Musicae Scientiae

http://msx.sagepub.com/

I like the way it sounds: The influence of instrumentation on a pop song's place in the charts

Joseph C. Nunes and Andrea Ordanini

Musicae Scientiae published online 3 September 2014

DOI: 10.1177/1029864914548528

Article

MUSICÆ SCIENTIÆ

Musicae Scientiae
1–18

DOI: 10.1177/1029864914548528
msx.sagepub.com
SAGE

I like the way it sounds: The influence of instrumentation on a pop song's place in the charts

Joseph C. Nunes
University of Southern California, USA

Andrea Ordanini
Bocconi University, Italy

Abstract

The way a song sounds depends on, among other things, what instruments are audible to the listener. We document the systematic relationship between instrumentation, or the particular combination of instruments audible in a hit song, and its relative market appeal vis-à-vis its place on the charts. This is accomplished using mixed analytical methods under ecologically valid conditions. The data come from *Billboard*'s Hot 100 weekly popularity rankings from the past 55 years. We compare number 1 singles with songs that never climbed above number 90. First, using Qualitative Comparative Analysis (QCA), we identify two configurations of core instruments that result in a song being more popular such that these configurations are present far more consistently among the set of number 1 hit songs. We also identify three configurations of core instruments that result in a song being less popular such that these configurations are present far more consistently among songs that stagnate at the bottom of the charts. What stands out is popular (less popular) configurations always include (exclude) background vocals. Further, our qualitative results reveal number 1 songs tend to include a greater number of instruments. Second, we utilize logit regression to document how the absolute number of distinct instrument types perceptible in a song affects the chances it will be a number 1 hit as opposed to stay at or below number 90. Our results suggest songs that do not follow conventional instrumentation and instead include an atypically low or high number of instruments are more likely to stand out, becoming number 1 hits. Looking across time, a greater number of instruments increased the chances of being a number1 song during mid-1970s, the 1980s and the 1990s, while fewer instruments increased the chances during the 1960s and in the late 2000s.

Keywords

Billboard's Hot 100 charts, instrumentation, musical tastes, orchestration, popular music, popularity, timbre, music preferences, sound

Corresponding author:
Joseph C. Nunes, University of Southern California, Marshall School of Business, 14718 Sutton Street, Los Angeles, CA 90089-0443, USA.
Email: jnunes@marshall.usc.edu

Why are some songs more popular than others? At present, we know relatively little about how different acoustic elements of a song influence whether it will rise to number 1 or end up lingering at the bottom of the pop music charts. When asked why he or she likes a particular song, the person with an untrained ear may respond simply by saying "I like *the way it sounds*," referring to the song's global sound quality or what might be considered an ineffable combination of instruments and vocals (Radocy & Boyle, 1997). In fact, a major determinant of the way a song sounds is the instrumentation, or the selection of particular instruments in the composition that are audible to the listener. This research investigates how a song's popularity, reflected by its relative position in the charts, is associated with the specific combination of instruments employed.

When writing a song, the musicians, or producer, must consider each instrument's distinctive and recognizable sound commonly referred to as its *timbre*. Timbre has been called a catch-all "for everything that cannot be labeled pitch or loudness" (McAdams & Bregman, 1979, p. 34) and is the psychoacoustic property that allows listeners to distinguish between different types of sound production, whether voices, string instruments, wind instruments, or percussion instruments.[1] The timbre associated with a specific instrument enables listeners to determine whether the same note is being played on a guitar or on a piano (Handel, 1989; McAdams, 1993). Even when listeners don't perform perfectly with respect to identifying specific instruments, such as a flute versus a recorder, they are adept at identifying instrument "types" (Giordano & McAdams, 2010).

Combining various instruments (e.g., a guitar with a piano) creates a distinctive mix of sounds, which have been shown to be identifiable as specific "timbre mixtures" (Alluri & Toiviainen, 2009; Aucouturier, 2006; Aucouturier, Pachet, & Sandler, 2005). A song's timbre mixture is often indicative of its genre, with the electric guitar, feedback and keyboards sound of Psychedelic Rock and the enormous energy and extreme low frequencies of Hip-Hop (bass, drums) both easily and quickly recognizable to listeners. In short, timbre mixtures refer to a "global sound quality of the recording" that enables listeners to identify, classify, and categorize pieces of music (Schellenberg, Iverson, & McKinnon, 1999, p. 642).

In fact, participants in studies by Gjerdingen and Perrott (2008) were able to identify popular songs from excerpts of only 100 and 200 milliseconds, leading these researchers to conclude the timbral qualities of everyday music appear to trigger memory in ways rhythm, pitch, and tonality cues cannot. It therefore appears the same perceptual mechanisms are at work whether the listener is processing polyphonic or monophonic instrument sounds (Alluri & Toiviainen, 2010).[2] Some psychoacoustic scholars have gone as far as arguing timbre is "the most important and ecologically relevant feature of auditory events" (Menon et al., 2002, p. 1742). Our claims are more modest; we propose that timbre mixtures matter to listeners of pop music and find evidence supporting a systematic relationship between specific instrument combinations and the chances a song reaches number 1 in *Billboard*'s Hot 100 charts as opposed to never climbing above the rank of number 90.

Although musicologists, music psychologists, and musicians assume timbre mixtures influence how much someone likes a song, limited empirical research on the role of timbre's influence on music preferences has been conducted to date (Teo, 2003). In addition, many of the findings conflict. For example, in laboratory studies using classical western music excerpts, researchers have found instrumental timbre is preferred over vocal timbre (Darrow, Haack, & Kuribayashi, 1987; LeBlanc, 1981). However, the opposite appears to be true when excerpts are taken from the popular musical style at the time (Shehan, 1981). Other research has investigated preferences using instrumental excerpts performed on either a "band" instrument (trumpet, clarinet, and bassoon) or a "non-band" instrument (violin, guitar, and piano). These investigators found fourth-graders in the USA exhibit a slight preference toward band instruments, while in Greek schools the preference leaned toward non-band instruments (Cutietta & Foustalieraki, 1990). Clearly, musical style, culture, and other factors will drive preferences.

Our research set out to examine the relationship between the type of instruments and the number of instruments audible and a song's relative popularity among the mass market for pop music in the USA. We do so under ecologically valid conditions by applying both qualitative and quantitative analysis techniques to a unique dataset that includes the USA's most popular pop songs from the past 55 years. We assess relative popularity using rankings on *Billboard*'s Hot 100 singles chart and by comparing those songs that reached number 1 (hereafter Top songs) with a comparison set of those songs that, while making it onto the Hot 100, never climbed above number 90 (hereafter Bottom songs). By restricting our sample as such, we attempt to control for exogenous factors that may have facilitated or inhibited a song from attaining national exposure; once on the Hot 100, a song has reached a significant and sizable audience and its ranking in large part reflects its relative market appeal. By relying only on songs popular enough to enter the Hot 100 chart, we believe this approach provides a conservative test of the role of instrumentation on popularity.

We are not the first to investigate acoustic factors presumed to contribute to a song's becoming a chart-topping hit. Academics in the University of Bristol's Intelligent Systems Laboratory have looked at the UK top 40 singles chart from the past 50 years using learning algorithms as part of what is commonly referred to as "Hit Song Science" (Dhanaraj & Logan, 2005). They examined song features including tempo (five categories), time signature (three categories), song duration (three categories), loudness, and the simplicity of the chord sequences. They conclude that what distinguishes hits (top 5 songs) from non-hits (songs from 30–40) is that hits are simpler and are getting louder and longer over time. Their results are not without critics. Researchers at Sony (Pachet & Roy, 2008), who, using a 32,000-title proprietary database of songs that each included between 20 and 632 features such as genre, style, language, etc., report finding acoustic classifiers are not a good predictor of a song's relative popularity (categorized simply as low, medium, and high).

Our goal is far more circumscribed. We seek to identify any combination of instruments among those audible to the listener that characterize the most popular songs, the number 1 hits (Top songs), and those combinations of instruments that characterize less popular, albeit still hit songs, represented by those that never rose above number 90 (Bottom songs). We are able to identify two specific configurations "sufficient" to make it to the top of the chart (i.e., 88% of songs with these instrument combinations make it to number 1 and account for nearly 9% of all Top songs). We also identify three specific configurations sufficient for a song to stay at the bottom of the chart (82% of songs with these instruments stay at or below number 90 and account for 23% of all Bottom songs). It is critical to convey that each configuration defines a "core" set of instruments and that other instruments can be added without changing the result. In other words, if the core instrument types were a clean guitar and synthesizer, adding another instrument sound or sounds such as a distorted guitar and/or bass would not change the outcome.

The intention of our first analysis is to focus on foundational or core instruments that distinguish Top songs from Bottom songs and vice versa. However, we will also examine how the absolute number of distinct instrument types perceptible in a song impacts its popularity. Our subsequent analysis tests whether adding additional instruments on top of the core instruments can help in separating Top songs from Bottom songs. We find a positive curvilinear effect such that the chance of being a number 1 hit increases more than proportionally as the number of instruments audible in a song increases.[3] Further analysis reveals that the results from both analyses are qualified by time, that is the year in which the song appeared in the chart. Taken together, our findings suggest a strong underlying influence of instrumentation such that both specific timbre mixtures and the number of audible instruments appear to contribute to determining the relative popularity of a hit song.

Sample selection

Billboard magazine is the preeminent source for assessing a song's popularity, and the publication's Hot 100 singles chart has been the leading indicator of various songs' relative popularity since its inception in 1958 (Bradlow & Fader, 2001). Any current song in any genre that the radio is playing or that people are buying is eligible for the ranking, considered in the music industry the "best benchmark we have to measure the bigness of hits" (Molanphy, 2013).

We identified each of the 1,029 songs that reached number 1 on *Billboard*'s Hot 100 between the chart's launch in 1958 and the end of August 2012 when our data collection ended. These songs comprise the set of Top songs. We also identified each of the 1,451 songs that made it onto the Hot 100 charts but never climbed above the number 90 ranking. These songs comprise the set of Bottom songs. The set of Bottom songs is 41% larger partly because it includes 10 times the number of positions (number 90 to number 100 as opposed to just number 1) and partly because these songs enter and fall off the charts more frequently. Thus, the full sample includes 2,480 songs with somewhat less than half (41%) being Top songs.

We secured audio recordings of as many of the 2,480 songs as possible, being extremely careful to obtain only the version that made it onto the charts. In other words, we did not rely on versions re-recorded by the original artist, covers, or remixes, unless it was one of these that made it onto the chart (e.g., Current's disco version of Bill Conti's Theme from Rocky "Gonna Fly Now" made it to number 98 in 1977). This resulted in a set of 2,399 songs for which recordings could be obtained (97% of all songs).

We employed a team of graduate students at one of the premier music schools in the USA to code the types of instruments and vocals audible on each recording. Each was trained and worked under the direction of a fourth-year PhD student in Music Theory whose dissertation focused on American pop music. To insure the reliability of these measures, we drew a set of 100 songs randomly and had these songs coded by two judges (Krippendorff's $\alpha = 0.74$ CI [0.606, 0.869]). The coders differed most with regard to guitar tone, or what they considered a clean electric guitar as opposed to a distorted electric guitar. This is due to the spectrum of sounds between clean and distorted and an inherent subjectivity in where to draw the line.

The types of instruments and descriptive statistics for the instruments deemed audible in the 2,399 songs are shown in Table 1.

Vocal timbre is determined in part by physical characteristics including age, race, and gender. Further, gender and race have been reported to impact a listener's response to popular music (Frith, 1983). A team of research assistants identified the age of the primary vocalist at the time the song was recorded utilizing various sources (e.g., books on popular music, biographies, Wikipedia). In addition, two judges independently identified the gender and race of each lead vocalist. Gender addresses the dominant voice or voices on the track and was determined by the audible voices on the recording and checked against the name and picture of the artist. Inter-judge reliability was good (Cohen's $\kappa = 0.79$, 95% CI [0.729, 0.848]). When there was disagreement, it primarily concerned identifying who was the "lead" vocalist. Race (ethnicity), when not explicitly identified in biographical information accessible online, was judged based on pictures of the artist accessed in Google images. The judges coded the lead vocalist's race either as Caucasian or not Caucasian. Inter-judge reliability was extremely high (Cohen's $\kappa = 0.95$, 95% CI [0.923, 0.985]).

Table 1. Occurrences of different instrument types in top and bottom songs.

Instrument type (Property space)	Number (%) of occurrences	Number (%) in Top songs	Number (%) in Bottom songs
Lead vocals	2,314 (96%)	1,001 (43%)	1,313 (57%)
Background vocals	1,095 (46%)	618 (56%)	477 (44%)
Bass guitar	2,117 (88%)	862 (41%)	1,255 (59%)
Clean guitar	931 (39%)	358 (38%)	573 (62%)
Distorted guitar	436 (18%)	194 (45%)	242 (55%)
Acoustic guitar	523 (22%)	206 (39%)	317 (61%)
Piano	719 (30%)	260 (36%)	459 (64%)
Electric piano	370(15%)	240 (65%)	130 (35%)
Synthesizer	619 (29%)	307 (50%)	312 (50%)
Organ	214 (9%)	82 (38%)	132 (62%)
Strings	547 (23%)	254 (46%)	293 (54%)
Horns	496 (21%)	186 (37%)	310 (63%)
Drums	2,300 (96%)	973 (42%)	1,327 (58%)
Fiddle	50 (2%)	10 (1%)	40 (3%)

Note: N = 2,399 songs (1,028 Top and 1,371 Bottom).

Our investigation into the impact of instrumentation on popularity integrates two different analytic techniques: (1) a qualitative analysis examining specific combinations of instruments and their association with the place a song reached in the charts; and (2) a quantitative analysis examining how the number of different instrument types audible in a song affected its chances of becoming a Top (vs. Bottom) song. A mixed methods approach generally is viewed as providing more effective investigative tools than an isolated qualitative or quantitative approach (Tashakkori & Teddlie, 2003).

Qualitative analysis

Method

For the qualitative analysis, we employed crisp-set QCA. The premise underlying QCA is that for certain types of outcomes, the best causal explanation involves complex combinations of the causal variables (for an overview, see Ragin, Berg-Schlosser, & de Meur, 1996). For these types of outcomes, the direct impact of single variables is not very informative (as might be the case in a typical regression); instead, the outcome depends on how the variables are combined, creating one or more specific "recipes." Thus, the data are never disentangled at the level of the individual variable but instead are kept together as the "bundle" associated with a specific outcome. Methodologically, QCA allows the researcher to derive which bundles or combinations lead to particular outcomes.[4]

Given our data, QCA is perfectly suited for our purposes as it allows us to identify whether specific combinations of instruments contribute to one of two outcomes. The specific combination of instruments in our data depend on whether or not an instrument type is audible, with each denoted as either present or absent, while the two outcomes of interest are defined as whether the song was a Top song or a Bottom song. Using set theory principles and Boolean algebra logic, QCA identifies the conditions—those specific instrument combinations—sufficient to lead to a song becoming a Top rather than a Bottom song, and vice versa. Note that

QCA determines its configurations based on judging both which instruments *should* be included and which instruments *should not* be included in the mix.

The first step in our analysis is to define the property space, in this case the list of all possible binary combinations of individual instruments. Originally, 14 different principal instrument types were identified as audible in our sample. Our intention was not to identify the full set of instruments actually played on a song (this could have been accomplished in other ways) but rather to identify which instrument types are audible to the listener. An instrument needed to make more than a very brief appearance to be included but did not need to be present throughout the entire song.

In some instances, categorization required some discretion on the part of the judges. For example, sounds that were "synthetic" sounding or clearly artificially produced were coded as a synthesizer. This included keyboards that were obviously not an electric piano or another acoustic keyboard instrument (e.g., harpsichord). We should also point out that synthesizers and electric pianos were not available in the 1950s and early 1960s, which precludes their presence in the very earliest part of the sample. Drum machines, while often distinguishable from actual drums in that they are artificial sounding, were not coded as synthesizers but were collapsed into drums. Violins, cellos, and other "orchestral" string instruments were lumped together, as were horns. Categories such as strings and horns included their synthesized equivalent only if the sound appeared natural. Our primary objective was consistency. Further details as to how instrument types were coded are available from the authors.

Upon closer inspection, we found lead vocals and drums are almost always present, appearing in 96% of songs. Conversely, fiddle was almost never present, appearing in only 2% of songs (See Table 1). These three specific instrument types would not help discriminate between combinations appearing in Top versus Bottom songs and consequently were excluded from the analysis. With the remaining 11 types of instruments, there are 2^{11} or 2,048 possible combinations that could have presented themselves. Each instrumental configuration represents a possible timbre mixture. Our set of songs included 403 (20%) of all possible combinations.

It is interesting to consider which instrument combinations occur most frequently in the songs in our dataset (across both Top and Bottom songs). The two most common combinations each include a synthesizer, in addition to lead vocals and drums, and their relative ranking is determined by either adding a bass guitar (number 1) or not (number 2). Importantly, no other instruments are present. The third and fourth most common combinations again include lead vocals and drums but add bass and a clean electric guitar (perhaps what most people consider the archetypal pop band). They are differentiated by number 3 excluding background vocals and number 4 including background vocals. Again, no other instruments are present. While these are the most common combinations, our interest was in which combinations are present in Top songs and not Bottom songs and which combinations are present in Bottom songs and not Top songs.

Thus, the second step in our analysis was to identify which configurations among the 403 present in our dataset are systematically associated with each outcome of interest. This is accomplished by calculating the proportion of cases consistent with being a Top song and not a Bottom song for each configuration and by calculating the exact opposite, the proportion of cases consistent with being a Bottom song and not a Top song. When consistency (songs that are in one configuration set and in one outcome set) is very high, that configuration is considered a "sufficient" condition for that outcome. Consistency, simply defined, is the percentage of cases in a particular configuration that also belong to the outcome of interest.

We began by setting a frequency cutoff such that only configurations present in 10 or more songs would be considered sufficient for the outcome. This was done to rule out measurement error and very rare combinations or events (Fiss, 2007). This resulted in 68 potential instrument combinations remaining in our data. Next, we fixed the consistency threshold for considering a configuration as "quasi" sufficient at .75 (i.e., at a minimum 75% of the cases of a specific configuration must be associated with the outcome of interest) according to QCA guidelines (Ragin, 2000).

The third and final step is the logical reduction of the solutions. Using Boolean algebra rules and counterfactual analysis, we used QCA software to prune the consistent configurations of logically redundant elements. In this way, we focused on the most parsimonious solutions available (Fiss, 2011) as our goal was to identify any "core" instrumental combinations associated with being a Top versus Bottom song and vice versa.

Results

We conducted two separate analyses, one for Top songs and one for Bottom songs. Sufficient conditions were first identified using fsQCA software (http://www.u.arizona.edu/~cragin/fsQCA/software.shtml); their consistency was then tested against the benchmark using the "fuzzy" suite in Stata. The results identified two core instrument configurations sufficient for being a Top song and three core instrument configurations sufficient for being a Bottom song. These five combinations are presented in Table 2. When interpreting these results, it is important to recognize that a variety of different instrument combinations drawn from the 68 potential instrument combinations are nested within these configurations. In other words, the sufficient configurations provided in Table 2 indicate the core instruments, and adding any omitted instruments (not specifically excluded) to form a specific instrument combination in our data would not affect the consistency (i.e., inclusion in being a Top or Bottom song).

The two core instrument configurations sufficient for being a Top song were present in 102 songs in our dataset with 90 being Top songs (~88% consistency). The 90 songs with these core instrument types account for 9% of all Top songs. These include: (1) the presence of background vocals, a synthesizer, and a clean guitar; and (2) the presence of background vocals, a synthesizer, and a distorted electric guitar. The specific songs including these configurations are provided in Appendix 1 (see supplemental online section). Due to the fact they are "core" configurations, we should point out again that other instruments can be layered on top. In addition, doing so implies that these solutions may have some overlap (e.g., a song with background vocals, a synthesizer, a clean guitar, and a distorted guitar would fit into both configurations). Songs shared by both configurations are highlighted in Appendix 1.

The three core instrument configurations sufficient for being a Bottom song were present in 410 songs in our sample with 336 (~82% consistency) being Bottom songs. The 336 songs with these core instrument types account for 23% of all Bottom songs. Interestingly, the core configurations sufficient for identifying a Bottom song all exclude background vocals (recall QCA indicates which instruments should *not* be included). In addition to excluding background vocals, these configurations include: (1) an acoustic guitar, acoustic piano, and no strings; (2) a clean guitar and acoustic piano; and (3) a bass guitar, synthesizer, and no electric piano. The specific songs including these three configurations are provided in Appendix 2 (see supplemental online section). Again, overlapping songs are highlighted.

It is worth reiterating that configurations characterizing Top and Bottom songs are distinguished by the presence and absence of background vocals, respectively. This may not seem all that surprising after watching the 2013 documentary *Twenty Feet from Stardom*, the story of the back-up singers behind some of the greatest musical legends of the 21st

Table 2. Instrument combinations that lead to becoming a Top song and a Bottom song.

Top songs	Instrument type 1	Instrument type 2	Instrument type 3	Instrument type 4	Consistency	F*	*p*
Configuration 1	Background vocals	Clean guitar	Synthesizer		0.90	16.65	0.000
Configuration 2	Background vocals	Distorted guitar	Synthesizer		0.88	6.85	0.009
Bottom songs	Instrument type 1	Instrument type 2	Instrument type 3	Instrument type 4	Consistency	F*	*p*
Configuration 1	(Background vocals)	Acoustic guitar	Acoustic piano	(Strings)	0.89	10.19	0.001
Configuration 2	(Background vocals)	Clean guitar	Acoustic piano		0.837	6.90	0.009
Configuration 3	(Background vocals)	Bass guitar	(Electric piano)	Synthesizer	0.80	4.21	0.040

Note. Exclusion of an instrument is denoted by parentheses.
*Benchmark proportion = 0.75.

Table 3. Distribution of songs with QCA Top and Bottom configurations across time.

Year	Top songs (number 1 Songs)		Bottom songs (no. 90–no. 100 songs)		
	Configuration 1	Configuration 2	Configuration 1	Configuration 2	Configuration 3
1958-1969	0	1	8	52	2
1970–1979	6	7	18	45	11
1980–1989	41	34	1	12	78
1990–1999	10	4	6	11	84
2000–2012	5	2	14	3	86
Total	62	48	47	123	261

century (e.g., Bruce Springsteen, Stevie Wonder, Sting, and the Rolling Stones). As pointed out in the Oscar-winning film, background vocals can elevate the perceived quality of a song's sound; a single vocal line can often sound lonely, while in contrast the dynamic interaction between a lead singer and one or more background singers may sound fuller, providing the appearance of the presence of an ensemble. Of course, this reasoning is purely speculative on our part.

We should point out that the configurations identified by QCA are not even distributed across time (see Table 3). It appears by looking at Table 3 that for Top configurations 1 and 2, these were most common in the 1980s and 1990s. This may be due, at least in part, to the rise in popularity of synthpop (aka electropop and technopop), which features the synthesizer as the dominant (i.e., core) musical instrument. With regard to Bottom configurations, configuration 1 is common in the 1960s and 1970s, while configuration 2 is common in the 1980s, 1990s and 2000s. The relative unpopularity of clean guitar and acoustic piano as core instruments may reflect the relative rise of rock and fall of folk best exemplified by Bob Dylan's controversial use of an electric guitar at the 1965 Newport Folk Festival. With respect to configuration 3, subgenres that emerged from punk such as grunge and indie rock may have contributed to the decline in popularity of the synthesizer as a core instrument. While our data track the popularity of these configurations; these interpretations are purely speculative on our part and require deeper historical analysis.

Another interesting insight provided by our qualitative findings is that some instruments, notably clean electric guitar and synthesizer, are present in configurations supporting both success and relative failure dependent on which other instruments they are combined with. The presence in both categories supports the relevance of looking at recipes or timbre mixtures as opposed to individual timbres associated with single instruments. The role of a single instrument may be negligible because of causal asymmetry (i.e., discerning separately which recipes lead to being a Top song and which lead to being a Bottom song).

To illustrate how our results can correspond with outcomes derived using more traditional methods, we used the configurations provided for Top songs to anticipate a potential interaction effect between background vocals and synthesizer on the chance a specific song became a top (vs. bottom) one. A logit model, controlling for the aforementioned covariates, confirms the presence of a significant interaction ($b_{int} = 2.52$; $z = 3.28$, $p < .001$) such that the positive effect of background vocals is, on average, enhanced by the presence of a synthesizer (See Figure 1). Without the benefit of QCA, the researcher would be confronted with the task of exploring which among all of the possible interactions with 11 types of instruments are most likely to lead to the outcome of interest.

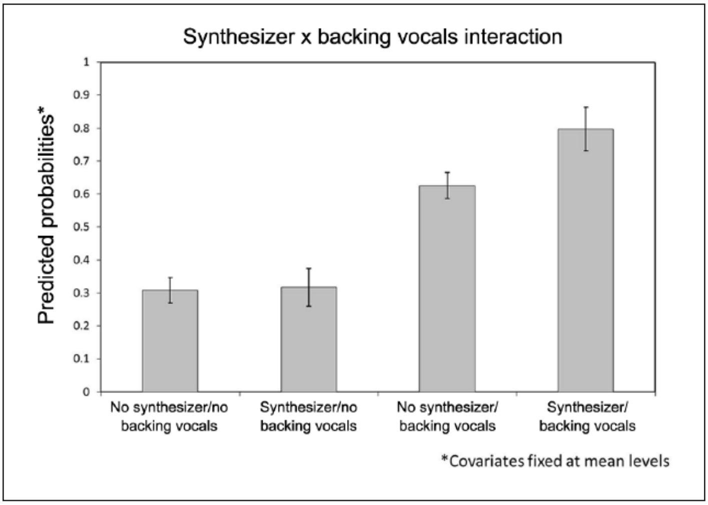

Figure 1. Predicted probabilities of being a number 1 song with variations of the presence and absence of both synthesizer and background vocals. Error bars indicate 95% CI.

Discussion

Songs that reach number 1 on *Billboard*'s Hot 100 exhibit extraordinary popularity and are, from the industry's and market's perspective, bigger hits than songs that stay at or near the bottom of the charts. Using a comprehensive sample that included all of the number 1 hits (Top songs) and a comparison set of all songs that never climbed above number 90 (Bottom songs), our analysis identified two distinct recipes, or core instrument configurations, consistently associated with Top songs and three recipes, or configurations, consistently associated with Bottom songs.

The results provide considerable evidence that there exist specific core instrument combinations that are more preferred and less preferred by the market, albeit, often in specific time periods, and that they are either associated with songs that reach the pinnacle of popularity or those that fall short. Common across core configurations sufficient to be a Top song is the inclusion of background vocals, while common across core configurations sufficient to be a Bottom song is the exclusion of background vocals. This appears to confirm the importance of background vocals in increasing one's chances of reaching the top of the charts.

Our analysis reveals that—historically—the aforementioned combinations have consistently led to the associated outcomes. Our analysis helps explain the relative success of a percentage of, but not nearly all, Top songs. And there are always exceptions (i.e., the core configurations are at times associated with the opposite outcome). Reasons other than the choice of instruments contribute to a song's popularity, or lack thereof. We do not want to overstate the implications of our findings. Certainly marketing, artist preference, genre, catchy hooks, and other factors play a role in determining a song's relative success. For example, the star power of Rihanna may overcome any effect of instrumentation. The total number of instruments might be another factor.

While we believe we have identified specific core instrument combinations a significant portion of the market tends to prefer more versus less, we do not know how many instruments these listeners prefer. In other words, in addition to the three instrument types in our two core

Table 4. Distribution of songs with the different number of instrument types.

Number of instrument types in song	Number of top songs	Number of bottom songs	Number of total songs
1	1	0	1
2	10	13	23
3	75	74	149
4	177	297	474
5	284	426	710
6	280	360	640
7	125	161	286
8	64	33	97
9	10	7	17
10	2	0	2
Total	1,028	1,371	2,399

configurations for Top songs, how many other instruments, if any, should be added? A preliminary inspection of our qualitative results reveals that Top songs include a greater number of instruments. Sufficient Top songs contain, on average, 1.63 more audible instruments than the remaining songs ($M_{\text{sufficient}} = 6.86$ v. $M_{\text{remaining}} = 5.24$, $F = 163.2$, d.f. = 1, $p < 0.001$). Whether and how the total number of instruments audible in a song impacts preferences across the entire sample is taken up in the subsequent quantitative analysis.

Quantitative analysis

Method

We utilize logit regression to test the relationship between the number of different instrument types audible and the probability of being a Top versus a Bottom song. The number of instrument types audible for each song is our principal independent variable and was coded as "Number." The squared term "Number-sq" was included to account for a potential non-linear relationship. Three types of control variables were included in our analysis. First, we consider key demographics features of the artist. The lead vocalist's age in years was coded as "Age" in our model. For gender, the reference category is "Male" while the category "Female" indicates whether the lead vocalist is female, and "Both" indicates whether the lead vocals are sung by both a man and woman. Each performer's race was also classified in terms of being "Caucasian" or not, which, as mentioned earlier, was based on research by judges who relied on contemporaneous photographs of the principal performers whenever necessary. We excluded songs for which we could not locate reliable information on the performer. This left 2,112 songs (88% of the songs used in the QCA) including 995 of the original 1,029 Top songs (97%) and 1,117 of the original 1,451 Bottom songs (77%). Descriptive statistics with respect to the number of instrument types for these songs are reported in Table 4.

Second, some artists entered in our dataset more than one time, either with a Top or a Bottom song, or both. Therefore, songs cannot be considered fully independent, and we account for within-artist correlation by employing a robust standard error clustered at the artist level for testing our parameters. Third, our dataset spans 55 years, and some omitted contextual variables can affect the chances of rising in the charts over time. We capture time dependence (account for changes over time) through a fixed time effect model that includes (54) year

Table 5. Impact of the Number of Instruments on the Probability of Becoming a Top Song.

Variable	Model 1 (yearly dummy)			Model 2 (cubic polynomial)		
	Coefficient	z*	Robust std. err.	Coefficient	z*	Robust std. err.
Age	0.944**	−7.29	.008	0.954**	−6.54	.007
Both	1.082	0.43	.200	1.052	0.28	.187
Female	0.960	−025	.157	0.985	−0.10	.154
Caucasian	1.095	0.70	.142	1.231	1.69	.151
Number	0.437**	−3.59	.109	0.435**	−3.71	.098
Number-Sq.	1.091**	4.13	.023	1.088**	4.14	.022
Year Dummies	Included					
Year				1.131**	3.20	.043
Year-Sq.				0.998	−1.38	.002
Year-Cubic				1.000	0.09	.000
Constant	27.88**	4.75	19.535	6.558**	2.73	4.511

Note. Exponentiated coefficients.
*Standard errors adjusted for 1,449 individual artists.
**$p < .01$.

dummy variables (see Model 1, Table 5). As a robustness check, we also model time dependence using a cubic polynomial time specification (see Model 2, Table 5), which is particularly appropriate for binary data and to evaluate trend effects (Carter & Signorino, 2010).

Results

The regression output utilizing robust standard errors clustered at the artist level is presented in Table 5. The number of instruments ("Number") shows a positive curvilinear effect on whether a song is preferred in terms of its likelihood of being a Top versus Bottom song. This curvilinear effect of the number of instruments is confirmed through a likelihood ratio test, both against the model with covariates only ($\chi^2 = 24.40$, $p < .001$) and a model with only a linear effect of the number of instruments ($\chi^2 = 17.40$, $p < .001$). The effect is robust across different time specifications (see Models 1 and 2). These results confirm what we observed earlier; in general and controlling for time, the greater the number of instruments that are audible (sometimes referred to as "richness" or "density"), the greater the likelihood of being a Top as opposed to a Bottom song. The estimated model has good predictive capability (area under the ROC curve = .703; the raw classification accuracy is 64.5%) and goodness-of-fit level (Hosmer-Lemenshow $\chi^2_{(8)} = 7.7$, $p = n.s.$).

The predicted probabilities for being a Top song with different numbers of instrument types are shown in Figure 2. What stands out is the likelihood of being a Top song is minimal when there are three to five instrument types audible, and the likelihood increases as the number of instruments either decreases or increases. Thus, distinctive instrumentation (i.e., very few or very many, which are indeed less common) seems to stand out in a positive way. An important result worth highlighting is how the predicted probabilities for being a Top song increase more than proportionally when additional instruments are added after the sixth. This is made apparent in Figure 3, which illustrates the marginal effect of adding one additional instrument; it reveals how the sixth instrument begins adding a significant positive effect to the probability of being a Top rather than Bottom song and how each additional instrument significantly increases this probability.

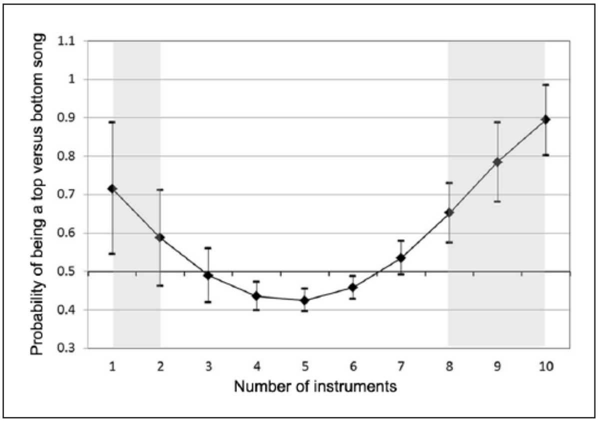

Figure 2. Predicted probabilities of being a number 1 song at different numbers of instrument types. Error bars indicate 95% CI.

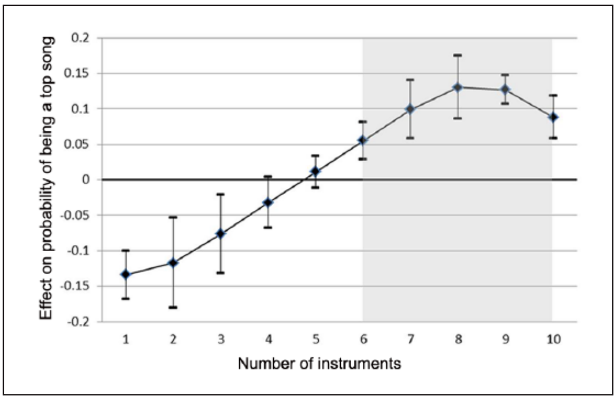

Figure 3. Average marginal effect of adding one additional instrument type on the probability of being a number 1 song. Error bars indicate 95% CI.

As an additional step, to investigate the extent to which our results were time-specific, we reran the regression including interaction terms for the number of instrument types audible for each song and the year of the song, specifying time with the cubic polynomial. A test of the interaction effect is highly significant [$\chi^2(3) = 124.6$, $p < .00$] suggesting that preferences regarding the inclusion of more or less instruments changed over time. The marginal impact of adding another instrument to the mix also varied across time (see Figure 4). Figure 4 illustrates how fewer instruments increased the chance of being a Top song in the 1960s and in the past decade or so (2000s), while adding instruments increased the chance of being a Top song during mid-70s, 1980s, and 1990s.

Discussion

Overall, our quantitative analysis reveals an interesting additional finding regarding instrument combinations. If we consider the typical song in our data includes five instruments ($\mu = 5.3$), we have strong evidence that songs with more or less than the average number of

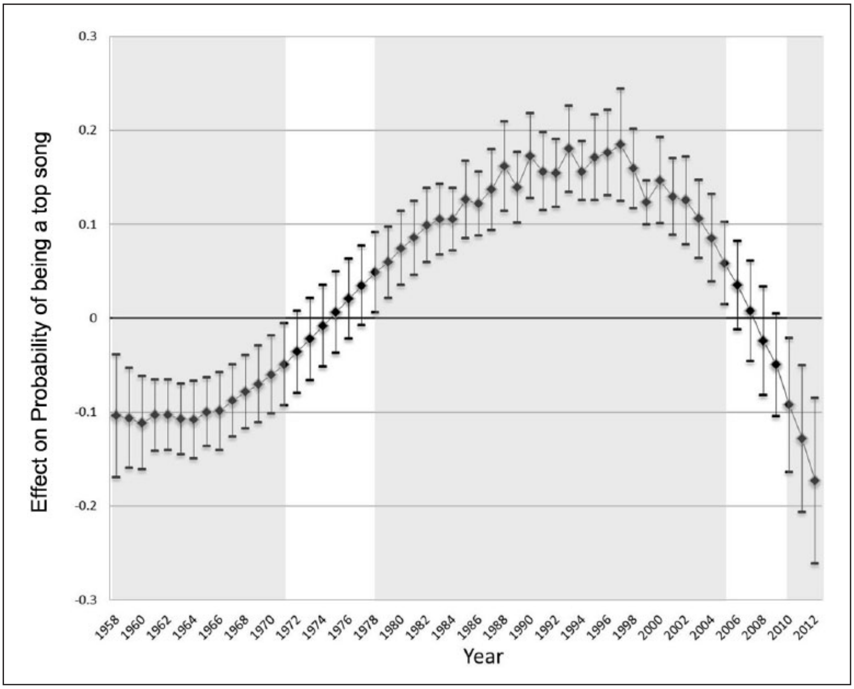

Figure 4. Marginal effect of adding one additional instrument type (on average) on the probability of being a number 1 song—plotted over time. Error bars indicate 95% CI.

instrument types stand a greater chance of being more popular (i.e., a Top as opposed to a Bottom song). Perhaps the most straightforward interpretation of our results is that songs by artists who utilized a surprisingly low or high number of instruments—at specific points in time—tended to stand out. This effect is most striking when one looks at the increasing likelihood of a song being a Top song as additional instruments are added beyond the fifth instrument. Another point worth highlighting is that, while people appear to have preferred relatively simple instrumentation during the 1960s, their preferences shifted toward songs including a greater number of instruments in the 1970s. Such a preference for "richer" songs also characterized the two following decades, while in recent years consumer preferences seemed to steer again toward songs with less instruments.

Conclusion

In their work examining how musical properties influence preferences, Rentfrow et al. (2012, p. 164) suggest it would be informative to code musical pieces for different instrumental families to "gain even more precise information about the nature of preference factors." This is what we have attempted to do using relative popularity as a proxy for the market's preference. We do so using data on instrumentation from some of the nation's most popular songs from the past 55 years. Rather than gauge respondents' preferences in a lab based on song snippets, we relied on *Billboard*'s rankings as a measure of popularity and therefore in all likelihood relative preferences. In this way, we were able to document the systematic relationship between the number and type of instruments (i.e., timbre mixtures) in a song and its relative popularity in an ecologically valid way.

We identify two configurations that result in a song being more popular, even among the nation's most popular songs. These particular configurations are present far more consistently among chart-topping number 1 singles as opposed to songs that made the charts but never climbed above number 90 in the rankings. We also identify three configurations associated with songs that have stagnated at the bottom of the charts as opposed to climbing to the top. Conversations with a number of professional musicians regarding our results confirmed that they are not inconsistent with any musical norms they employ in the profession and are, in fact, seen as sensible. One observation that was particularly relevant was the following: in pop music the mid-range frequency of a song is typically given to a harmony instrument, and that instrument will often be a guitar, synthesizer, or electric piano. Even if a song has many layers of instruments, the core nearly always includes a keyboard instrument or a guitar. Our results appear consistent with this observation.

A few musicians were surprised to learn only 56% of number 1 songs included background vocals, something they thought in hindsight would be even more ubiquitous. Further, this research documents the positive impact of adding additional layers of instruments to a song over and above the five instrument sounds most typically present. It may be that people like the "richer" sounds that often accompany a greater number of instruments, especially for pop music. Similarly, a single instrument stands out with listeners as well. It may also be that the mass market has tired of four and five instrument sounds and is looking for something different.

Music critics frequently lament the perceived homogenization of popular music. According to Reuters (2012), a research team led by artificial intelligence specialist Joan Serra at the Spanish National Research Council examined the music archive known as the Million Song Dataset and determined pop songs have become intrinsically louder and blander over time in terms of the chords, melodies, and types of sound used. On the one hand, our findings regarding the impact of specific instrument configurations on preferences could be viewed as consistent with this commentary. Consider that of the 2,048 possible instrument combinations, only 20% appear even once in our sample, and of those having greater than 10 songs, there were only 68 combinations. This concentration appears to reflect a limited diversity in instrumentation in pop music. On the other hand, our findings could be considered as revealing the benefits bestowed upon artists who attempt to make their sound more distinct. The artist is free to, and perhaps should, add additional instrument sounds on top of one of the two successful core configurations; this could be viewed as adding a personal embellishment to a generic recipe. Music performed in a style that deviates from the norm has been shown to heighten the pleasure of listening (Hargreaves & Casetell, 1987).

Of course, our recommendations for orchestrating a hit song are unavoidably overly general and should be interpreted as they are intended, suggestions based on a retrospective view of what has tended to work and not work in the past. We acknowledge that orchestration in terms of instrument types is only a small part of the story. Genre is important as commercially successful rap songs and dance tracks require very different instrumentation as compared to traditional rock music and style considerations must be taken into account. In addition, a lot of effort currently is being invested in creating novel instrumental sounds as well as experimenting with different sound mixes. We could not account for studio production techniques. Future research could explore the role of producers (e.g., American record producer Frederick Jay "Rick" Rubin) and studios that provide distinctive sounds (e.g., Sound City Studios reportedly had a very particular sound when it came to recording drums). We also recognize many factors unrelated to the music itself often are involved in determining what songs are popular, which the current studies did not (and could not) control for. Consider that the International Federation of the Phonographic Industry (IFPI) estimates that it takes approximately $700,000 to $1,400,000 to break an artist in a major recorded music market, which includes $200,000 to $500,000 in marketing and promotion (http://www.ifpi.org).

An important next step would be to replicate and refine our findings utilizing well-controlled lab experiments. Background vocals could be added and removed and various instrument combinations for the same song could be tested systematically. Interestingly, the authors have learned from discussions with executives at True Music in Thailand that the company is conducting just such real-world marketplace experiments utilizing their local commercial music streaming service. They are monitoring plays and listener feedback for different versions of the same song. In this way, True Music has real-time, online market data regarding how different orchestrations impact a song's popularity and thus consumer preferences. It would behoove researchers who are interested in which acoustic qualities of a song may be preferred to work with a firm such as True. This is something we are eager to do.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Supplemental material

Appendices 1 and 2 (song configuration lists) are available on the journal website: msx.sagepub.com/supplemental.

Readers interested in the raw data should contact the authors directly.

Notes

1. We acknowledge that we are simplifying things and that timbre is usually correlated with other psychoacoustic features such as loudness and pitch (e.g., Balzano, 1986) that also distinguish an instrument's role in a song.
2. "Polyphonic" here means the presence of more than one instrument and should not be confused with the term "polyphony" in music theory referring to several simultaneous melodies of equal importance.
3. By instrument, we mean the type of instrument. We make no distinction between whether one or two acoustic guitars are audible.
4. Contrary to traditional correlational analyses, in QCA if a particular combination is systematically associated with the presence of a particular outcome, it does not mean that the absence of this combination leads to the absence of the outcome. QCA is based on principles of causal asymmetry, meaning that the presence (e.g., being a Top song) and the absence of an outcome (e.g., not being a Top song but a Bottom song), respectively, may require different causal explanations.

References

Alluri, V., & Toiviainen, P. (2009, August). In search of perceptual and acoustical correlates of polyphonic timbre. In J. Louhivuori, T. Eerola, S. Saarikallio, T. Himberg, & P.-S. Eerola (Eds.), *Proceedings of the 7th Triennial Conference of the European Society for the Cognitive Sciences of Music* (pp. 5–10). Jyväskylä, Finland: University of Jyväskylä.

Alluri, V., & Toiviainen, P. (2010). Exploring perceptual and acoustical correlates of polyphonic timbre. *Music Perception: An Interdisciplinary Journal*, *27*(3), 223–242.

Aucouturier, J. J. (2006). *Dix expériences sur la modélisation du timbre polyphonique* [Ten experiments on the modeling of polyphonic timbre]. Unpublished doctoral dissertation, Universite Paris 6, France.

Aucouturier, J. J., Pachet, F., & Sandler, M. (2005). "The way it sounds": Timbre models for analysis and retrieval of music signals. *IEEE Transactions on Multimedia*, *7*(6), 1028–1035.

Balzano, G. J. (1986). What are musical pitch and timbre? *Music Perception*, *3*(3), 297–314.

Bradlow, E. T., & Fader, P. S. (2001). A Bayesian lifetime model for the "Hot 100" *Billboard* Songs. *Journal of the American Statistical Association*, *96*(454), 368–381.

Carter, D. B., & Signorino, C. S. (2010). Back to the future: Modeling time dependence in binary data. *Political Analysis*, *18*(3), 271–292.

Cutietta, R. A., & Foustalieraki, M. (1990). Preferences for select band and non-band instrument timbres among students in the United States and Greece. *Bulletin of the Council for Research in Music Education*, *105*, 72–80.

Darrow, A., Haack, P. A., & Kuribayashi, F. (1987). Descriptors and preferences for eastern and western musics by Japanese and American nonmusic majors. *Journal of Research in Music Education*, *35*(4), 237–248.

Dhanaraj, R., & Logan, B. (2005, September). Automatic Prediction of Hit Songs. In *Proceedings of Ismir* (pp. 488–491). London, UK: ESCOM.

Fiss, P. C. (2007). A set-theoretic approach to organizational configurations. *Academy of Management Review*, *32*(4), 1180–1198.

Fiss, P. C. (2011). Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal*, *54*(2), 393–420.

Frith, S. (1983). Popular music 1950–1980. In G. Martin (Ed.), *Making music: The guide to writing, performing, and recording* (pp. 18–48). London: Muller.

Giordano, B. L., & McAdams, S. (2010). Sound source mechanics and music timbre perception: Evidence from previous studies. *Music Perception*, *28*(2), 155–168.

Gjerdingen, R. O., & Perrott, D. (2008). Scanning the dial: The rapid recognition of music genres. *Journal of New Music Research*, *37*(2), 93–100.

Handel, S. (1989). *Listening: An introduction to the perception of auditory events*. Cambridge, MA: MIT Press.

Hargreaves, D. J., & Casetell, K. C. (1987). Development of liking for familiar and unfamiliar melodies. Bulletin of the *Council for Research in Music Education*, *91*, 65–69.

LeBlanc, A. (1981). Effects of style, tempo, and performing medium on children's music preference. *Journal of Research in Music Education*, *29*(2), 143–156.

McAdams, S. (1993). Recognition of sound sources and events. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp.146–198). Oxford: Clarendon Press.

McAdams, S., & Bregman, A. (1979). Hearing musical streams. *Computer Music Journal*, *3*(4), 26–43.

Menon, V., Levitin, D. J., Smith, B. K., Lembke, A., Krasnow, B. D., Glazer, D., et al. (2002). Neural correlates of timbre change in harmonic sounds. *NeuroImage*, *17*(4), 1742–1754.

Molanphy, C. (2013). How the hot 100 became America's hit barometer. NPR *The Record*, [web log]. Retrieved from http://www.npr.org/blogs/therecord/2013/08/16/207879695/how-the-hot-100-became-americas-hit-barometer

Neville, M. (Dir.). (2013) *Twenty feet from stardom* [film]. USA: Gil Friesen Productions, Tremolo Productions.

Pachet, F., & Roy, P. (2008, September). Hit song science is not yet a science. In J. P. Bello, E. Chew, & D. Turnbull (Eds.), *Proceedings of ISMIR* (pp. 355–360). Philadelphia, PA. Retrieved from: http://www.lulu.com/shop/juan-pablo-bello-and-elaine-chew-and-douglas-turnbull/proceedings-of-the-9th-international-conference-on-music-information-retrieval-conference-edition/paperback/product-3548598.html

Radocy, R. E., & Boyle, D. (1997). Functional applications of music in contemporary lifes. In R.E. Radocy, & D. Boyle, *Psychological Foundations of Musical Behavior* (pp. 31–64). Springfield IL: Charles C. Thomas Publishers.

Ragin, C. (2000). *Fuzzy-set social science*. Chicago: University of Chicago Press.

Ragin, C., Berg-Schlosser, D., & de Meur, G. (1996). Political methodology: Qualitative methods. In R. Goodin & Hans-Dieter Klingemann (Eds.), *New handbook of political science* (pp. 749–768). New York: Oxford University Press.

Rentfrow, P. J., Goldberg, L. R., Stillwell, D. J., Kosinski, M., Gosling, S. D., & Levitin, D. J. (2012). The song remains the same: Replication and extension of the MUSIC model. *Music Perception*, *30*(2), 161–185.

Reuters (2012, July 26). Pop music too loud and all sounds the same: official. Retrieved from http://www.reuters.com/article/2012/07/26/us-science-music-idUSBRE86P0R820120726

Schellenberg, E. G., Iverson, P., & McKinnon, M. C. (1999). Name that tune: Identifying popular recordings from brief excerpts. *Psychonomic Bulletin & Review*, *6*(4), 641–646.

Shehan, P. K. (1981). Student preferences for ethnic music styles. *Contributions to Music Education*, *9*, 21–28.

Tashakkori, A., & Teddlie, C. (Eds.). (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.

Teo, T. (2003). Relationship of selected musical characteristics and musical preferences (a review of literature). *Visions of Research in Music Education*, *3*. Retrieved from http://www-usr.rider.edu/~vrme/