Melvill-Smith, Sorcha, Krause, Amanda E., and North, Adrian C. (2023) ***Song popularity and processing fluency of lyrics.*** **Psychology of Music, 51 (3) pp. 986-1000.**

Published citation:

Melvill-Smith, Sorcha, Krause, Amanda E., & North, Adrian C. (2023) Song popularity and processing fluency of lyrics. *Psychology of Music*, *51* (3), 986-1000. https://doi.org/10.1177/03057356221118400

Song Popularity and Processing Fluency of Lyrics

Corresponding author:
Prof Adrian C North, Curtin University, Kent Street, Bentley, WA 6120, Australia. Email: adrian.north@curtin.edu.au

Abstract

Several studies show that a concept or object is more popular when it is easier to process. The present research applies this notion of processing fluency to the lyrics of all 271 top 5 songs on the United Kingdom chart for each week from 1999-2014. The processing fluency of the lyrics was computer scored for readability, presence of rhyme, and complexity, and popularity was assessed in terms of peak chart position and duration of tenure on the chart. After controlling for the energy of the musical component of the song, analyses showed that factors relating to the processing fluency of the lyrics predicted peak popularity but not duration of chart tenure. Significant relationships were observed between peak popularity and both rhyme saturation and basic lyric readability properties.

Keywords: Processing fluency, lyrics, popularity, readability, rhyming

Song Popularity and Processing Fluency of Lyrics

Several theories indicate that the subjective ease with which perceptual information is processed is related positively to the popularity of the object or concept in question. The notion of processing fluency is arguably the most prominent account of this in the literature (Oppenheimer, 2008), and research has considered this in the context of media such as advertisements, visual art, and text (e.g., Leder, 2013; Mayer & Landwehr, 2018; Schwarz, 2004), and in consumer behaviour (e.g., Aydin, 2018; Liebers et al., 2019). On this basis, the present research investigated whether the processing fluency of song lyrics can account for the commercial popularity of those songs.

Processing fluency has previously been measured using a variety of methods, including media priming, visual clarity, and language complexity (e.g., Alter & Oppenheimer, 2009; Im et al., 2010; Schwarz, 2004; Winkielman & Cacioppo, 2001). For example, Song and Schwarz (2009) found that instructions written in easier-to-read fonts (e.g., Calibri, 12pt) were related to participants' perceptions of their competence to carry out these instructions, and Hansen et al. (2008) found that statements written on paper with higher colour contrast to the font, which makes them easier to read, were regarded as more trustworthy than those printed on paper with lower colour contrast. This ease of processing also has positive implications for various aspects of people's evaluations and affective responses to the stimuli in question. For instance, Winkielman and Cacioppo (2001) found that when images were visually primed and exposed for a longer time, participants processed them more easily and enjoyed them more than images that were either not primed or exposed for a shorter time; Song and Schwarz (2009) found that food additives that were harder to pronounce were perceived as less familiar and riskier; and Dragojevic and Giles (2016) found that voices that were easier to understand gave rise to more positive emotional reactions.

A smaller number of studies show specifically that processing fluency is also associated with consumer's retail decision-making. For example, Im et al. (2010) found that improved image quality on online shopping sites lead to greater customer satisfaction. Herrmann et al. (2013) found that the use of basic scents (one scent note) in comparison to complex scents (multiple scent notes) correlated with the amount of money spent in a store. Jaud and Melnyk (2020) found that wine bottle labels with pictures were processed more easily than labels with text only, and were also associated with higher purchase intent and perceived product quality.

The recent growth of media streaming and subscription models represents a fundamental shift in how businesses approach advertising as well as audio, text-based, and visual mass media. Therefore, it is unsurprising that there is renewed research interest in predicting consumers' responses to these media (Raza & Nanath, 2020). The findings of this are consistent with those obtained in other domains, namely that improving processing fluency by using priming, specific literary devices, and ease of comprehension are all associated with greater liking and preference judgements (Filkukova & Klempe, 2013; Liebers et al., 2019; Newman & Schwarz, 2018). Aydin (2018) found a significant relationship between processing fluency and increased purchase intention in response to advertising. Liebers et al. (2019) found that when participants saw competing product commercials, conceptual priming of selected products resulted in more favourable judgements and increased purchase intent. Filkukova and Klempe (2013) explored rhyming as a measure of advertisement slogan processing fluency: rhyming slogans were perceived as more likeable, original, easy to recall, trustworthy, and convincing than non-rhyming slogans. With regard to other media, Newman and Schwarz (2018) showed that poorer quality radio broadcast audio was associated with the material being perceived as less fluent and less pleasant, and with the presenter being judged as less intelligent. Various other studies have demonstrated that readability scores are associated positively with the popularity of text-based media (Basyuk, 2018; Gao et al., 2019 Pancer et al., 2019; Wright, 2003). For instance, Pancer et al. (2019) observed a positive relationship between the readability of social media postings and subsequent popularity and engagement (as measured by the number of shares and likes).

A number of other recent studies have attempted to predict the sales chart performance of songs based on various structural and acoustic features of the music (e.g., Askin and Mauskapf, 2017; Herremans et al., 2014; Interiano et al., 2018; Ren and Kauffman, 2017; and Zangerle et al., 2020) and lyrics (Dhanaraj and Logan, 2005; Putter et al., 2021; Raza and Nanath, 2020). However, two other studies of music popularity have considered variables similar to processing fluency. Margulis (2013) explored participants' aesthetic responses to varying repetitive sequences of classical musical samples. Repetitive samples were judged as more interesting, enjoyable, and artistic than the non-repetitive samples. Bannister (2018) analysed data from the top singles on the Billboard Hot 100 from 1958 to 2018. Current top-performing songs featured time signatures and key signatures that were typical of past top-performing songs, suggesting that audience familiarity with these features may be a factor underlying song popularity. Findings such as these are consistent with Martindale et al. s (1988) cognitive theory of aesthetics, which suggests that liking for naturalistic artistic objects is related positively to their meaningfulness, which has been operationalised as how easily the person can interpret the stimulus in question. Several researchers have provided supporting evidence (Martindale et al., 1990; North et al., 2017; Nunes et al., 2015; Whitfield, 1983), which again implies that ease of processing is associated with greater liking.

In the light of the evidence reviewed here, it is unsurprising that a handful of studies suggest indirectly that the processing fluency of music lyrics can predict the popularity of those songs. North, Krause, and Ritchie (2020) considered 1414 sets of lyrics which were scored according to the extent to which each was typical of the corpus. The typicality of the lyrics was related positively to the chart performance of the songs, and while typicality does not necessarily equate with processing fluency it is arguable that the two concepts are related. In a similar vein, Nunes et al. (2015) demonstrated that lexical repetition (the repetition of words without regard for grammar) in song lyrics influenced a song's likelihood of reaching the top of the Billboard Hot 100 singles chart. Increased repetition, and thus potentially greater processing fluency, was associated with greater song popularity. Similarly, Hahn et al. (2018) found that pre-literate children preferred rhymed song lyrics over non-rhyming song lyrics; and Singhi and Brown (2015) showed that the rhyming features of lyrics could predict hits better than could musical features. Schedl (2019) investigated the relationships between objective measures of processing fluency (repetitiveness, readability, and duration) and the popularity of song music and lyrics, showing that popular songs are repetitive but not always simple regarding readability. Varnum et al. (2021) reported that lyrical simplicity was related positively to the novelty of the music in a song and to the popularity of that song. Varnum et al. (2021) propose that as more novel music is produced, the lyrics of these songs become increasingly simple in order to compensate for the greater processing load imposed by the music. In short, several studies seem to indicate that various aspects of the processing fluency of lyrics are implicated in the popularity of the song in question.

The need for research concerning processing fluency and music lyrics is further indicated by several studies which show that the specific nature of the impact of processing fluency is influenced by several properties common to music lyrics. These studies highlight the need to determine whether processing fluency is related to listener responses to lyrics, and if so what form this relationship might take when different forms of processing fluency are considered. For instance, Miall and Kuiken (1994) showed that rhyme and alliteration increased reading time, and so by implication appear to hamper processing, but also improved affective responses. The authors argue that there are several explanations of this, such as rhyme encouraging the perceiver to reflect on the text (increasing reading time) but also understand the emotional connotations better (increasing affective responses). In a similar vein, Menninghaus et al. (2015) found that rhymes facilitate prosodic processing but hamper semantic comprehension, so that it may be appropriate to distinguish between perceptual (or prosodic) fluency and conceptual (or semantic) fluency.

Schotanus et al. (2015) found that the popularity of psalms could be predicted by balanced motion (e.g., line length, lines per stanza) and the degree of repetition they contained, and by implication their processing fluency. However, the authors also note that some aspects of balanced motion and repetition were related to popularity more closely than were others. In conjunction, findings such as these indicate the need to consider different forms of processing fluency when studying artistic products, and suggest that different measures of processing fluency may give rise to different relationships with popularity. Similar arguments are made by Schotanus (2021) concerning analysis of the popularity of two Bob Dylan songs.

Previous research findings indicate that specific aspects of what we might collectively term processing fluency can predict positive affective responses and commercial popularity, and that lyrics contribute to song popularity. Moreover, several variables indicative of processing fluency may underlie the relationship between lyrics and song popularity, although there is also evidence that different measures of processing fluency may give rise to differing relationships with popularity and enjoyment. The current research therefore aimed to directly assess the relationship between various manifestations of lyric processing fluency and song popularity. Various aspects of rhyming, readability, and word complexity were employed as measures of the processing fluency of song lyrics due to their prominence in the literature. Four measures of the songs' commercial success were used: 'peak popularity' represented the song's peak chart position and 'hit duration' represented the duration of the song's tenure on the charts, and values for both were employed for the United Kingdom only and also for the United Kingdom and United States combined. There were four hypotheses.

H1) There will be a positive relationship between various aspects of the processing fluency of the lyrics and UK - US peak popularity.

H2) There will be a positive relationship between various aspects of the processing fluency of the lyrics and UK - US hit duration.

H3) There will be a positive relationship between various aspects of the processing fluency of the lyrics and UK peak popularity.

H4) There will be a positive relationship between various aspects of the processing fluency of the lyrics and UK hit duration.

To test these hypotheses, this research utilised a computerised dataset of 271 popular songs from 1999 to 2014, accounting for all those songs that reached the top five singles sales charts in the United Kingdom chart each week throughout the time period in question.

Method

Research Design

The research utilised a correlational design. Six predictor variables were used to capture distinct aspects of processing fluency, namely Rhyme Saturation, Multisyllabic Rhymes, Short Unique Rhymes, Basic Lyric Readability Properties, Computerised Reading Difficulty, and Complexity. Four criterion variables were measured, namely UK-US Peak Popularity, UK-US Hit Duration, UK Peak Popularity, and UK Hit Duration. A significant amount of research in the field of experimental aesthetics shows that various aspects of responses to music are influenced by the extent to which the music possesses so-called arousal potential, energy, or high information content (e.g., Berlyne, 1971; Heyduk, 1975; Vitz, 1966). It is also reasonable to suspect that lyricists write words that will complement the music in question (see North et al, 2020), suggesting that some attempt to control for variations between each song's musical components is warranted. Therefore to control for any confounding effects of the song's acoustic components, a control variable, 'Energy', was utilised.

Dataset and Lyrics

The raw dataset is a subset of a larger dataset used in previous studies (e.g., North, Krause, Sheridan, & Ritchie, 2019; North, Krause & Ritchie, 2020). Music industry professionals catalogued song meta-data such as genre, artist, and record company from over 400,000 record labels to produce this master dataset containing over 38 million pieces of music. The top five songs in the United Kingdom official chart for each week from 1999 to 2014 ($N = 271$; www.officialcharts.com) were employed as the subset of the master dataset used for this research. A song appeared in the dataset only once, even if it spent multiple weeks in the top five. The lyrics were collected from web-based sources (primarily www.azlyrics.com) and were crosschecked with a second internet source to verify consistency. In the instance of discrepancies between versions (often due to multiple versions of songs, such as remixes), a third source was consulted, and the version intended for radio airplay was identified. Various processes were employed to clean the lyrics and so ensure that the computer could interpret them consistently: these involved spell checking; and batch search and replace word processing operations, so that instances of vernacular such as “cudda” were replaced with “could have”, standard contractions were expanded (so that, for instance, “couldn’t” was replaced with “could not”), and instances of “chorus”, “repeat verse”, or “x3” were replaced with the full relevant text.

For the hierarchical multiple regression analyses used in the present research, Tabachnick and Fidell (2013) suggest that the total population (*N*) should ideally be 50 + 8(*k*) when testing a full regression model, and alternatively 104 + *k* when testing individual predictors. It is common to calculate both and use the larger *N* of the two. An a-priori power analysis was conducted using G*Power 3.1 (Faul et al., 2009) based on the findings of North, Krause, Sheridan, and Ritchie (2020), Nunes et al. (2015), and Varnum et al. (2021). The G*power analysis suggested a sample size of 103 based on a medium Cohen s f^2 effect size and a target power of .80. Following Tabachnick and Fidell (2013), 50 + 8(7) = 106 and 104 + 7 = 111, and so the present dataset of 271 songs was deemed sufficient to identify any significant relationships (Cohen, 1988).

Measures

Multiple computerised data analysis programs were utilised to produce processing fluency data for the lyrics, as detailed below.

Rhyming variables. The rhyming aspect of processing fluency was measured using the Rhyme Analyzer program (Hirjee & Brown, 2010). This quantifies different rhyming qualities of songs such as rhymes per line, end rhyme scores, and complex rhymes. A full list of the variables is shown in Appendix 1. Given the number of variables measured, the data for all the Rhyme Analyzer variables were subjected to exploratory principal components analysis with oblique rotation, which identified three factors explaining 44.49% of the variance, namely Rhyme Saturation, Multisyllabic Rhymes, and Short Unique Rhymes (see Appendix 2). Factor scores were calculated and used in subsequent analyses. The three sets of factor scores had small inter-correlations (ranging from $r = -.29$ to .30) indicating that different constructs of rhyme were identified. Rhyme Saturation describes the extent to which rhymes appear in the song. A higher score suggests more rhyming in the lyrics, and thus greater processing fluency. The variables loading onto this factor included rhymes per syllable, line internals per line, chaining per line, rhyme density, rhymes per line, perfect rhymes, and links per line. Multisyllabic Rhymes recorded instances of rhymes with two or more syllables. An example of this is the best rhyme is that/a cat rhymes with hat”. A higher score indicates the presence of more Multisyllabic Rhymes in the lyrics, indicating greater processing fluency. Short Unique Rhymes concerns average word length and the proportion of words in a rhyming pair’s second line that do not appear in the first line. An example of a smaller Short Unique Rhyme would be The cat in a hat/The cat on a mat”, where more similar one-syllable words are present. A larger score for Short Unique Rhymes indicates that the rhymes contain more syllables and unique words, indicating lower processing fluency.

Readability variables. The readability aspect of processing fluency was measured using the online application Readable (https://readable.com). This assesses popular readability formulae such as the

Flesch Reading Ease Score and The Gunning Fog Index, and readability features such as Sentence Count and Letters per Word. A full list of variables is shown in Appendix 3. There is currently no clear consensus on which readability formula is (a) the most valid method of capturing readability and (b) the most appropriate for different research designs (van Oosten et al., 2010; Wang et al., 2013). Given this lack of consensus and the number of variables measured by Readable, the data was subjected to an exploratory principal components analysis with oblique rotation, which produced two factors, Basic Lyric Readability Properties and Computerised Reading Difficulty which explained 45.49% of the variance (see Appendix 4). Factor scores were calculated and used in subsequent analyses. The correlation between the two sets of factor scores was small ($r = .10$) indicating that they assessed different aspects of readability. Basic Lyric Readability Properties captured the presence of conventional readability characteristics in the lyrics. A lower factor score for Basic Lyric Readability Properties indicates that the lyrics are easier to comprehend and require less reading expertise. Thus, a lower score indicates greater processing fluency. The specific variables loading onto this factor included word count, syllable count, words with one syllable, words with six letters, unique word count, and grammar error count. Computerised Reading Difficulty captured overtly computerised, formula-driven measures of the lyrics reading difficulty. A lower factor score for Computerised Reading Difficulty indicates that the lyrics require a lower objective reading level/age, and hence, higher processing fluency.

Complexity. The complexity of the lyrics was assessed using Diction 7.0 (Hart et al., 2013), which has been used previously in several studies (e.g., Cook & Krupar, 2010; see extensive list at http://www.dictionsoftware.com/published-studies). The program measures 36 semantic variables, although only complexity is clearly related to processing fluency, representing a simple measure of the mean number of characters per word. The variable reflects Flesch s (1951) complexity assumption that more complex text increases abstractedness and reduces comprehensibility. A higher score indicates more characters per word and therefore lower processing fluency.

Energy. Energy scores captured the music's acoustic and structural properties, and were used as a control variable. Scores were included in the master data set, and are calculated via a process described in North, Krause, Sheridan, and Ritchie (2018, 2019). In brief, this involves a trained artificial intelligence system analysing each piece of music using a series of algorithms based on music perception research which utilise factors such as beats per minute, tone, and pitch to determine how energetic the music is. Numerous classic theories (e.g., Berlyne, 1971; Osgood, 1952; Russell, 1978) highlight that factors analogous to 'energy' are influenced by structural and acoustic properties of music and underlie evaluative responses across a broad range of domains.

Popularity Scores. The research used two measures of popularity, namely Peak Popularity and Hit Duration, each of which was calculated separately for both the United Kingdom only and United Kingdom and United States together, leading to four variables in total, namely UK - US Peak Popularity, UK - US Hit Duration, UK Peak Popularity, and UK Hit Duration. Greater popularity is indicated by higher scores on each measure. The popularity scores were included in the project's master data set using a process detailed in North, Krause, Sheridan, and Ritchie (2017). In brief, the Peak Popularity measures represented the highest position the song reached on a variety of sales and airplay charts. The respective charts were weighted according to their geographical range of coverage and generality, so that for instance an appearance on the UK singles chart was assigned a weighting of 1 whereas an appearance of the song on an album that appeared in the UK album chart was assigned a weighting of 0.5. For each song per chart, the

Peak Popularity score was calculated as 1 divided by the sum of (peak chart position multiplied by chart weighting). Hit Duration scores were determined by the total number of weeks that each song appeared on the charts, irrespective of chart position. Note that although the songs included in the research were selected on the basis of their appearance in the top 5 singles chart, the popularity data included information from both singles and albums charts, since the appearance of a song on a successful album is clearly also relevant to its popularity.

Results

Examination of the Mahalanobis distance scores (minimum = 0.90, maximum = 58.24; Tabachnick & Fidell, 2013) identified multivariate outliers. This was expected because the research explicitly concerns songs with extremely high levels of popularity. High skewness and Kurtosis for all of the criterion variables were identified, and so following statistical conventions (Allen & Bennett, 2012; Tabachnick & Fidell, 2013) the four criterion variables were subjected to logarithmic (log10) transformation to minimise the skewness of the data. Collinearity statistics for tolerance and variance inflation factors (Tolerance > 0.1, VIF 10) were all within acceptable bounds (Allen & Bennett, 2012; Tabachnick & Fidell, 2013). The normality of residuals was assessed visually through the normal Q-Q plots of standardised residuals. The criterion variable plots indicated that the data points followed a normal distribution. However, as expected, some slight violations of assumptions were detected for the predictor variables. Linearity and homoscedasticity were visually assessed through the scatterplot of standardised residuals against standardised predicted values. Overall, there were no discernible patterns within the scatterplots, indicating that the assumptions of linearity and homoscedasticity were met.

A hierarchical multiple regression analysis was conducted to test whether the lyric processing fluency scores could predict UK - US Peak Popularity. Energy was entered on step 1, and was non-significant, R^2 =.000, F(1, 269) = 0.03, p = .875. On step 2, the lyric processing fluency variables were added, accounting for an additional significant 11% of the variance in UK - US Peak Popularity, ΔR^2 =.110, ΔF(6, 263) = 5.42, p < .001. In combination, all of the predictor variables explained a significant 11% of the variance in UK - US Peak Popularity, R^2 =.110, F(7, 263) = 4.65, p < .001. Further details are presented in Table 1. Specifically, Rhyme Saturation scores were related positively to UK – US Peak Popularity, whereas the Basic Lyric Readability Properties were related negatively to UK – US Peak Popularity, indicating that more easily processed lyrics were more popular.

- TABLE 1 HERE -

A second hierarchical multiple regression analysis was conducted to test whether the lyric processing fluency scores could predict UK - US Hit Duration. Energy was entered on step 1, and was non-significant, R^2 =.005, F(1, 269) = 1.34, p = .249. On step 2, the lyric processing fluency predictor variables were added, accounting for a non-significant additional 4.4% of the variance in UK - US Hit Duration, ΔR^2 =.044, ΔF(6, 263) = 2.03, p = .062. In combination, all of the predictor variables explained a non-significant 4.9% of the variance in UK - US Hit Duration, R^2 =.049, F(7, 263) = 1.94, p = .064 (see Table 2).

- TABLE 2 HERE -

A third hierarchical multiple regression analysis was conducted to test whether the lyric processing fluency scores could predict UK Peak Popularity. Energy was entered on step 1, and was non-significant, R^2 =.001, $F(1, 269) = 0.25$, $p = .614$. On step 2, the lyric processing fluency predictor variables were added, accounting for an additional significant 10.6% of the variance in UK Peak Popularity, ΔR^2 =.106, $\Delta F(6, 263) = 5.15$, $p < .001$. In combination, all of the predictor variables explained a significant 10.7% of the variance in UK Peak Popularity, R^2 =.107, $F(7, 263) = 4.49$, $p < .001$. Further details are presented in Table 3. Rhyme Saturation scores were related positively to UK Peak Popularity, whereas the Basic Lyric Readability Properties were related negatively to UK Peak Popularity, indicating that more easily processed lyrics were more popular.

- TABLE 3 HERE -

A final hierarchical multiple regression analysis was conducted to test whether the lyric processing fluency scores could predict UK Hit Duration. Energy was entered on step 1, and was non-significant, R^2 =.000, $F(1, 269) = 0.02$, $p = .886$. On step 2, the lyric processing fluency predictor variables were added, accounting for an additional non-significant 4.6% of the variance in UK Hit Duration, ΔR^2 =.046, $\Delta F(6, 263) = .2.11$, $p = .053$. In combination, all of the predictor variables explained a non-significant 4.6% of the variance in UK Hit Duration, R^2 =.046, $F(7, 263) = 1.81$, $p = .086$ (see Table 4).

- TABLE 4 HERE -

Discussion

This study investigated whether various factors relating to the processing fluency of lyrics could explain song popularity. Both measures of Peak Popularity were related to variables that would influence the processing fluency of the lyrics, so that lyrics that were easier to process were associated with songs reaching higher chart positions. This supports hypotheses 1 and 3. Both measures of Hit Duration were not related to the processing fluency of the lyrics, so that lyrics that were easier to process were not associated with songs spending longer periods of time in the charts. This does not support hypotheses 2 and 4.

The significant results concerning Peak Popularity are attributable primarily to two of the processing fluency variables, namely Rhyme Saturation and Basic Lyric Readability Properties. As noted earlier, Rhyme Saturation describes the extent to which rhymes appear in the song, and captures features such as rhymes per syllable, rhyme density, rhymes per line, perfect rhymes, and links per line. Basic Lyric Readability Properties captures features such as word count, syllable count, words with one syllable, words with six letters, unique word count, and grammar error count. Lyrics that are simpler in these terms achieve a higher peak position on the charts. It is also interesting that popularity was not predicted by the remaining processing fluency variables, which included several measures used extensively in previous research: the present results show that specifically peak popularity is related to rhyme saturation and basic readability properties, rather than that all aspects of popularity are related to all aspects of processing fluency. This is consistent with the research cited above showing that it is important to consider specific aspects of processing fluency and popularity rather than treating ‘processing fluency’ and ‘popularity’ as single variables.

The differing results concerning Peak Popularity versus Hit Duration suggest that easier readability and higher rhyme saturation in the lyrics are related to short-term spikes in popularity rather than popularity over more extended periods of time. It is tempting to speculate that this may arise because these properties of the lyrics lead to the songs being quickly perceived as boring or repetitive, so that while the lyrics are quickly understood they are also quick to lose the high level of popularity that results from this processing fluency. Future research could investigate the mediating role of processing fluency in the effect of repetition on liking for lyrics (cf. Heyduk, 1975; Johnston & Childers, 2021).

In this context it is particularly interesting to relate the present findings to those of two of the studies cited above. First, Nunes et al. (2015) showed that lyric processing fluency, as measured by lexical repetition, contributed to a song's higher chart position and faster chart success; but that excessive lyric repetition was associated with decreased popularity. Our speculative interpretation of the differing findings here concerning Peak Popularity and Hit Duration is arguably consistent with Nunes et al.'s findings. Note that such an argument is also consistent with Huron s (2013) habituation-fluency theory of repetition, which argues that internal repetition within a piece of music leads to processing fluency, and in turn a positive hedonic response, but that this repetition also leads to habituation which itself reduces a listener s responsiveness to the music. These ideas may inform future research on processing fluency in the context of lyrics, particularly given Huron s (2013, p.7) argument that, With the possible exception of dance and meditation, there appears to be nothing else in common human experience that is comparable to music in its repetitiveness". In the context of music and lyrics, processing fluency has the potential to be an extremely complex concept that interacts with several other musical features, so that for example the repetitiveness of the music interacts with processing fluency to influence the point in time at which peak popularity occurs.

Second, we observed a significant relationship between some aspects of lyric processing fluency and Peak Popularity: in contrast, North, Krause, and Ritchie (2020) identified a positive relationship between the typicality of lyrics relative to the remainder of the cohort and the duration of the chart tenure of the song in question, and that the relationship between typicality and duration of tenure was stronger than that between typicality and peak chart position. If we consider the present findings in the context of North et al.'s they suggest that peak popularity is related to the simplicity of the lyrics (at least in terms of rhyme saturation and basic readability) whereas the duration of popularity is related more closely to lyrics that are typical. There may be a subtle distinction between typicality on the one hand and rhyme saturation and basic readability on the other, but this distinction may be important in explaining whether a song enjoys a longer period of popularity versus a higher peak in popularity.

Before concluding we should acknowledge several limitations of the present research. First, for the sake of clarity we should stress that there is no suggestion that processing fluency could or should be the primary determinant of the commercial success of a song. In the light of the numerous psychological and marketing factors (e.g., advertising, radio airplay, artist management) that influence commercial success, however, it is interesting that rhyme saturation and basic readability can predict the peak chart position that a song will reach. Second, the present data do not allow us to determine the extent to which any given individual's liking for a song is influenced by processing fluency. Similarly the present research concerned only very successful pop songs, and we cannot draw any conclusions concerning whether the findings reported here generalise to the millions of less popular tracks within other genres or cultures to which someone might listen; or whether the present findings can be generalised to future chart performance. Third, our conclusions concerning processing fluency are only as valid as the variables and computer software used to operationalise this. With regard to readability, for example, the use of complex metaphors within lyrics could mean that even simple words and grammatical structures can nonetheless be found in lyrics that are difficult to comprehend. Similarly, lyrics inevitably use different forms of sentence structure and grammar compared to conventional prose, whereas computerised measures of processing fluency are typically developed for analysis of the latter. The inability of computers to cope well with metaphors and similar artistic devices clearly hampers all current research on music lyrics.

This leads to a fourth point, namely that the use of computerised measures throughout makes it impossible for the present research to distinguish between objective properties of the lyrics that might be expected to influence processing fluency and a listener's subjective experience of finding it difficult to comprehend or otherwise process those lyrics as a consequence of the properties in question. Computerised measures make the implicit assumption that objectively complex lyrics are subjectively difficult to process, and that objectively simple lyrics are subjectively easy to comprehend. It would be extremely interesting to conduct further research that compared measures of the processing fluency of lyrics obtained by computerised methods against similar measures obtained from human participants. We suspect that the two sets of data may differ when considering lyrics containing metaphors, which make reference to broader cultural events and movements, or which otherwise require comprehension beyond the literal meaning of the words and use of standard grammar. Finally, the present results do not tell us why specific aspects of processing fluency influence popularity. For example, do rhymes make the lyrics more memorable, easier to understand and interpret, or have some other specific effect altogether?

The current findings demonstrate that basic readability and rhyme saturation of lyrics can predict a song s peak chart position, and so provide evidence that one form of popularity is related to two aspects of processing fluency. This finding contributes to our understanding of processing fluency and the prediction of music popularity, and of course indicate more generally that lyrics themselves influence the popularity of music. Future research may attempt to replicate or extend these findings to different samples of music, to explore the more precise relationships between particular lyrical features and particular measures of popularity, or to investigate the relationship between processing fluency and repetition in predicting differing forms of popularity.

References

Allen, P., & Bennett, K. (2012). *SPSS: a practical guide version 20.0*. Cengage Learning Australia. https://books.google.com.au/books?id=iL9jLwEACAAJ

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, *13*(3), 219-235. https://doi.org/10.1177/1088868309341564

Askin, N., & Mauskapf, M. (2017). What makes popular culture popular? Product features and optimal differentiation in music. *American Sociological Review*, *82*(5), 910-944. http://www.jstor.org.dbgw.lis.curtin.edu.au/stable/26426411

Aydin, A. E. (2018). Processing fluency: Examining its relationship with thinking style and purchase intention. *Journal of Marketing Communications*, *24*(6), 588-598. https://doi.org/10.1080/13527266.2016.1167765

Bannister, M. (2018, March 9). The Billboard Hot 100: Exploring six decades of number one singles. GitHub. https://github.com/mspbannister/dand-p4-billboard/blob/master/Billboard_analysis__100417_.md>.

Basyuk, T. (2018). Influence of readability on popularization of internet resources. *Computational Linguistics and Intelligent Systems*, *2*, 146-156.

Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century Crofts.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Taylor & Francis Group.

Cook, S. L., & Krupar, K. (2010). Defining the twentieth century and impacting the twenty-first: Semantic habits created through radio and song. *ETC: A Review of General Semantics*, *67*(4), 412-434. http://www.jstor.org/stable/42579071

Dhanaraj, R., & Logan, B. (2005, September 11-15). Automatic prediction of hit songs [paper presentation]. *Proceedings of the 6th International Conference on Music Information Retrieval*, London, UK. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.86.4054&rep=rep1&type=pdf

Dragojevic, M., & Giles, H. (2016). I don't like you because you're hard to understand: The role of processing fluency in the language attitudes process. *Human Communication Research*, *42*(3), 396-420. https://doi.org/10.1111/hcre.12079

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Filkuková, P., & Klempe, S. H. (2013). Rhyme as reason in commercial and social advertising. *Scandinavian Journal of Psychology*, *54*(5), 423-431. https://doi.org/https://doi.org/10.1111/sjop.12069

Flesch, R. (1951). *The art of clear thinking*. New York: Harper.

Gao, X., Dera, J., Nijhof, A. D., & Willems, R. M. (2019). Is less readable liked better? The case of font readability in poetry appreciation. *PLOS ONE*, *14*(12), e0225757. https://doi.org/10.1371/journal.pone.0225757

Hahn, L. E., Benders, T., Snijders, T. M., & Fikkert, P. (2018). Infants sensitivity to rhyme in songs. *Infant Behavior and Development*, *52*, 130-139. https://doi.org/10.1016/j.infbeh.2018.07.002

Hansen, J., Dechêne, A., & Wänke, M. (2008). Discrepant fluency increases subjective truth. *Journal of Experimental Social Psychology*, *44*(3), 687-691. https://doi.org/10.1016/j.jesp.2007.04.005

Hart, R. P., Carroll, C. E., & Spiars, S. (2013). *Diction 7.0: the text analysis program*. Austin: Digitext.

Herremans, D., Martens, D., & Sörensen, K. (2014). Dance hit song prediction. *Journal of New Music Research*, *43*(3), 291-302. https://doi.org/10.1080/09298215.2014.881888

Herrmann, A., Zidansek, M., Sprott, D. E., & Spangenberg, E. R. (2013). The power of simplicity: processing fluency and the effects of olfactory cues on retail sales. *Journal of Retailing*, *89*(1), 30-43. https://doi.org/10.1016/j.jretai.2012.08.002

Heyduk, R. G. (1975). Rated preference for musical compositions as it relates to complexity and exposure frequency. *Perception and Psychophysics*, *17*(1), 84-90. https://doi.org/10.3758/bf03204003

Hirjee, H. (2010). *Rhyme, rhythm, and rhubarb: Using probabilistic methods to analyze hip hop, poetry, and misheard lyrics*. [Master's thesis, University of Waterloo]. UWSpace. http://hdl.handle.net/10012/5419

Hirjee, H., & Brown, D. (2010, August, 9-13). Rhyme Analyzer: an analysis tool for rap lyrics [paper presentation]. *Proceedings of the 11th International Conference on Music Information Retrieval, Utrecht, Netherlands*. https://ismir2010.ismir.net/proceedings/late-breaking-demo-23.pdf

Huron, D. (2013). A psychological approach to musical form: the habituation–fluency theory of repetition. *Current Musicology*, *96*, 7-35.

Im, H., Lennon, S. J., & Stoel, L. (2010). The perceptual fluency effect on pleasurable online shopping experience. *Journal of Research in Interactive Marketing*, *4*(4), 280-295. https://doi.org/10.1108/17505931011092808

Interiano, M., Kazemi, K., Wang, L., Yang, J., Yu, Z., & Komarova, N. L. (2018). Musical trends and predictability of success in contemporary songs in and out of the top charts. *Royal Society Open Science*, *5*(5), 171274. https://doi.org/10.1098/rsos.171274

Jaud, D. A., & Melnyk, V. (2020). The effect of text-only versus text-and-image wine labels on liking, taste and purchase intentions: the mediating role of affective fluency. *Journal of Retailing and Consumer Services*, *53*, 101964. https://doi.org/https://doi.org/10.1016/j.jretconser.2019.101964

Johnston, R. R., & Childers, G. M. (2021). Musical pantophagy: is there an effect on changes in preference resulting from repeated exposure? *Psychology of Music*. https://doi.org/10.1177/0305735620987299

Leder, H. (2013). Next steps in neuroaesthetics: which processes and processing stages to study? *Psychology of Aesthetics, Creativity, and the Arts*, *7*(1), 27–37. https://doi.org/10.1037/a0031585

Liebers, N., Breves, P., Schallhorn, C., & Schramm, H. (2019). Fluency in commercial breaks: the impact of repetition and conceptual priming on brand memory, evaluation, and behavioral intentions. *Journal of Promotion Management*, *25*(6), 783-798. https://doi.org/10.1080/10496491.2018.1536620

Margulis, E. H. (2013). Aesthetic responses to repetition in unfamiliar music. *Empirical Studies of the Arts*, *31*(1), 45-57. https://doi.org/10.2190/EM.31.1.c

Martindale, C., Moore, K., & Borkum, J. (1990). Aesthetic preference: anomalous findings for Berlyne's psychobiological theory. *The American Journal of Psychology*, *103*(1), 53-80. https://doi.org/10.2307/1423259

Martindale, C., Moore, K., & West, A. (1988). Relationship of preference judgments to typicality, novelty, and mere exposure. *Empirical Studies of the Arts*, *6*(1), 79-96. https://doi.org/10.2190/MCAJ-0GQT-DJTL-LNQD

Mayer, S., & Landwehr, J. R. (2018). Quantifying visual aesthetics based on processing fluency theory: four algorithmic measures for antecedents of aesthetic preferences. *Psychology of Aesthetics, Creativity, and the Arts*, *12*(4), 399-431. https://doi.org/10.1037/aca0000187

Menninghaus, W., Bohrn, I. C., Knoop, C. A., Kotz, S. A., Schlotz, W., and Jacobs, A. M. (2015). Rhetorical features facilitate prosodic processing while handicapping ease of semantic comprehension. *Cognition*, *143*, 48-60.

Miall, D. S. & Kuiken, D. (1994). Foregrounding, defamiliarization, and affect: response to literary stories. *Poetics*, *22*, 389-407.

Newman, E. J., & Schwarz, N. (2018). Good sound, good research: how audio quality influences perceptions of the research and researcher. *Science Communication*, *40*(2), 246-257. https://doi.org/10.1177/1075547018759345

North, A. C., Krause, A. E., & Ritchie, D. (2020). The relationship between pop music and lyrics: a computerized content analysis of the United Kingdom s weekly top five singles, 1999–2013. *Psychology of Music*, *49*(4), 735-758. https://doi.org/10.1177/0305735619896409

North, A. C., Krause, A. E., Sheridan, L. P., & Ritchie, D. (2017). Energy, typicality, and music sales: a computerized analysis of 143,353 pieces. *Empirical Studies of the Arts*, *35*(2), 214-229. https://doi.org/10.1177/0276237416688063

North, A. C., Krause, A. E., Sheridan, L. P., & Ritchie, D. (2018). Energy, popularity, and the circumplex: a computerized analysis of emotion in 143,353 musical pieces. *Empirical Studies of the Arts*, *36*(2), 127-161. https://doi.org/10.1177/0276237417704339

North, A. C., Krause, A. E., Sheridan, L. P., & Ritchie, D. (2019). Popularity, mood, energy, and typicality in music: a computerized analysis of 204,506 pieces. *Psychology of Aesthetics, Creativity, and the Arts*, *13*(1), 89-109. https://doi.org/10.1037/aca0000165

North, A. C., Krause, A. E., Sheridan, L. P., & Ritchie, D. (2020). Comparison of popular music in the United States and the United Kingdom: computerized analysis of 42,714 pieces. *Psychology of Music*, *48*(6), 846-860. https://doi.org/10.1177/0305735619830185

Nunes, J. C., Ordanini, A., & Valsesia, F. (2015). The power of repetition: repetitive lyrics in a song increase processing fluency and drive market success. *Journal of Consumer Psychology*, *25*(2), 187-199. https://doi.org/10.1016/j.jcps.2014.12.004

Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences*, *12*(6), 237-241. https://doi.org/10.1016/j.tics.2008.02.014

Osgood, C. E. (1952). The nature and measurement of meaning. *Psychological Bulletin*, *49*, 197-237.

Pancer, E., Chandler, V., Poole, M., & Noseworthy, T. J. (2019). How readability shapes social media engagement. *Journal of Consumer Psychology*, *29*(2), 262-270. https://doi.org/10.1002/jcpy.1073

Putter, K., Krause, A. E., and North, A. C. (2021). Popular music lyrics and the COVID-19 pandemic. *Psychology of Music*. https://doi.org/10.1177/03057356211045114

Raza, A., & Nanath, K. (2020, July 16-17). Predicting a hit song with machine learning: is there an apriori secret formula? [paper presentation]. *Proceedings of the 2020 International Conference on Data Science Artificial Intelligence and Business Analytics, Medan, Indonesia*. https://doi.org/10.1109/DATABIA50434.2020.9190613

Readable.com by Added Bytes Ltd (n.d.). Readability formula. Retrieved November 19, 2021, from https://readable.com/features/readability-formulas/

Ren, J., & Kauffman, R. J. (2017, June 5-10). Understanding music track popularity in a social network [paper presentation]. *Proceedings of the 25th European Conference on Information Systems ECIS, Guimarães, Portugal*. https://ink.library.smu.edu.sg/sis_research/3960

Russell, J. A. (1978). Evidence of convergent validity on the dimensions of affect. *Journal of Personality and Social Psychology*, *36*, 1152-1168.

Schedl, M. (2019, May 13-17). Genre differences of song lyrics and artist wikis: an analysis of popularity, length, repetitiveness, and readability [paper presentation]. *Proceedings of the World Wide Web Conference, San Francisco, CA, USA*. https://doi.org/10.1145/3308558.3313604

Schotanus, Y. P. (2021). Experiencing Dylan: the effect of formal structure and performance on the popularity and interpretation of two Dylan songs. in T. Gurke & S. Winnett (eds.), *Words, music, and the popular* (pp.39-64). Cham: Palgrave Macmillan.

Schotanus, Y., Koops, H. V., Edworthy, J. R. (2018). Interaction between musical and poetic form affects song popularity: the case of the Genevan Psalter. *Psychomusicology*, *28*, 127-151.

Schwarz, N. (2004). Metacognitive experiences in consumer judgment and decision making. *Journal of Consumer Psychology*, *14*(4), 332-348. https://doi.org/https://doi.org/10.1207/s15327663jcp1404_2

Singhi, A., & Brown, D. G. (2015, Jun 16-19). Can song lyrics predict hits [paper presentation]. *Proceedings of the 11th International Symposium on Computer Music Multidisciplinary Research, Plymouth, United Kingdom*. https://cs.uwaterloo.ca/~browndg/CMMR15data/CMMR2015paper.pdf

Song, H., & Schwarz, N. (2009). If it's difficult to pronounce, it must be risky: fluency, familiarity, and risk perception. *Psychological Science*, *20*(2), 135-138. https://doi.org/10.1111/j.1467-9280.2009.02267.x

Tabachnick, B., & Fidell, L. (2013). *Using Multivariate Statistics: Pearson New International Edition PDF EBook*. Pearson Education, Limited. http://ebookcentral.proquest.com/lib/curtin/detail.action?docID=5175291

van Oosten, P., Tanghe, D., & Hoste, V. (2010, May 19-21). Towards an improved methodology for automated readability prediction [paper presentation]. *Paper presented at the 7th Conference on International Language Resources and Evaluation, Paris, France*. http://hdl.handle.net/1854/LU-1055826

Varnum, M. E. W., Krems, J. A., Morris, C., Wormley, A., & Grossmann, I. (2021). Why are song lyrics becoming simpler? A time series analysis of lyrical complexity in six decades of American popular music. *PLOS ONE*, *16*(1), Article e0244576. https://doi.org/10.1371/journal.pone.0244576

Vitz, P. C. (1966). Affect as a function of stimulus variation. *Journal of Experimental Psychology*, *71*, 74-79.

Wang, L. W., Miller, M. J., Schmitt, M. R., & Wen, F. K. (2013). Assessing readability formula differences with written health information materials: application, results, and recommendations. *Research in Social and Administrative Pharmacy*, *9*(5), 503-516. https://doi.org/10.1016/j.sapharm.2012.05.009

Whitfield, T. W. A. (1983). Predicting preference for familiar, everyday objects: an experimental confrontation between two theories of aesthetic behaviour. *Journal of Environmental Psychology*, *3*(3), 221-237. /https://doi.org/10.1016/S0272-4944(83)80002-4

Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality And Social Psychology*, *81*(6), 989-1000. https://doi.org/10.1037/0022-3514.81.6.989

Wright, P. (2003). Criteria and ingredients for successful patient information. *Journal of Audiovisual Media in Medicine*, *26*(1), 6-10. https://doi.org/10.1080/0140511031000091211

Table 1

Hierarchical Regression Results for UK - US Peak Popularity

Variable	*B*	95% CI for *B*		SE *B*	β	*t*
		LL	*UL*			
Step 1						
Constant	-0.69	-0.93	-0.45	0.12		-5.65***
Energy	0.00	-0.00	0.00	0.00	.01	0.16
Step 2						
Constant	-1.92	-3.74	-0.11	0.92		-2.08*
Energy	0.00	-0.00	0.00	0.00	-.03	-0.44
Complexity	0.37	-0.14	0.87	0.26	.13	1.42
Rhyme Saturation	0.13	0.03	0.23	0.05	.16	2.48*
Multisyllabic Rhymes	-0.06	-0.18	0.07	0.06	-.07	-0.93
Short Unique Rhymes	-0.07	-0.18	0.04	0.06	-.09	-1.33
Basic Lyric Readability Properties	-0.20	-0.30	-0.10	0.05	-.25	-4.06***
Computerised Reading Difficulty	-0.03	-0.19	0.13	0.08	-.04	-0.37

Note. Overall model: R^2 =.110, $F(7, 263) = 4.65$, $p < .001$; CI = confidence interval; *LL* = lower limit; *UL.* = upper limit.
$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 2

Hierarchical Regression Results for UK - US Hit Duration

Variable	*B*	95% CI for *B*		SE *B*	β	*t*
		LL	*UL*			
Step 1						
Constant	1.65	1.49	1.81	0.08		19.88**
Energy	0.00	-0.00	0.00	0.00	.07	1.16
Step 2						
Constant	1.10	-0.19	2.38	0.65		1.68
Energy	0.00	-0.00	0.00	0.00	.03	0.49
Complexity	0.17	-0.19	0.53	0.18	.09	0.93
Rhyme Saturation	0.09	0.01	0.16	0.04	.16	2.34*
Multisyllabic Rhymes	-0.03	-0.12	0.06	0.04	-.06	-0.74
Short Unique Rhymes	-0.05	-0.12	0.03	0.04	-.08	-1.14
Basic Lyric Readability Properties	-0.05	-0.12	0.02	0.04	-.10	-1.47
Computerised Reading Difficulty	0.02	-0.09	0.13	0.06	.04	-.35

Note. Overall model: R^2 =.049, $F(7, 263) = 1.94$, $p = .064$; CI = confidence interval; *LL* = lower limit; *UL.* = upper limit.
$^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Table 3

Hierarchical Regression Results for UK Peak Popularity

Variable	*B*	95% CI for *B*		SE *B*	β	*t*
		LL	*UL*			
Step 1						
Constant	-0.73	-0.95	-0.51	0.11		-6.50***
Energy	0.00	-0.00	0.00	0.00	-.03	-0.50
Step 2						
Constant	-1.51	-3.19	0.17	0.85		-1.78
Energy	-0.00	-0.00	0.00	0.00	-.07	-1.90
Complexity	0.24	-0.23	0.71	0.24	.09	0.99
Rhyme Saturation	0.10	0.01	0.19	0.05	.14	2.09*
Multisyllabic Rhymes	-0.07	-0.19	0.04	0.06	-.10	-1.23
Short Unique Rhymes	-0.10	-0.20	0.00	0.05	-.14	-1.96
Basic Lyric Readability Properties	-0.18	-0.27	-0.09	0.05	-.25	-3.93***
Computerised Reading Difficulty	0.01	-0.14	0.16	0.08	.02	0.18

Note. Overall model: R^2 =.107, $F(7, 263) = 4.49$, $p < .001$; CI = confidence interval; *LL* = lower limit; *UL.* = upper limit.
*$p < .05$, **$p < .01$, ***$p < .001$.

Table 4

Hierarchical Regression Results for UK Hit Duration

Variable	*B*	95% CI for *B*		SE *B*	β	*t*
		LL	*UL*			
Step 1						
Constant	1.61	1.45	1.76	0.08		20.23***
Energy	-9.64	-0.00	0.00	0.00	-0.01	-0.14
Step 2						
Constant	1.67	0.44	2.90	0.62		2.68**
Energy	-0.00	-0.00	0.00	0.00	-0.06	-0.93
Complexity	-0.00	-0.35	0.34	0.18	-0.00	-0.01
Rhyme Saturation	0.09	0.03	0.16	0.04	0.18	2.69**
Multisyllabic Rhymes	-0.03	-0.12	0.05	0.04	-0.06	-0.74
Short Unique Rhymes	-0.06	-0.13	0.02	0.04	-0.11	-1.56
Basic Lyric Readability Properties	-0.02	-0.08	0.05	0.03	-0.03	-0.44
Computerised Reading Difficulty	0.03	-0.08	0.14	0.06	0.06	0.52

Note. Overall model: R^2 =.046, $F(7, 263) = 1.81$, p = .086; CI = confidence interval; *LL* = lower limit; *UL.* = upper limit.
*$p < .05$, **$p < .01$, ***$p < .001$.

Appendix 1 - Overview of Rhyme Analyzer variables

Variable	Description
Syllables per Line	Mean number of syllables per line
Syllables per Word	Mean word length in syllables
Syllable Variation	Standard deviation of line lengths in syllables
Novel Word Proportion	Mean percentage of words in the second line in a pair not appearing in the first
Rhymes per Line	Mean number of detected rhymes per line
Rhymes per Syllable	Mean number of detected rhymes per syllable
Rhyme Density	Total number of rhymed syllables divided by total number syllables
End Pairs per Line	Percentage of lines ending with a line-final rhyme
End Pairs Grown	Percentage of rhyming couplets in which the second line is more than 15% longer (in syllables) than the first
End Pairs Shrunk	Percentage of rhyming couplets in which the second line is more than 15% shorter (in syllables) than the first
End Pairs Even	Percentage of rhyming couplets neither grown nor shrunk
Average End Score	Mean similarity score of line final rhymes
Average End Syllable Score	Mean similarity score per syllable in line final rhymes
Singles per Rhyme	Percentage of rhymes that are one syllable long
Doubles per Rhyme	Percentage of rhymes that are two syllables long
Triples per Rhyme	Percentage of rhymes that are three syllables long
Quads per Rhyme	Percentage of rhymes that are four syllables long
Longs per Rhyme	Percentage of rhymes that are longer than four syllables
Perfect Rhymes	Percentage of rhymes with identical vowels and codas
Line Internals per Line	Number of rhymes with both parts falling in the same line divided by total number of lines
Links per Line	Mean number of link rhymes per line
Bridges per Line	Mean number of bridge rhymes per line
Compounds per Line	Mean number of compound rhymes per line
Chaining per Line	Total number of words or phrases involved in chain rhymes divided by total number of lines

Adapted from Hirjee (2010)

Appendix 2 - Principal components analysis of Rhyme Analyzer variables

Rhyme Analyzer variables	Factor loading		
	1	2	3
Rhymes per Syllable	.99		
Line Internals per Line	.95		
Chaining per Line	.94		
Rhyme Density	.86		
Rhymes per Line	.85	.32	
Perfect Rhymes	.66	-.48	
Links per Line	.50		
Syllables per Line		.95	
Syllable Variation		.74	
Compounds per Line		.65	
Bridges per Line		.57	-.40
Novel Word Proportion			.84
Longs per Rhyme			-.71
Syllables per Word			.56
Eigenvalue	6.10	2.46	2.12
% of Variance	25.43	10.25	8.81

Note. Factor loadings < .3 are suppressed.

Appendix 3 - Overview of Readable variables

Variable	Description
	Readability formulae
Flesch-Kincaid Grade Level	Originally created to analyse the readability of technical material. Suitable for the grading of a wide range of writing and for all ages. The Flesch-Kincaid Grade Level is equivalent to the US grade level of education. It shows the required education to be able to understand a text.
Flesch Reading Ease Score	This was originally created in education research but is now used for a variety of purposes (e.g., government bodies and policy writers). Scores are typically between 0 and 100. Higher scores indicate better readability.
Gunning Fog Index	Suitable for a wide range of material and especially for business writing and publishing. Grading suitable for all ages. The Gunning Fog formula generates a grade level between 0 and 20. It estimates the education level required to understand the text.
Coleman-Liau Index	A character-calculated formula suitable for a wide range of texts and for all ages, and especially useful within education. Coleman-Liau is a readability formula which shows the reading level of a text. It uses sentences and letters as variables.
Smog Index	SMOG stands for 'Simple Measure of Gobbledygook'. It is a readability framework that measures how many years of education the average person needs to have to understand a text. It is best for texts of 30 sentences or more. Applicable to a range of texts and ages but has been proven particularly useful in healthcare literature.
Automated Readability Index	The ARI assesses the U.S. grade level required to read a piece of text and is suitable for a wide range of texts and ages but has proven especially useful in technical writing. It differs from other formulae by counting characters rather than syllables. The more characters, the harder the word.
Forcast Grade Level	The FORCAST formula is designed to analyse technical documents. These could be training manuals, forms, and surveys. The formula is: Grade level = 20 – (N / 10), where N = number of single-syllable words in a 150-word sample.

Powers Sumner Kearl Grade	The Powers Sumner Kearl formula calculates a US grade level. It uses the number of words, sentence length and number of syllables as variables. The formula is as follows: Reading age = 0.0778 (average sentence length) + 0.0455 (number of syllables) + 2.7971.
Rix Readability	Rix is calculated using word length and sentence length. It is an evolution from the Lix formula and produces a grade level.
Raygor Readability	The Raygor estimate graph is a graph-based readability formula. The US grade level is calculated using sentences and letters as variables. This formula is best used for grading middle school level texts.
Fry Readability	Fry originally created the formula to analyse high school educational texts, but it can be used for all US grade levels of text, including college level. Fry is a graph-based formula using sentences and syllables as variables. Fry plots the text on a graph corresponding to the score.
Readable Rating	Readable's bespoke rating system grades text from A to E for readability. Text aimed at the general public should be grade B or better.
CEFR Level	CEFR stands for the Common European Framework of Reference for Languages, and is used to evaluate language skills. It is used by government bodies to help their content reach people with varying language comprehension.
IELTS Level	IELTS stands for The International English Language Testing System, and is a standardized test for English language proficiency.
Spache Score	Spache is a readability test for English and is best for texts up to fourth-grade level. It compares the words in a text to a set list of everyday words, which someone up to fourth grade generally can understand.
New Dale-Chall Score	The New Dale-Chall readability score measures a text against a number of words considered familiar to fourth graders. The more unfamiliar words used, the higher the reading level.
Lix Readability	Lix uses word length and sentence length as variables. It was formulated to consider the readability of non-English reading materials. Scores are typically between 0 and 100. Lower scores indicate better readability, with a good score being 40 or below.
Lensear Write	The Lensear Write readability formula is a text-based formula. It scores monosyllabic words and strong verbs. Scores are typically between 0 and 100. Good readability is a score of 70-80, with higher scores indicating overly simple text and lower scores indicating poor readability.
Readability issues	

Sentences > 30 Syllables	Percentage of sentences containing more than 30 syllables.
Sentences > 20 Syllables	Percentage of sentences containing more than 20 syllables.
Words > 4 Syllables	Percentage of words containing more than 4 syllables.
Words > 12 Letters	Percentage of words containing more than 12 letters.
Text issues	
Characters Per Word	Mean characters per word
Syllables Per Word	Mean syllables per word
Words Per Sentence	Mean words per sentence
Words Per Paragraph	Mean words per paragraph
Sentences Per Paragraph	Mean sentences per paragraph
Content	
Reach	Reach is a measure of the proportion of the target audience that can read the content easily.
Tone	Tone analysis indicates how formal or how conversational the text looks. Conversational text uses more pronouns and fewer prepositions, among other differences.
Sentiment	Sentiment analysis gives an idea of whether the text uses mostly positive language, negative language, or neutral language. For longer pieces, the text is split into three to give sentiment analysis for the beginning, middle and end of the piece.
Personalism	Personalism is the measure of the degree to which the text refers to the reader, rather than themselves.
Text statistics	
Character Count	Letters (A-Z only, not including numbers and punctuation).
Syllable Count	Total syllables
Word Count	Total words
Unique Word Count	Total unique words
Sentence Count	Sentences are assumed to end with periods, question marks, exclamation marks or line breaks.

Paragraph Count	Paragraphs are defined as ending with a line break.
Timing	
Reading Count	The time it would take for the average person to read the text, at a rate of 225 words per minute
Speaking Time	The time it would take for the average person to say the text aloud, at a rate of 125 words per minute.
Text composition	
Adjectives	Percentage of adjectives
Adverbs	Percentage of adverbs
Conjunctions	Percentage of conjunctions
Determiners	Percentage of determiners
Interjections	Percentage of interjections
Nouns	Percentage of nouns
Proper Nouns	Percentage of proper nouns
Prepositions	Percentage of prepositions
Pronouns	Percentage of pronouns
Qualifiers	Percentage of qualifiers
Verbs	Percentage of verbs
Unrecognised Words	Unrecognised words are words which could not be classified by the program.
Non-Words	Non-words are items which are not words, including email addresses and URLs.
Writing issue	
Passive Voice Count	Passive voice can make the text less engaging to the reader compared to the active voice.
Adverb Count	Adverbs are worth avoiding where possible, as they can usually be replaced by more active verbs.
Cliché Count	Clichés are phrases that have been overused and can appear unoriginal.
Content analysis (percentage of words in the text that are found in the associated program dictionary)	
Keyword Density Analysis	Google search friendly words

Gender Analysis	Identifies whether the text ‘looks’ like it was written by a man or woman.
Profanity Detector	Swear words and abusive words
Buzzword Detector	Strategic buzzwords
Stop Word Detector	Words that search engines consider unimportant
Hedge Word Detector	Words used in writing to lessen the impact of something, or to introduce ambiguity and uncertainty
Lazy Word Detector	Words deemed as uncompelling or not bold
Names Detector	Finds referenced names in text
Transition Word Detector	Linking words

Note. Table information retrieved from Readability formulas (n.d.). Readable.com by Added Bytes Ltd. Retrieved November 19, 2021, from https://readable.com/.

Appendix 4 - Principal components analysis of Readable variables

"Readable" output variables	Factor loading	
	1	2

Word Count	.99	
Syllable Count	.98	
Sentence Count Flesch	.98	
Words With One Syllable	.98	
Paragraph Count	.93	
Sentence Count Flesch	.92	
Sentence Count	.92	
Composition Pronoun Count	.90	
Composition Verb Count	.85	
Words With Six Letters	.73	.48
Composition Noun Count	.72	
Composition Conjunction Count	.71	
Composition Preposition Count	.71	
Unique Word Count	.66	
Composition Determiner Count	.63	
Spache Difficult Words	.56	
Composition Adjective Count	.52	
Composition Unknown Count	.51	
Grammar Error Count	.36	
Flesch Reading Ease		-.93
SMOG Index		.92
Lix Score		.90
Syllables Per Word		.89
Powers Sumner Kearl Grade		.86
Flesch Kincaid Grade Level		.85
Gunning Fog Score		.83
FORCAST Grade Level		.82
Rix Score		.80
Letters per Word		.75

Words With Seven Letters	.64	.67
Automated Readability Index		.64
Coleman Liau Index		.62
Fry Grade Level		.60
Eigenvalue	14.37	10.20
% of Variance	26.62	18.88

Note. Factor loadings < .3 are suppressed.