Developing Interactive Educational Songs for Introductory Statistics

Lawrence M. Lesser, Dennis K. Pearl, John J. Weber III, Dominic M. Dousa, Robert P. Carey, and Stephen A. Haddad

ABSTRACT
This article describes the process used to develop and assess an NSF-funded instructional innovation: an online collection (https://www.CAUSEweb.org/smiles/) of 28 interactive songs of high esthetic quality designed to span literature-based learning objectives of introductory statistics that develop statistical literacy and reasoning. The interactive songs are also designed to reduce statistics anxiety and require little time or instructor expertise. The songs are interactive in that the interface solicits (and provides hints and feedback on) student contributions (concepts or examples) and then plays back the song with student inputs integrated and highlighted. After providing a brief background, this article describes requirements, challenges, and opportunities in educational songwriting for the statistical sciences, then describes the intervention and how its special nature affected the development process. Pilot studies at a research university and at a majority Black two-year college showed that students found the innovation to be a good tool to help their learning, reduce their anxiety about statistics, have an easy to follow interface, and use high quality songs. Analysis of log files from the use of the software shows some evidence of better performance on assessments after use and informs improvements of the automated feedback. Supplemental materials for this article are available online.

1. Introduction
This article describes the process to develop and assess an instructional innovation: the interactive song. After providing a brief background, we describe requirements, challenges, and opportunities in educational songwriting for the statistical sciences. In subsequent sections, we describe the intervention and how its special nature affected the development process.

We focus on songs with lyrics designed to support the learning of introductory statistics. As reviewed in Crowther et al. (2016), the use of song in higher education spans a wide variety of disciplines within and beyond STEM and its possible benefits include reduced stress or anxiety, improved recall, and increased motivation or engagement. A well-written jingle or song with rhymed lyrics can help content make a more memorable impression on students than that same content delivered in prose form.

The main goal of our project is to develop interactive song resources that require little instructor time or expertise, but that reduce statistics anxiety (Chew and Dillon 2014) and have high impact on developing statistical literacy and reasoning (delMas 2002). Lesser (2012) and Rogness, Kaplan, and Fisher (2016) discuss the value of concise high bang-for-the-buck statistics activities.

A key purpose of this article is to illustrate the process of developing this distinctive intervention of interactive songs for teaching statistics. By “interactive,” we mean building supplemental online activities that include student contributions that integrate with the song. Here, students interact with prompts to provide examples and demonstrate conceptual knowledge (a recommendation of the 2016 GAISE College Report) about the learning objective at hand. Their responses are used to fill in components of a song highlighted (in the displayed lyric if not also with a synthetic voice) during playback. We dubbed this intervention SMILES: Student-Made Interactive Learning with Educational Songs.

SMILES resources aim to be: engaging to students and require their active input; machine gradable to provide rapid feedback; online to allow out-of-class use and avoid an instructor time burden; easy to use to avoid frustration or undue ramp-up time; efficient in making impact per student time on task; grounded in core learning objectives; usable in a manner within the instructor’s comfort zone; and amenable to experimental testing of their efficacy for our goals of increased learning and reduced anxiety. Also, recorded songs are intended to have high esthetic value by having had most of the composing done by music professionals and almost all of the performing done by music majors in a state-of-the-art commercial music studio.
2. Prior Work

Our earlier NSF-funded experiment found that songs outperformed other modalities such as cartoons in terms of learning gains by college introductory statistics students who were randomized to have such items inserted into mini-readings in their course learning management system. That paper (Lesser, Pearl, and Weber 2016) suggested possible explanations for songs’ superior performance, including stronger mnemonic potential, greater engagement or time on task (e.g., listening to a song vs. reading a cartoon), multiple input channels (i.e., auditory as well as visual), more release of dopamine (Salimpoor et al. 2011), or a greater level of activeness (e.g., inviting students to sing along).

Also, many dimensions or elements of a song (rhyme, rhythm, melody, etc.) activate listeners at least implicitly to predict what will come next. Considering all of this, the researchers saw promise in focusing on songs and exploring how to increase their apparent relatively higher level of activeness and engagement even more. The idea of making content interactive is not new, considering diverse examples such as the NSF-funded E-MATE project to develop interactive e-textbooks (Qaissaunee and Parr 2016), an automated songwriting site (Pogue 2005), the 40-book Choose Your Own Adventure series, interactive song videos such as Roy Zimmerman’s “What Would You Rather Get for Christmas?,” and the Oddcast (host.oddcast.com/demo_ttsing/) “text-to-sing” demo. However, it appears to be novel to communicate educational content via interactive songs.

3. Songwriting in Mathematical/Statistical Sciences

Background and benefits of using song in the mathematical and statistical sciences have been discussed (Lesser 2000, 2001, 2014, 2018a, 2018b) and there is a mathematical sciences presence among initiatives for educational songwriting in STEM, including the 1999–2011 activity of the Science Songwriters Association (http://www.science-groove.org/SSA/), science song blogs (such as http://singaboutscience.org/wp/), a curated database of 7000+ science songs (http://singaboutscience.org/; Crowther 2012), and song-in-STEM conferences (https://www.CAUSEweb.org/voices/). There have also been songwriting initiatives specific to the mathematical sciences, such as the National Museum of Mathematics launching (in 2015) annual international mathematics song competitions. Also, CAUSE (Consortium for the Advancement of Undergraduate Statistics Education) has conducted seven statistics song contests since 2007 and curates a searchable collection (https://www.CAUSEweb.org/cause/resources/fun/) of over 800 educational fun items that includes almost 200 statistics songs, almost all of which have soundfiles that may be played online or downloaded.

3.1. Challenges

As Lesser, Pearl, and Weber (2016) note, STEM songwriting challenges include:

(2) checking of jargon (e.g., technical words such as heteroscedasticity) for singability and rhymeability,
(3) balancing the concrete and abstract,
(4) keeping the tone inviting for novice learners so it goes beyond insiders,
(5) aligning to a specific learning objective—that is, not a generic “STEM is cool; stay in school” anthem, nor a song that merely contains (but does not illuminate) content terms, and
(6) keeping the length short to maintain interest while optimizing memorability and the use of class time.

Related to (6) are findings from a study of MOOC videos that showed almost complete engagement when videos were under six minutes (Guo, Kim, and Robin 2014). Just as best practices (Berk 2009; Aragon and Wickramasinghe 2016) for the use of video for educational purposes recommend that instructors keep videos to no more than 5–10 min to optimize student attention and engagement, this limit happens to be very consistent with the time we observed (see Section 4) that was involved in listening to and engaging with a song, including time spent singing/reading its lyrics and responding to accompanying questions.

Some challenges were song-specific. For example, in the writing of the project song “Chi-Squared Dance,” we needed several qualitative variables which each had only two syllables and an unambiguous number of categories so that the degrees of freedom number later in the song would be correct. For example, the first variable that came to mind as apparently meeting these conditions was “gender,” but it is not clear that we can view gender as having only two options (Fausto-Sterling 2000; Hardin and Miller 2019). All SMILES songs mentioned in this article can be accessed at https://www.CAUSEweb.org/smiles/songs. Publicly launched in May 2018, the collection includes a “Build a Song” option to create an interactive song and a “Composer Version” to hear the songwriters’ previously composed studio version.

3.2. Prosody

Prosody involves how music and lyrics interact and reinforce each other (Harper 2018). For example, the SMILES song “Chi-Squared Dance” sets the lyric phrase “large gap” to a large melodic leap, sets the phrase “a long right tail” to a long descending melodic phrase, and uses square dance music to playfully invoke the title. Some variations in melody were based directly on inputs students chose, such as how “Target Practice” had two possible melodic endings, where the lyric phrase referring to the estimator’s performance aligned with whether a student chose an unbiased estimator (i.e., the sample mean) or an estimator (sample range) that underestimates its population counterpart.

3.3. Interactiveness

In rough orderings of increasing interactiveness (e.g., discussed in Lesser 2017) with respect to the role of students for traditional use of song in teaching, a fairly highly interactive level of song has students contribute inputs for the lyric. There are, however, several ways in which inputs can be gathered.
Some commercially released songs have call-and-response, such as Billy Jonas’ “One!” (https://www.youtube.com/watch?v=JXi65Ms2Xyw) in which an audience can complete end rhymes in real time, even if hearing the song for the first time. Examples of math songs that particularly lend themselves to such completion tend to have important content words completing end rhymes, such as “Circle Song” (Lesser 2014), whose opening couplet is: “Take your finger ’round a jar, circumference equals 2πr.” Another example is “Correlation Song” (https://www.CAUSEweb.org/cause/resources/fun/songs/correlation-song), whose middle couplet is “If the fit supports a line, its slope and r would share the sign.” With our current project, we did not want to limit ourselves to such a “complete the rhyme” approach because of the undue constraints that would place on rhyme scheme and on the richness of inputs that could be solicited.

To help students gain comfort with the patterns and conventions of statistical language, a more general possible intervention might show some lyrics with a number of words omitted in several places (not unlike the cloze procedure used in reading assessment; Gellert and Elbro 2013) that the student must fill in (with some fill-ins having more than one reasonable completion). Fill-in-the-blank songwriting goes back many decades as a therapeutic tool (Schmidt 1983), so it may not be surprising that what appears to be the first fill-in-the-blank lyric published in a mathematical sciences periodical (at the end of Lesser 2000, inspired by Robb 1996) is an icebreaker script to solicit learners’ math identities and learning styles. We ruled out this method of letting students see the song as they contributed inputs for it because we wanted students to have anticipation rather than distraction and not have some answers tipped off by other visible parts of the lyric (Lesser et al. 2018).

Instead, our approach was largely inspired by the Mad Libs™ phrasal template word game (e.g., http://www.madlibs.com/books/), in which individuals in advance blindly supply words for given parts of speech (a plural noun, adjective, exclamation, adverb, a verb ending in “ing,” etc.) or sometimes another type of category (e.g., part of the body) that another individual uses to complete a story and then read the humorous result back to the first individual. One of the project songs most aligned to the Mad Libs style is “No One Counted On Simpson’s Paradox” where students are asked (see Appendix A) to provide examples of subpopulations with different values on a confounder without the use of statistical terminology.

The Mad Libs game has been adapted for educational purposes in several disciplines, including statistics. Trumpower (2010) adapted it to create statistics problems in which students blindly supply names of variables in response to a prompt for a specific type of variable (e.g., dichotomous variable, interval/ratio variable, etc.). Depending on the type of problem, students are asked for numerical values that will be plugged into the data summary table students will then analyze (mostly conceptually). He anecdotally claims several pedagogical advantages: that students will have no preconceptions about the (make-believe) results, but more interest in the context (since they chose variables of interest to them), and more freedom to suggest explanations without fear of being wrong.

Other SMILES songs use content-specific prompts without the usual Mad Libs focus on parts of speech, such as “Height of Confidence” where students are asked how sample size, population standard deviation, and confidence level affect the width of a confidence interval while still other songs provide hints (such as those in “Super Bowl Poll”—see screenshot in Appendix A).

The use of interactive songs is arguably consistent with various pedagogical calls for active learning. For example the GAISE College Report ASA Revision Committee (2016) states that active learning “allows students to discover, construct, and understand important statistical ideas as well as to engage in statistical thinking” (p. 18). There is also broad empirical support for active learning in the STEM education research literature, reflected in a position statement (cbmsweb.org/wp-content/uploads/2016/07/active_learning_statement.pdf) and the meta-analysis (Freeman et al. 2014) of 225 studies that concluded that active learning increases exam performance.

Interactive songs also have many aspects of personalized learning (Cavanagh 2014). For example, the songs allow students to feel ownership (when they choose songs and within-song examples of interest to them). Also, students’ individual needs and skills are taken into account by students being given feedback and having access to hints as needed. And students can pause, rewind, and replay songs as desired. These features of inviting student input and choice are noted as attractive to millennial students (Wilson and Gerber 2008). Also, we conjecture that songs for which students contribute inputs will be more memorable because of the cognitive psychology phenomenon known as the generation effect, which stipulates that generated information is remembered better than provided information (Goldman and Kelley 2009).

4. Songwriting With Inserts

Because of the lack of precedent for this project and the need for it to be robust and effective simultaneously in so many pedagogical, technological, and esthetic dimensions, the team knew that cycles of feedback would be needed throughout the process, and this iterative structure is represented by Figure 1. The top row diagrams the typical flow of the songwriting phase and the bottom row diagrams the accompanying evaluation efforts. This work is part of an NSF-funded project that includes consultation with an Advisory Board and external evaluator who are experienced teachers in varied institutional settings (see https://www.CAUSEweb.org/smiles/team) and leaders in statistics education reform, research and curriculum development.

4.1. Benefits

Even without music, there is much educational precedent and support for students contributing pieces of content into some structure that leaves spaces to be filled. For example, there is the “sentence frame” technique in which the instructor helps all learners (especially English language learners) by providing a sentence with key content words missing that a student must provide to complete a definition or report a conclusion. An example from Wagler and Lesser (2011, p. 824) is “Z is the number of ______ that a value is above the ______.” The sentence frame in Lesser and Winsor (2009, p. 6) to summarize the conclusion of a hypothesis test directly inspired our project song “A
Fitting Conclusion.” This technique is potentially helpful to any student, especially in the context of teaching English learners who may need such scaffolding when speaking in front of their peers.

Another example of scaffolding is when blanks are left in some text and students are given a word bank, consisting of words they may choose from, possibly some words more than once. An example of this in science songwriting is how certain words were omitted from the lyrics in the song “The Scientific Method” (Eldon and Eldon 1995).

Active learning (a principle of good practice in Chickering and Gamson 1987) is supported by the process of students completing the lyric by choosing examples and/or reasoning about concepts. The student role also evokes many of the verbs (e.g., construct, compose, combine, adapt) associated with higher levels of Bloom’s Taxonomy, thus elevating song-enhanced learning beyond the more common (and more passive) recall type of use.

4.2. Addressing Challenges

It soon became clear it would not generally work to use intact existing statistics songs (e.g., from the CAUSEweb Fun Collection) and identify words to omit that could be given by students. Most of those existing songs had one or more pitfalls because the songs:

- parodied a song that is unduly profane, sacred, obscure, or is copyrighted, and thus limiting use;
- were not focused deeply on a single statistics learning objective, such as including statistical terms without conceptual grounding;
- were too lengthy;
- had content too advanced for an introductory course;
- had, in constrained rhyming positions, important statistics examples that may be unlikely to be supplied in response to open-ended prompts;
- consisted of elaborated examples so that potential student inputs would need to be entire sentences rather than concise bits of input;
- were limited to lower-order thinking or recall (Lummis et al. 2017), rather than including conceptual understanding and/or real-world interpretation; and
- were not useful for a novice learner (even if interesting artistically or for community-building among insiders).

Awareness of student inputs therefore had to inform the writing of the new songs from the start. This involved many considerations and implications, including:

1. Key content words are ideally in important positions such as end rhymes (for emphasis and recall). For more open-ended student inputs, however, we could not count on such inputs being consistent with a not-yet-revealed rhyme scheme and so input places generally must avoid rhyming positions, as does this version of Mad Libs: [http://www.madglibs.com/libs/173.html](http://www.madglibs.com/libs/173.html).

2. Asking for an input containing a large number of words or syllables makes it unlikely students will come up with something useful whereas limiting students to just a few syllables can frustrate students by being too restrictive for what they want to communicate. We ended up capping student inputs at 3 or 4 syllables, as more could compromise the intelligibility of the text due to the fixed space in the music allotted to the response, and potentially result in an awkward distortion of pitch or rhythm.

3. Fixed-choice options must be checked for singability. For example, the nominal variable insert option “best class” was changed to the much more singable “your major.” In general, vowels are easiest for singing a sustained note and next easiest are syllables ending with a “liquid” or “nasal” phoneme such as L, R, M, N, or NG.

4. To maximize the intelligibility of the synthetic voice singing student inputs, the pitch and rhythm of the melodies at the points where the inserts are placed were written to make
smooth transitions between human and synthesized voices, and maintain the musical integrity of the melody. In particular, to allow as much time as possible for the longest available response, we lengthened the rhythm of some words (the ones that corresponded to the points where recorded sounds would be inserted) and quickened the rhythm of others (especially words before the insertion point). For pitch, we tried to avoid skips, and even tried to avoid having more than one note in the insertion point to help make the process of programing the pitch of the inserted material as straightforward as possible.

**General challenges.** As mentioned previously, songs that go beyond mere recall are harder to write, and more so when there will be student inputs in the song. For example, due to the statistical emphasis on analysis and inference with context, project songs attempt to link concepts to context in the prompts and/or in the song itself. Our songs are primarily at Travis Weiland’s medium level of the importance of context to a student task (Table 1 of Weiland 2016). A higher level of context integration is in our SMILES song “It Might Not Be That Bad,” in which knowledge of disease incidence/mortality rates is crucial to the screening test performance ideas often used to motivate Bayes’ theorem.

It is also generally valuable for educational songs to be concise (as seen at [https://www.CAUSEweb.org/smiles/songs](https://www.CAUSEweb.org/smiles/songs), half the songs are under 90 sec) and memorable (consistent with the work of Jakubowski et al. 2017). For an example of the latter, two of the project songs use a melody (“Twinkle, Twinkle, Little Star”) specifically described by Senthilingam (2016) as highly memorable because of its rhythmic pattern and simple musical contour.

### 4.3. Robustness and Responsiveness to Student Inputs

While developing feedback to student responses, we needed to consider more dimensions (e.g., spelling, number of syllables, identification of synonyms, nature of hints, pluralization, etc.) than we originally expected. Some synonyms were content-related (e.g., normal, Gaussian, bell-shaped) and some were not (bigger, larger, greater). Most student responses to a prompt required us to write its own specific set of feedback. Here, we drew heavily from the Natural Language Toolkit (by Steven Bird and Edward Loper of the University of Pennsylvania), Princeton’s WordNet, Carnegie Mellon’s Pronouncing Dictionary, and Stanford’s Natural Language Processing.

To give the interface more robustness by allowing a greater range of student inputs, the lyrics were constructed in some places to not merely have fixed blanks to fill in, but also places where even adjacent words or melody could be modified in response to the input, such as in the song “Target Practice” described in Section 3.2.

Each song developed for the project requires:

- lyrics aligned with content objectives (see Section 4.6) and connecting to real-world context if possible;
- that it is built for inputs (see Section 4.2);
- music that is original (or public domain) and in a genre in which lyrics are not too fast or buried;
- questions requesting students to submit a statisticalexample or statistical concept;
- feedback for each student response including hints, suggestions for students to consider after an incorrect input, software analysis of syllabic requirements for the input (required for a smooth placement of the song insert created from the student input);
- recording a version of each song with blank portions for the placement of the song insert based on the student responses;
- creating the song’s open-ended inserts using synthetic voice software;
- playback design of the song; and
- interface design.

Because songs were assumed to be more interactive when students felt like they had free choice when typing in answers, writers aimed for this first, with fixed pull-down menus of choices being a fallback option when an open-ended format was not practical. Sometimes a prompt could appear to be open-ended but in practice have a very high probability of predictably delivering a particular word or property. (For nonstatistical examples, asking “name a food that’s blue” almost surely yields “blueberries” and asking “name a one-syllable US state” can yield only “Maine.”) In “Height of Confidence,” we ask for one word to describe what happens to the width of a confidence interval under various changes (of n, s, or confidence level). Fortuitously, almost any reasonable word a student types will be a two-syllable word in trochee meter: shorter, lower, tighter, lesser, smaller, longer, wider, bigger, higher, etc.

Some discoveries of robustness involved mathematics: the SMILES song “Regression Rumba” includes the value of $r^2$ to the hundredths place based on a value of $r$ the student provided to the tenths place. It turns out that no two-digit squares of a one-digit number contain the only two-syllable digit (7), and thus generate an $r^2$ value of exactly three syllables (e.g., “point oh four” or “point one six”).

### 4.4. Optimizing Prompts to Solicit Student Inputs

Writing the song with places for student inputs is one thing, but a separate nontrivial step is developing wording for the prompts that solicit those inputs. A challenge was finding balance between not having students bogged down by doing a formal calculation or having to recall peripheral or background knowledge, forcing us to have clarity on whether a song was to be used to introduce a topic or summarize a topic—that is, whether it would precede or follow a mini-lesson.

The template for our first interactive song was piloted informally in a live technology-free setting—a September 2015 mixer for new and continuing students in a master of arts in teaching mathematics degree program. During this in-person event, the first author observed that small groups were struggling to interpret the wording of the prompts and generate answers that satisfied suggested targets in terms of number of syllables, syllabic stress, and rhyme sounds. In particular, a prompt soliciting an 8-syllable phrase ending in a particular rhyme sound proved to be too ambitious and that part of the song is now simply provided without student input. In general, being able to keep the number
of syllables flexible but small proved important in developing subsequent songs.

4.5. User Interface

The interactive online platform required preparing both an interface for delivering the student prompts as well as a song playback mechanism.

*Developing student prompts.* The prompts themselves require students to input a statistical concept or context so that the lyrics and song recording become accessible after correct responses to all prompts are given. The prompts are generally formatted as multiple-choice (sometimes in the form of matching) questions when eliciting statistical concepts or as open-ended word inserts (especially when requesting context or examples). The multiple-choice prompts are evaluated for correctness (sometimes accepting more than one answer as reasonable), and hints or other feedback are provided, depending on the student’s choices. Open-ended prompts require a wider variety of checks including spelling (auto-correcting close spellings and allowing for British spellings), grammar checks, screening for profanity, and syllable counts to ensure that the words used will fit into the song via the synthetic voice. Each hint or piece of feedback we provide corresponds to one of these conditions:

- not getting an answer could leave a student unduly “stuck” from continuing,
- a term or symbol is used that a student might not know,
- informal alternative wording of a question might help students handle the academic language,
- a word (e.g., acquit) might be unfamiliar to someone new to the English language and/or American society (e.g., our “Hypothesis on Trial” song),
- we want to teach the student something along the way by giving them a way to deduce the answer rather than repeatedly guess,
- an inputted answer may be out of range or inconsistent with another answer,
- seeing an example or visual may be helpful in understanding a definition, or
- to extend their knowledge, sometimes even for correct entries.

Checks for correctness in open-ended responses as well as the hints and feedback given are improving over time as we collect and analyze data from user inputs.

The initial pre-song prompts that students must answer correctly before hearing their version of the song include a total of 129 questions for our first 25 songs (i.e., about 5 per song). While 28% of the items offer hints, nearly every item gives feedback on incorrect responses. Also, 53 inputs involve free response and hence the use of synthetic voice on playback while the others involve forced choice answers that are highlighted on playback but generally sung (i.e., without synthetic voice) in the studio (Table 1).

*Playback with synthetic voice.* Once students appropriately enter all the inputs, they may continue on to the playback page with song lyrics and the song recording containing their inputs. The students can play the song multiple times, sing along with the lyrics line-by-line, or return to the input page and try other inputs. They can also listen to the composer’s version and use playback controls such as pause, skip-ahead and switching between the studio version and the student-created version.

We examined 28 speech synthesis engines with varying capabilities and licensing agreements. Fourteen of these engines were available as open source projects, eight as Software as a Service (SaaS) agreements, and the remaining six as paid Software Developer Kits (SDKs). Of the many text-to-speech systems available on the market, few are designed to support the necessary parameters for creating a singing voice. The Festival Speech Synthesis System/FestVox was chosen for its phoneme-based approach as well as its ability to fine-tune utterances with pitch, tone, and duration controls. Although Festival was created before Speech Synthesis Markup Language (SSML) was standardized, it offers support for singing XML markup, which provides many of the same features. One drawback to using diphone synthesis is that it produces less natural and lower quality voices. In the future, we hope to replace Festival with an engine that can fully support SSML and produce high-quality vocalizations.

4.6. Selection of Statistical Topics

Selection of the learning goals we attempted to target was informed by the GAISE College Report and/or the reasoning skills assessed in the Goals and Outcomes Associated with Learning Statistics (GOALS; Sabbag, Garfield, and Zieffler 2015) instrument, early versions of which the authors accessed. Examples of direct song connections with the latter include “Height of Confidence” (item #11 in GOALS-4 blueprint), “My Family’s Mean” (item #5), and “Everything’s Unusual” (item #20). Beyond connections with the broad principles (such as active learning mentioned in Section 3.3) of the 2016 GAISE College Report, SMILES songs also are grounded in specific details in the report. For example, the song “The Enlightened Teacher” relates to the “distinction of probabilistic sampling techniques from non-probabilistic ones” from GAISE Goal 5 (p. 10), “Throw That Out?” aligns with the call for “awareness of ethical issues associated with sound statistical practice” from Goal 9 (p. 11), and “No One Counted on Simpson’s Paradox” is supported by Goal 6 (pp. 10–11): “Multivariable relationships, illustrating Simpson’s Paradox or investigated via multiple regression, help students discover that a two-way table or a simple regression line does not necessarily tell the entire (or even an accurate) story of the relationship between two variables.”

| Type and # of prompts (with % of prompts with a hint) for 25 songs combined. |
|---------------------------------------------------------------|
| Free response input | Multiple choice input | Total |
|--------------------|----------------------|-------|
| Concept            | 22 (18%)             | 65 (17%) | 87 (17%) |
| Context            | 25 (56%)             | 4 (0%)   | 29 (48%) |
| Nonstatistical     | 3 (67%)              | 2 (0%)   | 5 (40%)  |
| Computation        | 3 (33%)              | 5 (60%)  | 8 (50%)  |
| Total              | 53 (40%)             | 76 (18%) | 129 (27%) |

*Twenty-four of these 53 items involve entering inputs about content or context that have automated checks against a list of responses considered correct or acceptable.*
As can be seen from https://www.CAUSEweb.org/smiles/songs (or Appendix B), the songs span a representative range of topics throughout a first course in statistics rather than concentrating just on, say, descriptive statistics. This breadth was intentional so that the set of songs could be a framework or supplement for a semester-long course. Beyond the topic, the content of the lyrics was also directly informed by literature, such as the previously mentioned sentence frame inspiring “A Fitting Conclusion.” Also, the merry-go-round analogy for correlation patterns in “Correlation Illustration” came from Evans (1986) and the target analogy for estimation came from Moore (1997).

5. Songwriters

5.1. External Songwriters

We sought outside songwriters with each writer having back-grounds in both statistics and songwriting (e.g., see https://www.CAUSEweb.org/smiles/team) to generate more complete coverage of statistics topics as well as to yield ideas and genres beyond what would readily come from our own process. Our initial inspiration for a songwriter collaborative was NIM-BioS’ Songwriter-in-Residence program, which was inspired by its BioSongs Project (http://www.nimbios.org/press/BioSongs) whose director later joined our project’s Advisory Board. But while the NIMBioS project required songwriters’ physical presence at its offices at least three days per week over a month to interact with NIMBioS scientists, this was not practical or necessary for our purposes since our songwriters had sufficient knowledge of the content involved (i.e., introductory statistics, not cutting-edge research) and we could use phone and email to discuss topic selection and song drafts. A Dropbox file allowed each writer to see what topics were already taken as well as to see the GOALS and GAISE documents to inspire ideas. In other words, we gave songwriters some latitude to see what topics sparked them. One constraint, however, was that to avoid dealing with possible copyright issues or seeking licenses, we required songs to be original (or at least parody songs in the public domain: imslp.org). A further benefit of not parodying songs is avoiding resistance or distraction if students have strong knowledge of the content involved (i.e., introductory statistics, not cutting-edge research) and we could use phone and email to discuss topic selection and song drafts. A Dropbox file allowed each writer to see what topics were already taken as well as to see the GOALS and GAISE documents to inspire ideas. In other words, we gave songwriters some latitude to see what topics sparked them. One constraint, however, was that to avoid dealing with possible copyright issues or seeking licenses, we required songs to be original (or at least parody songs in the public domain: imslp.org). A further benefit of not parodying songs is avoiding resistance or distraction if students have strong connections or reactions to the song being parodied (Baker 2015).

The way the first author identified suitable songwriters resembled the qualitative sampling scheme of snowball sampling. Several times, a brainstormed name or trusted colleague would in turn lead to a person who was unknown to the first author. Because her song “Amoeba Hop” has appeared on several albums and yielded an AAAS-award-winning children’s microbiology book, well-connected contemporary folk artist Christine Lavin was asked for leads and she recommended Tom Toce. Greg Crowther, in turn, suggested Monty Harper. Robbi Sherwin, a prior collaborator with the first author, recommended Amy Adler.

5.2. Internal Songwriters

Complementing the efforts of these four external songwriters was a duo of internal songwriters that each have a master’s degree in statistics and experience teaching statistics at the college level. While there was some tendency for them to divide primary focus by lyrics versus music, both writers contributed both lyrics and music. There were many examples where the education researcher (this article’s first author) brought in a “finished” song that was then significantly improved by the composition professor (fourth author) altering its harmonic structure, such as when “Hypothesis on Trial” was changed from an AABA form to AAAB form to reinforce the contrasting lyric content in the third verse. Another example was when the classical composer felt the text of “Regression Rumba” needed a final summarizing couplet for closure, which the education researcher then wrote. A total of 21 songs were written during the fall 2015 semester—half by the external songwriters (who each wrote 2 or 3) and half by the internal songwriters.

6. Pilot Study and Student Interaction with Prompts

6.1. Methods

Five of the SMILES songs were piloted in introductory General Education statistics courses at a two-year college and a research university in 2017 as part of an activity in a computer lab setting (see Table 2). Each school tested three songs compatible with their course syllabus (with one song tested at both schools) in a short time period. For example, the activity took a total of 18 minutes at the research university (averaging 6 min per song, meeting our goal of avoiding a burden on class time). Data were also collected on student actions with the interface (e.g., clicking on a button or entering a response) and on their responses to a survey. These were recorded in a log file built using xAPI statements (https://xapi.com) implemented using LearningLocker software by ht2labs and analyzed to evaluate student responses. Data rows in the log file contained session IDs along with each action the user made together with its time of occurrence.

6.2. Results

A synopsis of the student responses to Likert items on agreement with key project goals is provided in the bar graphs in Figure 2. Students self-reported the tool was helpful in reducing anxiety, increasing engagement with the material, being relevant to their learning, and having a user-friendly interface.

Comparing student responses to the pre-song prompts versus some post-song assessments on the same topics showed some short-term gain in understanding after engaging with each of those activities. Analysis of student entries or selections in response to the prompts versus post-song assessments showed
Figure 2. Student feedback from pilot study at research and two-year institutions. Statements for each graph: (A) "The interactive song activity is a good way to get students engaged with learning statistical topics." (B) "The interactive song activities we did in class were relevant to my learning of statistical topics." (C) "The web-based interface of prompts and playback in the song activities was user friendly." (D) "The interactive song activity is a good tool in helping to relieve student anxiety about statistical topics." (E) "The songs used in the activities we did in class were high quality."
Figure 2. (Continued)
the method to be effective. For example, in a pre-song question associated with the song “Levels of Measurement,” students were asked to identify variable types for specific real examples and 34% of students were correct in doing so. In the follow-up assessment, students were asked to identify the level of measurement for a pet dog’s allergy rating (low, medium, or high), weight, body temperature (degrees Centigrade), and breed—82% of these same students answered all four correctly. Similar observational data on pre-post knowledge were also found for the other two songs used in the study.

The Likert scale data from the two institutions were reinforced by the positive student comments at both institutions that were also collected as part of the survey. The most common comments by far were some versions of “helpful in learning” and "easy to use interface,” while the two most common suggestions were to “make songs more fun (e.g., catchier beat)” and “improve sound quality” (e.g., the synthetic voice). These student comments as well as suggestions from workshop attendees, and the project Advisory Board led us to modify the SMILES collection to include songs with

- faster tempos (which is tricky with synthetic voice inserts),
- more contemporary genres,
- more varied singers (e.g., including a male vocalist on some songs),
- shorter length, or
- familiar melodies (e.g., a parody) to facilitate learning or recalling the song.

This feedback also led the SMILES team to identify and repurpose four previously recorded songs using public domain (e.g., imslp.org or pdinfo.com) permission-granted melodies for use in the SMILES project. Our Advisory Board and Evaluator also reviewed our processes and in some cases, individual songs. One advisory board member (while visiting the first author’s campus on August 31, 2016) tried out the interface for two songs and made useful comments about the insert prompts and the synthesized voice which helped inform aspects of the design of the interface and the planned research.

Along with the student feedback presented in Figure 2 on the in-class use of the SMILES website https://www.CAUSEweb.org/smiles/, we also examined the log files of the website that delivered the prompts during that same time frame in 2017. These data come from anonymous users so we are unable to identify the originating student or computer—though it is almost certain that all users in the window of time collected were at the research university when the class labs were meeting since the url of the project’s website had not been widely distributed up to that point.

As seen in Table 3, nearly all students completed the prompts when they worked in class and were given dedicated time facilitated by the instructor. In contrast, when these same songs were assigned as out-of-class and ungraded work only 41%–62% of students completed all questions. Also, for the in-class use, performance clearly improved in terms of the percentage of students who responded correctly to pre-song prompts on their first attempt compared to the percentage responding correctly to unpaired post-song assessments on the same learning objective done online within 36 hr. Note that such a large gain in the percent of correct responses may be partially explained by the fact that only the post-song quiz was graded.

The log files also allowed us to see where students got stuck and repeatedly answered the same question without success. Such information showed us where to improve question wording or add more hints and feedback.

### 7. Discussion

There are many concrete literature-based recommendations of practices to maximize student learning from educational videos (Brame 2016) that seem to readily transfer to songs, in the domains of: cognitive load, student engagement, and active learning. Such applicable suggestions include: keeping each song brief, using conversational language, having accompanying interactive questions, and having interactive features that give students control. Thus, we attempted to honor these aspects in developing the SMILES intervention.

As discussed in Lesser (2018a), songs-with-inputs could be used by students as a stepping stone to creating their own songs because they can take a lyric they already have familiarity with and connection to and then students "are given the freedom to alter any wording provided in the original format and often end up rewriting whole sections of the song” (Robb 1996, p. 32). If this is too big a jump for a student (e.g., in light of challenges previously discussed), an instructor can suggest a lyric (with manageable length and simple structure) to work with and/or some places in the lyric to change and then let students do the rest. Students can find elsewhere online a royalty-free vocals-free backing track where they can sing with the instrumental track and thus create an entirely new short jingle or rap from scratch and share the result.

#### 7.1. The Importance of Transferability

We are unaware of another attempt to create and assess interactive educational songs on this scale. We feel this has yielded a very comprehensive educational resource for introductory statistics or for teaching statistical concepts to students in other disciplines. It is no less important to see this as a proof of concept that can be readily transferred to the teaching of other disciplines in STEM and beyond. Beyond the general transferability of using song across STEM reflected by the VOICES initiative (https://www.CAUSEweb.org/voices/), there was a demonstration of specific STEM songs (in biology) being readily adaptable.

---

**Table 3.** Completion rates and unpaired assessment results for songs in spring 2017 in-class pilot study at a research university.

| Song                  | % of (n) students completing prompts | Pre-song % correct | Post-song % correct | Learning objective |
|-----------------------|-------------------------------------|--------------------|---------------------|--------------------|
| "Levels of Measurement" | 70                                  | 93%                | 34%                 | Identify data type in context |
| "Height of Confidence"  | 71                                  | 99%                | 40%                 | Effect of n on CI level on CI width |
| "Super Bowl Poll"       | 61                                  | 74%                | 15%                 | MOE varies with square root of n |

Note: MOE = Margin of Error; CI = Confidence Interval.
for the SMILES approach (Lesser et al. 2018). By sharing the resources, pitfalls, and insights we discovered along the way, we hope that those who do apply our methods to their field will be able to make more efficient timelines and protocols without reinventing the wheel. This article's documentation of process and trajectory is perhaps especially meaningful given the distinctive pioneering nature of this innovation, and also serves to honor recent calls for increased transparency in the research and creative process. This project also models how musicians in the academy can contribute to scientific fields and work with their less musically trained colleagues in other disciplines in a manner that complements well their traditional performing and teaching activities.

7.2. Future Directions

To improve completion rates for the SMILES prompts seen in the out-of-class environment, we are working to build up the number of hints and the quality of feedback to avoid students floundering at the task. In addition to extending transferability to other disciplines as discussed in Section 7.1, we hope to further personalize the interactive songs and to alleviate any remaining hesitations instructors might have about their use. For example, a vocoder (a portmanteau of voice encoder) may be used to allow students to speak their own inputs into their device's microphone, which are then put on a carrier wave to have them sung synthetically using the student's own vocal inflection and thus in the process give even nonmusical students more ownership. For now, we have added a karaoke button that allows students to play a song with its vocals removed and sing their own vocal live on top of the instrumental music. Research shows that hearing oneself provides a further learning benefit (Forrin and MacLeod 2018).

Secondly, we have started to add resources for instructors beyond the multiple-choice assessments and readings already produced, such as basic sheet music (for instructors who wish to organize an in-class performance). We also plan to incorporate experiences from the field and lesson plans. For some songs, Shiny apps from collections such as https://sites.psu.edu/shinyapps/ are being identified for use after the song activity to reinforce principles (e.g., “Height of Confidence” could be followed by https://psu-eberly.shinyapps.io/Inference_for_Proportions/). This is part of a more general conversation about what resources best support STEM songs (Crowther and Lesser 2019). Finally, while the process of creating (and doing preliminary assessment of) this extensive, multistage, interdisciplinary intervention has been a worthy thing in its own right, we are also hoping to gather data from rigorous randomized experiments to compare performance under varied treatment conditions (e.g., interactive song vs. a pre-composed song or no song) using a protocol that avoids the experimenter effect that may have occurred in Hardiman et al. (2019).
Appendix A: Screenshots of Pre-Song Prompts for Two Songs

https://www.CAUSEweb.org/smiles/songs/super_bowl_poll
https://www.CAUSEweb.org/smiles/songs/simpsons_paradox

Super Bowl Poll

1. The NFL is a professional American football league of 32 teams. Pick your favorite team — or just any team you think might have a chance to win the Super Bowl (championship game): Select.

2. The margin of error for a sample proportion for a survey of 1000 people would be about %. Hint

Here is a graph for how the margin of error changes with sample size and the population proportion p:

3. If 17% is the sample percentage, then the margin of error you entered in the above item gives an interval estimate as low as % and as high as %.

4. If you multiplied the sample size by a factor of nine, that would Select the margin of error.
No One Counted On Simpson's Paradox

Please fill in the blanks below with words that are as short as possible. Refer to these examples if you need help. Show examples: Example 1  Example 2  Example 3.

1. Give the name of a group that people could leave or join, comprised of two mutually exclusive types of people or individuals. 

2. Give a general label for any individual in your group (plural noun). 

3. Give a label (plural noun) for one type of individual in your group, ideally a type likely to score the higher average measure of your variable.

4. Give a label (plural noun) for a second type of individual in your group, ideally a type likely to score the lower average measure of your variable.

5. Name a dependent variable (plural noun) - some quality of the individuals in your group that you can measure and that might change over time.

6. Give a relative adjective (ends in "er") that describes your variable having a higher value.

7. Give a relative adjective (ends in "er") that describes your variable having a lower value. Should be the opposite of question 3.
### Appendix B: Learning Objectives of Songs

| Song title                          | Topic/learning objective                                                                 |
|------------------------------------|-----------------------------------------------------------------------------------------|
| A Fitting Conclusion               | Apply relationships among alpha level, p-value, and the decision of a hypothesis test. |
| A Radical Approach                 | Understand that standard error changes with the square root of the sample size.        |
| ANOVA                              | Recognize, in context, the idea of ANOVA as comparing variance between groups to variance within groups. |
| Central Limit Theorem              | Recognize when the Central Limit Theorem applies.                                       |
| Chi-Squared Dance                  | Recognize, in context, parts of a chi-squared test for independence (null, degrees of freedom, and observed vs. expected rationale). |
| Correlation Does Not Imply Causation| Recognize and construct examples to illustrate how correlation does not imply causation. |
| Correlation Illustration           | Interpret positive, negative, and zero correlation and transfer knowledge about associations to new context/situations. |
| Don't Buy a Carpet Estimation Notation: It's Greek (and Roman) to Me | Recognize how a change in units affects the variance (since variance is expressed in squared units). |
| Everything's Unusual               | Reason how larger sample sizes decrease the p-value, all else being equal.               |
| Height of Confidence               | Reason about the factors that affect the width of a confidence interval (sample size, confidence level, standard deviation). |
| Help Me Make a Good Bar Graph      | Recognize good properties of a bar graph.                                              |
| Hypothesis on Trial                | Identify counterparts in the courtroom analogy for hypothesis test (innocence = null, acquit = fail to reject, etc.). Identify errors of Type I and II in context. |
| Inferential and Descriptive Statistics| Contrast inferential and descriptive statistics with respect to their different goals and typical tools or outputs. |
| It Might Not Be That Bad           | Applying Bayes' rule or examining marginal and conditional proportions in a table to see how, for rare conditions, most positive test results may be false positives. |
| Levels of Measurement              | Give 4 levels of measurement (nominal, ordinal, interval, ratio scales) in appropriate hierarchical order and identify examples of each level in a real-world context. |
| My Family's Mean                   | Reason about mean and median, and the effect of an outlier.                             |
| No One Counted On Simpson's Paradox| Explore Simpson's paradox and recognize how a third variable might drive the relationship between two others. |
| Probability Rules Rap              | Recognize and apply the basic rules of probability and distinguish between their uses.   |
| Regression Rumba                   | Interpret basics of regression, including: checking assumptions, interpretation of slope, r, and r². |
| Super Bowl Poll                    | Apply margin of error in the context of a poll question, including that variability decreases with the square root of sample size. |
| Target Practice                    | Identify examples of biased (e.g., range) and unbiased (e.g., mean) estimators.         |
| The Enlightened Teacher            | Identify factors that allow a sample of data to be representative of the population and distinguish between random and convenience samples. |
| The Null Hypothesis                | Identify the null hypothesis (as opposed to the alternative hypothesis) for the most common hypothesis tests. Identify factors to consider when deciding how outliers should be treated, as well as factors for deciding if a study is worthwhile. |
| Throw That Out?                    | Understand the definition of a p-value.                                               |
| What P-Value Means                 | Reason about the meaning of and distinguishing among observed, fitted, and residual values. |

### Supplementary Materials

The supplemental materials consists of a 3-minute video about Project SMILES.

### Acknowledgments

We thank the editors and anonymous reviewers for their very helpful feedback. We are also grateful for the many contributions from other members of our overall project (see www.CAUSEweb.org/smiles/team). A special thanks to GSU’s J. Garrett DeHart, whose videography of this project helped us win the grand prize in the mid/advanced career category of the NSF’s 2019 We Are Mathematics Video Competition.

### Funding

This work is supported by the first three authors’ National Science Foundation DUE EAGER grants #1544426 (PSU), #154437 (UTEP), and #1544243 (GSU). Opinions, findings, and conclusions are those of the authors and not necessarily those of the NSF.

### ORCID

Lawrence M. Lesser [http://orcid.org/0000-0001-5762-3987](http://orcid.org/0000-0001-5762-3987)

Dennis K. Pearl [http://orcid.org/0000-0003-1981-1826](http://orcid.org/0000-0003-1981-1826)

Robert P. Carey [http://orcid.org/0000-0003-3794-3849](http://orcid.org/0000-0003-3794-3849)

### References

Aragon, R., and Wickramasinghe, I. P. (2016), "What Has an Impact on Grades? Instructor-Made Videos, Communication, and Timing in an Online Statistics Course," *Journal of Humanistic Mathematics*, 6, 84–95.

Baker, F. A. (2015), *Therapeutic Songwriting: Developments in Theory, Methods, and Practice*, New York: Palgrave Macmillan.

Berk, R. A. (2009), "Multimedia Teaching With Video Clips: TV, Movies, YouTube, and mtvU in the College Classroom," *International Journal of Technology in Teaching and Learning*, 5, 1–21.

Brame, C. J. (2016), "Effective Educational Videos: Principles and Guidelines for Maximizing Student Learning From Video Content," *Cell Biology Education—Life Sciences Education*, 15, 1–6.

Cavanagh, S. (2014), "What Is 'Personalized Learning'? Educators Seek Clarity," *Education Week*, available at [https://www.edweek.org/ew/articles/2014/10/22/09pl-overview.h34.html](https://www.edweek.org/ew/articles/2014/10/22/09pl-overview.h34.html).

Chew, P. K. H., and Dillon, D. B. (2014), "Statistics Anxiety Update: Refining the Construct and Recommendations for a New Research Agenda," *Perspectives on Psychological Science*, 9, 196–208.

Chickering, A. W., and Gamson, Z. F. (1987), "Seven Principles for Good Practice in Undergraduate Education," *American Association for Higher Education Bulletin*, 39, 3–7, available at [http://www.aahea.org/articles/sevenprinciples1987.htm](http://www.aahea.org/articles/sevenprinciples1987.htm).

Crowther, G. J. (2012), "The SingAboutScience.org Database: An Educational Resource for Instructors and Students," *Biochemistry and Molecular Biology Education*, 40, 19–22, DOI: 10.1002/bmb.20567.

Crowther, G. J., and Lesser, L. M. (2019), "Just Add Water? A Conversation About Resources to Support Songs," in *3rd Annual Conference of Virtual Ongoing Interdisciplinary Collaborations on Educating with...*
Lesser, L. M. (2018b), “Modulating Misconceptions With Musical Means,” *International Journal of Science Education*, 38, 73–95.

delMas, R. C. (2002), “Statistical Literacy, Reasoning, and Learning: A Commentary,” *Journal of Statistics Education*, 10, 1–11. DOI: 10.1080/10691898.2002.11910679.

Eldon, D., and Eldon, D. (1995), *Lyrical Life Science Workbook*, Corvallis, OR: Lyrical Learning.

Evans, G. F. (1986), “Getting Through Statistics With the Help of Metaphors,” *Journal of Education for Business*, 62, 28–30.

Fausto-Sterling, A. (2000), *The Five Sexes, Revisited*, *The Sciences*, 40, 19–23.

Ferrin, N. D., and MacLeod, C. M. (2018), “This Time It's Personal: The Memory Benefit of Hearing Oneself,” *Memory*, 26, 574–579.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafo, N., Jordt, H., and Wenderoth, M. P. (2014), “Active Learning Increases Student Performance in Science, Engineering, and Mathematics,” *Proceedings of the National Academy of Sciences of the United States of America*, 111, 8410–8415, available at http://www.pnas.org/content/111/23/8410.full.

GAISE College Report ASA Revision Committee (2016), *Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report 2016*, available at http://www.amstat.org/education/gaise/GAISECollege_Full.pdf.

Gellert, A. S., and Elbro, C. (2013), “Cloze Tests May Be Quick, But Are They Dirty? Development and Preliminary Validation of Cloze Test of Reading Comprehension,” *Journal of Psychoeducational Assessment*, 31, 16–28.

Goldman, B. A., and Kelley, M. R. (2009), “The Generation Effect in the Context of Lyrical Censorship,” *Psi Chi Journal of Undergraduate Research*, 14, 72–77.

Guo, P. J., Kim, J., and Robin, R. (2014), “How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos,” in *LIMA Proceedings of the First ACM Conference on Learning at Scale*, Association for Computing Machinery, New York, pp. 41–50.

Hardman, M. M., JohnBull, R. M., Carran, D. T., and Shelton, A. (2019), “The Effects of Arts-Integrated Instruction on Memory for Science Content,” *Trends in Neuroscience and Education*, 14, 25–32.

Hardin, J., and Miller, J. (2019), “The Evolution of Variables and the Existence of Trans People,” *Amstat News*, 501, 37–38.

Harper, M. (2018), “Prosody in Songwriting,” in *2nd Annual Conference of Virtual Ongoing Interdisciplinary Collaborations on Educating With Song (VOICES)*, available at https://www.CAUSEweb.org/voices/2018/session/13.

Jakubowski, K., Finkel, S., Stewart, L., and Millenstien, D. (2017), “Dissecting an Earworm: Melodic Features and Song Popularity Predict Involuntary Musical Imagery,” *Psychology of Aesthetics, Creativity, and the Arts*, 11, 122–135.

Lesser, L. M. (2000), “Sum of Songs: Making Mathematics Less Monotone!,” *Mathematics Teacher*, 93, 372–377.

Lesser, L. M. (2001), “Musical Means: Using Songs in Teaching Statistics,” *Teaching Statistics*, 23, 81–85.

Lesser, L. (2012), “High-Speed Hypothesizel,” *Teaching Statistics*, 34, 10–12.

Lesser, L. M. (2014), “Mathematical Lyrics: Noteworthy Endeavors in Education,” *Journal of Mathematics and the Arts*, 8, 46–53.

Lesser, L. M. (2017), “A Continuum of Interactivity With Educational (Mathematics/Statistics) Songs,” in *2nd Annual Conference of Virtual Ongoing Interdisciplinary Collaborations on Educating With Song (VOICES)*, available at https://www.CAUSEweb.org/voices/2017/panel/1-3.

Lesser, L. M. (2018a), “Student-Created Songs in Statistics Class,” videoposter, in *4th Electronic Conference on Teaching Statistics*, available at https://www.CAUSEweb.org/cause/ecots/ecots18/posters/4-04.

Lesser, L. M. (2018b), “Modulating Misconceptions With Musical Means,” *Teaching Statistics*, 40, 79–82.

Lesser, L. M., Pearl, D. K., and Weber, J. J. (2016), “Assessing Fun Items’ Effectiveness in Increasing Learning of College Introductory Statistics Students: Results of a Randomized Experiment,” *Journal of Statistics Education*, 24, 54–62.

Lesser, L., Pearl, D., Weber, J., and Crowther, G. (2018), “Writing or Adapting Songs for Student Inputs to Make Interactive STEM Songs,” in *2nd Annual Conference of Virtual Ongoing Interdisciplinary Collaborations on Educating With VOICES*.

Lesser, L., and Winsor, M. (2009), “English Language Learners in Introductory Statistics: Lessons Learned From an Exploratory Case Study of Two Pre-Service Teachers,” *Statistics Education Research Journal*, 8, 3–32.

Lummis, S. N., McCabe, J. A., Sickles, A. L., Byler, R. A., Hochberg, S. A., Eckart, S. E., and Kahler, C. E. (2017), “Lyrical Memory: Memsonic Effects of Music for Musicians and Nonmusicians,” *Psi Chi Journal of Psychological Research*, 22, 141–150.

Moore, D. S. (1997), *Statistics: Concepts and Controversies* (4th ed.), New York: W. H. Freeman.

Pogue, D. (2005), “Mad Libs, in Song,” *New York Times*, available at https://www.nytimes.com/2005/12/01/technology/poguesposts/mad-libs-in-song.html.

Qissasneue, M., and Parr, K. (2016), “Engage Me and I Learn: Interactivity Brings Content to Life,” in *SoTL Commons Conference, Georgia Southern University*, available at http://digitalcommons.georgiasouthern.edu/sotlcommons/SoTL/2016/38/.

Robb, S. L. (1996), “Techniques in Song Writing: Restoring Emotional and Physical Well Being in Adolescents Who Have Been Traumatically Injured,” *Music Therapy Perspectives*, 14, 30–37.

Rogness, N., Kaplan, J. J., and Fisher, D. (2016), “The HILT-LAS Project: High Impact, Little Time Activities That Address Lexical Ambiguity in Statistics,” in *Proceedings of AAAS/NSF Envisioning the Future of Undergraduate STEM Education: Research and Practice*, Washington, D.C., April 27–29.

Sabbag, A. G., Benovoy, M., Larcher, K., Dagher, A., and Zatorre, R. J. (2011), “Anatomically Distinct Dopamine Release During Anticipation and Experience of Peak Emotion to Music,” *Nature Neuroscience*, 14, 257–262.

Schmidt, J. A. (1983), *Fundamentals of Statistics: Informed Decisions Using Statistical Fundamentals*, Upper Saddle River, NJ: Pearson.

Sullivan, M. (2014), *Fundamentals of Statistics: Informed Decisions Using Statistical Fundamentals (4th ed.)*, Upper Saddle River, NJ: Pearson.

Wilson, M., and Gerber, L. E. (2008), “How Generational Theory Can Improve Teaching: Strategies for Working With the ‘Millennials,’” *Currents in Teaching and Learning*, 1, 29–44.