songs and for each genre separately. The effect is significant for a genre if the confidence interval does not include zero. In some specifications (e.g., model 2 including artist, song, and time covariates, when using Jensen-Shannon divergence), pop has a significantly negative coefficient, suggesting that lyrical differentiation may sometimes hurt pop music success. R & B = rhythm and blues.

own genre). Lyrical differentiation again predicted success, whereas a song's similarity to other genres was nonsignificant (see Table S1 in the Supplemental Material). These analyses cast doubt on the possibility that similarity to other genres was driving success.

We also considered whether individual words, rather than the bundles of words used in our LDA-based approach or the LIWC dictionaries, might explain the results. However, even after dummies were included for the presence of each of the 100 most frequent words used across all songs, the effect for lyrical differentiation remained significant, b = 8.08, $t(4064)$ = 2.52, p = .01 (see Table 2, Model 4). In sum, across a variety of specifications, lyrically differentiated songs were more popular. Results also remained the same when we controlled for topical diversity (entropy) and alternate approaches to calculating topics (e.g., within genre) or differentiation (i.e., across time; see the Supplemental Material).

Identifying the effect

To best identify lyrical differentiation's effect, one would ideally keep all other song aspects the same, vary lyrical differentiation, and examine its influence on popularity. To approximate this, we examined songs that charted in two different genres at the same time. Whereas artist, lyrics, and all other song features are identical (it is the same song), lyrical differentiation will be greater in one genre than the other, providing a stricter test of lyrical differentiation's impact. If the same song by the same artist is more successful in the genre in which it is more lyrically differentiated (vs. one in which it is less differentiated), this would support the notion that lyrical differentiation, rather than some other factor, is driving popularity.

To test this possibility, we analyzed the 410 songs that appeared on two different genre charts at the same time using an analysis approach similar to difference-in-differences analysis. We calculated the lyrical differentiation of each song from each of its two genres and took the difference of these values.

Results underscored the prior findings, indicating that songs are more popular in genres in which they are more lyrically differentiated, b = 34.64, $t(408)$ = 3.00, p = .003. This, combined with additional analyses (see the Supplemental Material), including using a matched comparison group of less popular (i.e., nonranked) songs by the same artist from the same album, a model accounting for right truncation, and a two-stage Heckman selection model, cast doubt on the notion that selection can explain the results.

Table 2. Results From the Models Testing the Link Between Atypicality and Song Ranking

Variable	Model 1	Model 2	Model 3	Model 4
Lyrical differentiation	6.45** (1.99)	8.01** (2.45)	8.38*** (2.48)	7.84** (2.66)
Times charted		0.83*** (0.10)	0.69*** (0.11)	0.75*** (0.11)
Multigenre count		4.79*** (0.80)	4.58*** (0.69)	5.24*** (0.77)
Radio airplay		11.17*** (0.55)	11.10*** (0.55)	11.09*** (0.56)
LIWC dictionaries				
Word count			0.00 (0.00)	0.00 (0.00)
Six-letter words			0.02 (0.07)	–0.02 (0.07)
Cognitive words			–0.06 (0.07)	0.03 (0.08)
Affect words			0.00 (0.08)	0.05 (0.08)
Social words			–0.07 (0.05)	–0.01 (0.06)
Perceptual words			0.02 (0.08)	0.18 (0.12)
Motivation words			–0.02 (0.06)	–0.03 (0.06)
Temporal words			–0.05 (0.05)	–0.06 (0.06)
Relativity words			0.00 (0.05)	–0.05 (0.05)
Swear words			0.15 (0.20)	0.06 (0.22)
Control				
Artist/song	No	Yes	Yes	Yes
Topic	No	Yes	Yes	Yes
Time	No	Yes	Yes	Yes
Top 100 words	No	No	No	Yes
Intercept	23.34*** (0.70)	38.25*** (1.75)	39.95*** (2.48)	33.03*** (6.40)
Adjusted R^2	.023			
Marginal R^2		.142	.146	.175
Conditional R^2		.344	.347	.367

Note: Values given are unstandardized regression coefficients. Song ranking was reverse-coded. LIWC = Linguistic Inquiry and Word Count.
$^{**}p < .01$. $^{***}p < .001$.

Variation by genre

If lyrics are shaping success, as we suggest, then one might imagine these effects would vary by genre. Lyrics may matter less in dance music, for instance, where attributes that drive movement (e.g., the beat) may be more important than lyrics. Results were consistent with this prediction (see Fig. 2). A partitioned regression model showed that whereas lyrical differentiation was linked to popularity in most genres, the relationship was weaker in dance as well as pop music, which, almost by definition, is more about mainstreaming than differentiation (Frith, 1986).

Type of atypicality

Further analyses shed light on the type of differentiation linked to success. For example, are country songs with more differentiated lyrics more successful because they include more country-associated content (e.g., more "girls and cars" and other genre-typical lyrics than other country songs) or because they include less country-associated content?

Topics were ordered on the basis of the degree to which their use within a genre deviated from their use across genres. Then, for the five most and least typical topics for each genre, we separately calculated and aggregated directional (rather than absolute) lyrical differentiation. Including these two variables as predictors in the base specification indicates that successful songs use less of the more genre-typical topics, $b = -3.60$, $t(4196) = 3.28$, $p = .001$. In terms of what takes their place, use of more of the least genre-typical topics such as uncertain love, $t(4189) = 2.78$, $p = .005$, and dance movement, $t(4189) = 2.69$, $p = .007$, is generally linked to greater popularity. Less-used topics linked to success also vary by genre. For example, "street cred" features in only 2% of rock lyrics, yet rock songs that contain more lyrics about it are more popular, $b = 35.65$, $t(589) = 2.30$, $p = .02$. Additional results are reported in the Supplemental Material.

Topical versus style differentiation

Although topically differentiated songs are more popular, stylistic differentiation is another matter. Linguistic topic (or content) refers to what someone is discussing (e.g., cars, love, or money), but linguistic style refers to the small subset of words (e.g., prepositions and conjunctions) that relate to how a person writes or speaks (Ireland & Pennebaker, 2010). Stylistic differentiation, however, does not predict success, $b = 0.11$, $t(4198) = 0.05$, $p = .96$ (see the Supplemental Material). Further, the relationship between topical differentiation and popularity persisted even after we controlled for stylistic differentiation. This suggests that successful songs tend to be about different topics but not necessarily in a different style (although specific stylistic word features, on their own, may be linked to popularity).

Discussion

Although some researchers have argued that cultural success is impossible to predict, textual analysis of thousands of songs suggests that those whose lyrics are more differentiated from their genres are more popular. This dovetails with recent perspectives on the psychological foundations of culture (Kashima, 2008; Schaller & Crandall, 2004). When shared across individuals, psychological processes shape the practices and items that constitute culture. In this case, value for novelty or difference may underlie the link between lyrical differentiation and cultural success.

Might different aspects (e.g., instrumentation vs. lyrics or cinematography vs. script) play different roles in cultural adoption? In songs, for example, would melodic differentiation also be beneficial? Although extrapolation from our findings might suggest that the answer is "yes," elements such as instrumentation or melody may be particularly important in determining how people classify a song (e.g., the sound of a banjo signals country music). If so, lyrics may be freer to deviate from genre norms.

Different aspects of cultural products (e.g., music and lyrics) may also combine to shape differentiation and, thus, popularity. Although novelty can be good, there are also benefits of familiarity (Kunst-Wilson & Zajonc, 1980). Rather than achieving the right balance on one dimension alone, successful items may mix similarity and differentiation across dimensions (Berger, Bradlow, Braunstein, & Zhang, 2012)—similar enough to evoke the warm glow of familiarity but differentiated enough to feel new and exciting. Remixes, for example, add new lyrics to an old tune or old lyrics to a new tune. Differentiation on one dimension may be balanced by similarity on another.

As with any investigation, boundaries apply. Although we believe our results illustrate a broader pattern, we cannot speak outside of the years we examined. Future work might examine whether these effects generalize across cultures. For example, East Asians prefer less social differentiation, which might extend to preferences for novelty versus typicality. Similarly, other genres (e.g., opera) may show different patterns based on norms associated with them.

Finally, our findings highlight the value of natural language processing to study cultural dynamics. Advances in computational social science provide rich opportunities to extract cultural and behavioral insight from a variety of texts (e.g., books, news, and social media). Hopefully, this emerging tool kit will help provide deeper insight into why things catch on.

Action Editor

Philippe G. Schyns served as action editor for this article.

Author Contributions

Both authors contributed equally to this research and are listed alphabetically. Both authors designed the study and collected the data. G. Packard analyzed and interpreted the data with help from J. Berger. J. Berger drafted the manuscript, and G. Packard provided critical revisions. Both authors approved the final manuscript for submission.

Acknowledgments

The authors thank Shawn Zamechek for assistance with the latent Dirichlet allocation and Amir Goldberg, Oded Netzer, and Michael Macy for helpful feedback.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/0956797618759465

Open Practices

Data and analysis materials have not been made publicly available. Plans for analysis of the field data were not preregistered.

References

Akpinar, E., & Berger, J. (2015). Drivers of cultural evolution: The case of sensory metaphors. *Journal of Personality and Social Psychology*, *109*, 20–34.

Berger, J., Bradlow, E., Braunstein, A., & Zhang, Y. (2012). From Karen to Katie: Using baby names to study cultural evolution. *Psychological Science*, *23*, 1067–1073.

Berger, J., & Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research*, *49*, 192–205.

Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception and Psychophysics*, *8*, 279–286.

Bielby, W. T., & Bielby, D. D. (1994). All hits are flukes: Institutionalized decision making and the rhetoric of network prime-time program development. *American Journal of Sociology*, *99*, 1287–1313.

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, *55*, 77–84.

Boyer, P. (1994). *The naturalness of religious ideas: A cognitive theory of religion*. Berkeley: University of California Press.

Frith, S. (1986). Why do songs have words? *The Sociological Review*, *34*, 77–106.

Ireland, M. E., & Pennebaker, J. W. (2010). Language style matching in writing: Synchrony in essays, correspondence, and poetry. *Journal of Personality and Social Psychology*, *99*, 549–571.

Kashima, Y. (2008). A social psychology of cultural dynamics: Examining how cultures are formed, maintained, and transformed. *Social and Personality Psychology Compass*, *2*, 107–120.

Kunst-Wilson, W. R., & Zajonc, R. B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science*, *207*, 557–588.

Lieberson, S. (2000). *A matter of taste: How names, fashions, and culture change*. New Haven, CT: Yale University Press.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224–253. doi:10.1037/0033-295X.98.2.224

Norenzayan, A., Atran, S., Faulkner, J., & Schaller, M. (2006). Memory and mystery: The cultural selection of minimally counterintuitive narratives. *Cognitive Science*, *30*, 531–553.

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Retrieved from https://repositories.lib.utexas.edu/bitstream/handle/2152/31333/LIWC2015_LanguageManual.pdf

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006). Experimental study of inequality and unpredictability in an artificial cultural market. *Science*, *311*, 854–856.

Schaller, M., & Crandall, C. S. (2004). *The psychological foundations of culture*. Hillsdale, NJ: Erlbaum.

Simonton, D. K. (1980). Thematic fame, melodic originality, and musical zeitgeist: A biographical and transhistorical content analysis. *Journal of Personality and Social Psychology*, *38*, 972–983.

Zuckerman, M. (1979). *Sensation seeking: Beyond the optimum level of arousal*. Hillsdale, NJ: Erlbaum.