Student-Made Interactive Learning with Educational Songs (SMILES)

Dennis Pearl, Penn State University
Larry Lesser, The University of Texas at El Paso
John Weber, Perimeter College at Georgia State University

This Workshop

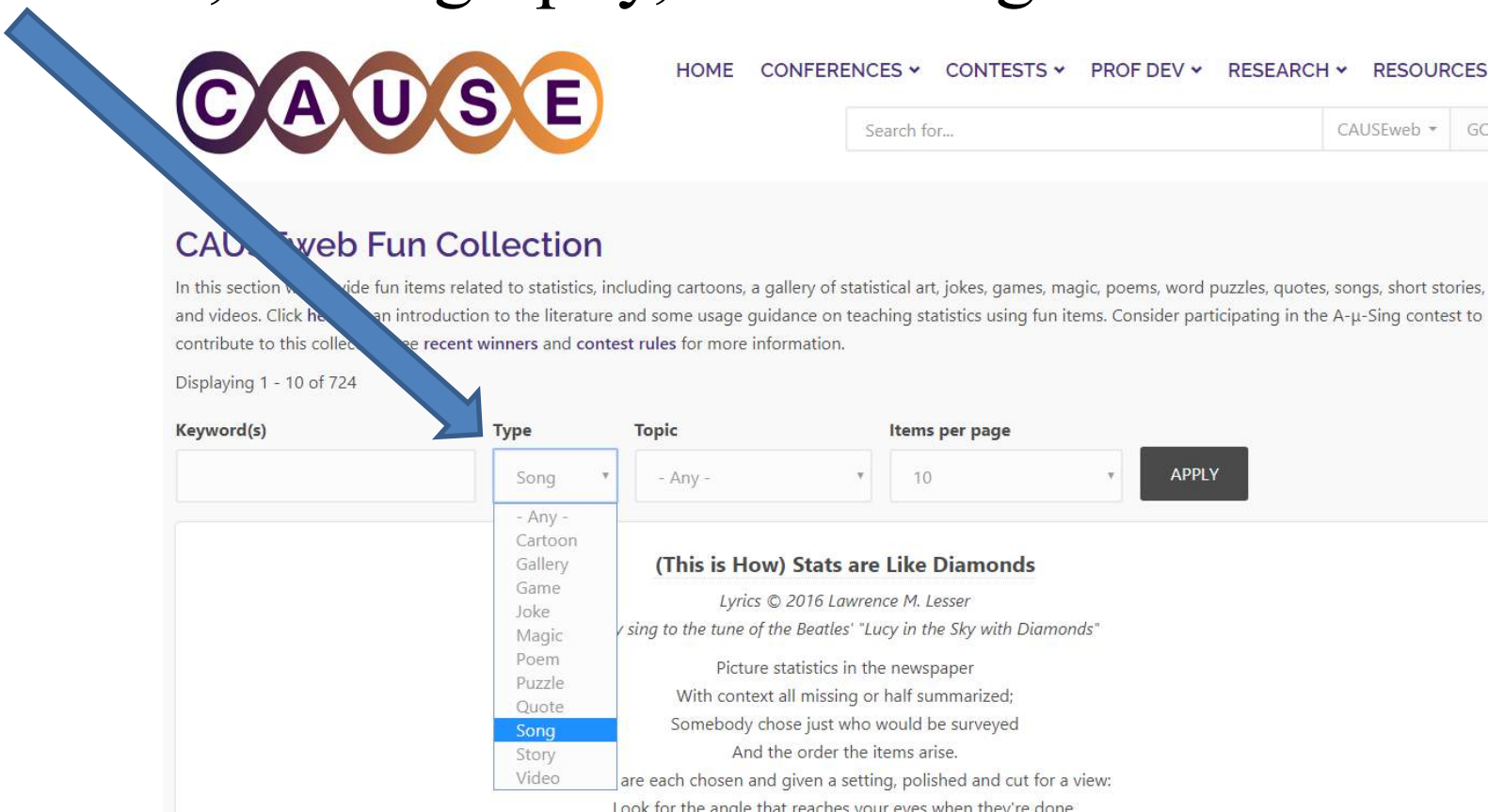
- Introductions
- Songs and other Fun for Teaching Statistics
- Why SMILES & the SMILES Library
- Your turn to play
- Critique and Discuss
- Results of Implementations
- The VOICES initiative

Resources & References we sent

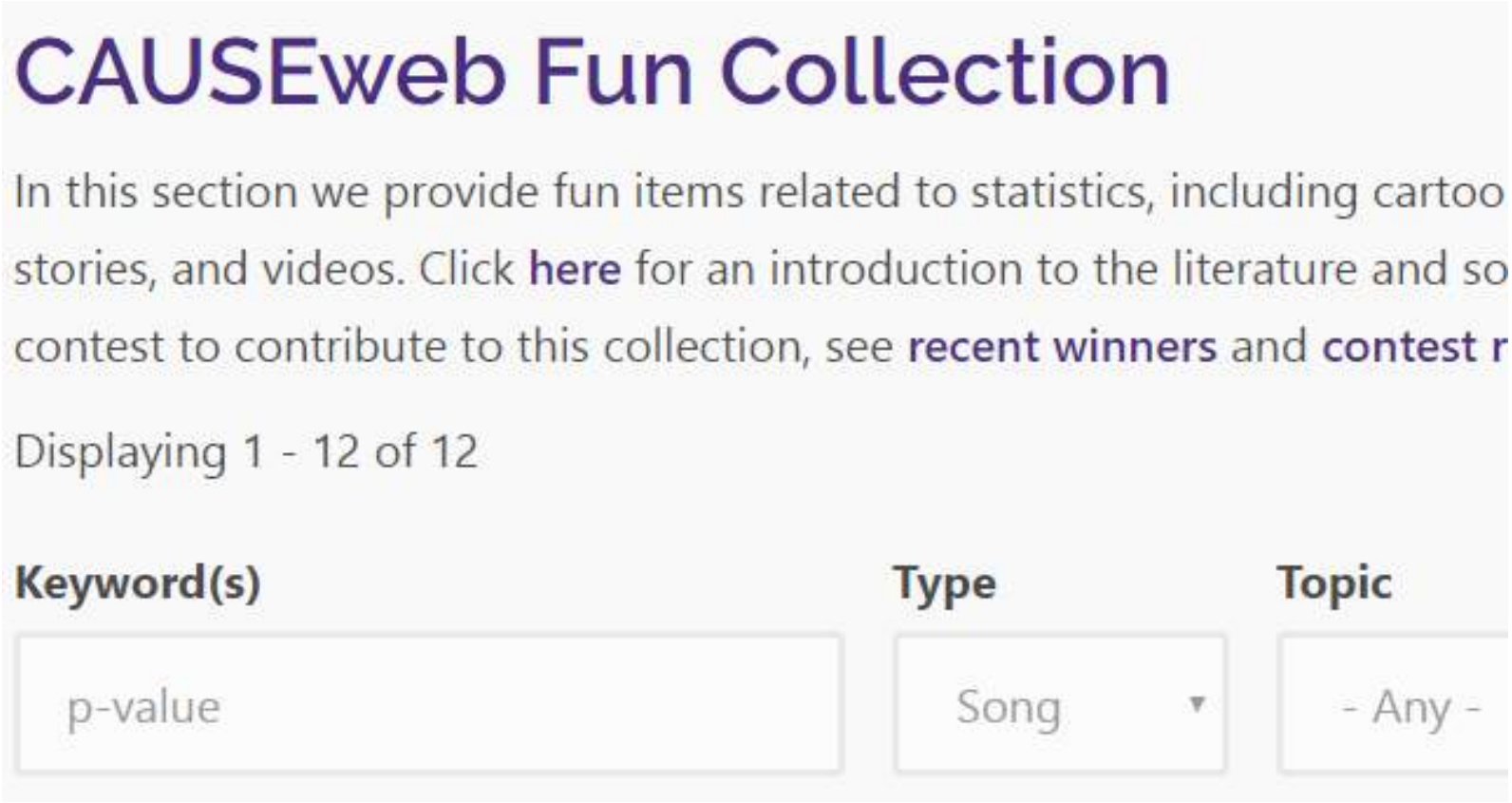
- www.CAUSEweb.org/resources/fun
- NSF Showcase video
- 2016 *JSE* paper on Effectiveness of Song www.tandfonline.com/doi/full/10.1080/10691898.2016.1190190
- 2017 VOICES presentation at www.causeweb.org/voices/2017/panel/1-3

causeweb.org

- site launched in 2005 (part of NSF's National Science Digital Library)
- Its curated, searchable 724-item fun collection includes 145 songs (almost all with soundfiles), 10 modalities, bibliography, & lesson guidance



Example: if you're teaching *p*-value



…and you choose this…

What P-Value Means

Lyric ©2005 Lawrence Mark Lesser;
May sing to tune of "Row, Row, Row Your Boat"

It is key to know
What p-value means --

It's the chance
(with the null)
you obtain
data that's

At least that extreme!

Download

Rating:

Average: 5 (1 vote)

…which has lesson guidance

Description (From the full CAUSEweb record description at: https://www.causeweb.org/cwis/r1248/song_what_p-value_means): Song is simply a quick jingle to help students recall the conceptual interpretation of a p-value. May be sung to tune of "Row, Row, Row Your Boat". Recorded June 26, 2009 at the OSU Whisper Room: Larry Lesser, vocals/guitar; Justin Slauson, engineer.

URL for lyric and soundfile: http://www.causeweb.org/resources/fun/db.php?id=86

Length of song: 10 seconds

Goal: helping students learn (and practice saying) the interpretation of a p-value

Target audience: students in any class that introduces p-values

Set-up: discuss a real-life vignette's probability such as "in a 10-child family, 9 babies were girls". Discuss what would be even more "extreme" (10 of 10 girls; if two-tailed, 9 or 10 boys as well) and unpack that there is an implicit null hypothesis that the probability of a birth being a girl is about 0.50, and that the 7 of 7 feels unusual because under the null hypothesis of independent births with P(girl) = 0.5 each time, it seems very unlikely to get 9 or more of the babies to be girls. Play the song.

In-class Use: Play the soundfile (http://www.causeweb.org/resources/fun/db.php?id=86) so students can hear the song (and read the lyric at the same time, making sure you select an updated browser that allows this). Then play it again and have the class sing along. For more adventurous classes, try singing it in a two-part round as "Row, Row, Row Your Boat" would be (i.e. after one half of the room finishes "It is key to know," the other half begins the song as done in the video at www.youtube.com/watch?v=CJrXuooX7hI.

Online self-paced use: instruct students to do a "set-up" reading, then click on the soundfile to play the song several times, then click on the "follow-up" reading

Follow-up: recap the pieces of the song to make sure students understand what is meant by "extreme" (try a different scenario to assess this and move away from the p=0.5 null hypothesis: a basketball player makes 7 of 10 free throws) and give examples that are one-tailed and examples that are two-tailed. Emphasize the conditional (i.e., "if the null hypothesis is true, then the probability of....") structure of the interpretation of a p-value.

Assessment: on the next midterm or quiz, try a relevant CAOS-pool item from the ARTIST database (https://apps3.cehd.umn.edu/artist/ or https://apps3.cehd.umn.edu/artist/tests/index.html); or here's a multiple-choice item adapted from Vogt, 2007 p. 13:

A p-value of .03 means:

1. There's a 3% chance the null hypothesis is wrong.
2. The probability that the result is due to chance (is a coincidence) is 3%.
3. A result of this size would occur by chance alone 3% of the time.
4. If the null hypothesis were true, the probability of getting a result at least this far away from the null hypothesis would be 3%.

STEM songs http://singaboutscience.org/

SEARCH OUR DATABASE OF 7000+ SONGS

Keyword/Phrase: Go!
Keyword/Phrase: Go!
Performer/Songwriter: Go!
Song Title: Go!
Song Template: Go!
(i.e., the original song upon which a parody is based)
Album Title: Go!
Website URL: Go!
(e.g., youtube.com [omit http://www.])

Additional (optional) search restrictions:

- [x] Only include songs intended for students between ages 0 and 99

ages" or adults will not be included if the range is narrowed to less than 0-

- [ ] Only include songs known to be parodies.
- [ ] Only include songs with online lyrics.
- [ ] Only include songs with free full-length online recordings.
- [ ] Only include songs with free full-length online videos.
- [ ] Don't include songs that may contain adult language.

Go! Clear all

randomized experiment (July 2016 *JSE*)

JOURNAL OF STATISTICS EDUCATION
2016, VOL. 24, NO. 2, 54–62
http://dx.doi.org/10.1080/10691898.2016.1190190

Assessing Fun Items' Effectiveness in Increasing Learning of College Introductory Statistics Students: Results of a Randomized Experiment

Lawrence M. Lesser[a], Dennis K. Pearl[b], and John J. Weber, III[c]

[a]Department of Mathematical Sciences, The University of Texas at El Paso, El Paso, Texas, USA; [b]Department of Statistics, The Pennsylvania State University, University Park, Pennsylvania, USA; [c]Department of Mathematics, Computer Sciences, and Engineering, Perimeter College at Georgia State University, Clarkston, Georgia, USA

ABSTRACT

There has been a recent emergence of scholarship on the use of fun in the college statistics classroom, with at least 20 modalities identified. While there have been randomized experiments that suggest that fun can enhance student achievement or attitudes in statistics, these studies have generally been limited to one particular fun modality or have not been limited to the discipline of statistics. To address the efficacy of fun items in teaching statistics, a student-randomized experiment was designed to assess how specific items of fun may cause changes in statistical anxiety and learning statistics content. This experiment was conducted at two institutions of higher education with different and diverse student populations. Findings include a significant increase in correct responses to questions among students who were assigned online content with a song insert compared with those assigned content alone.

KEY WORDS

CAUSEweb fun collection; Humor; Song; Statistics education research

diverse settings/populations for a statistical literacy course

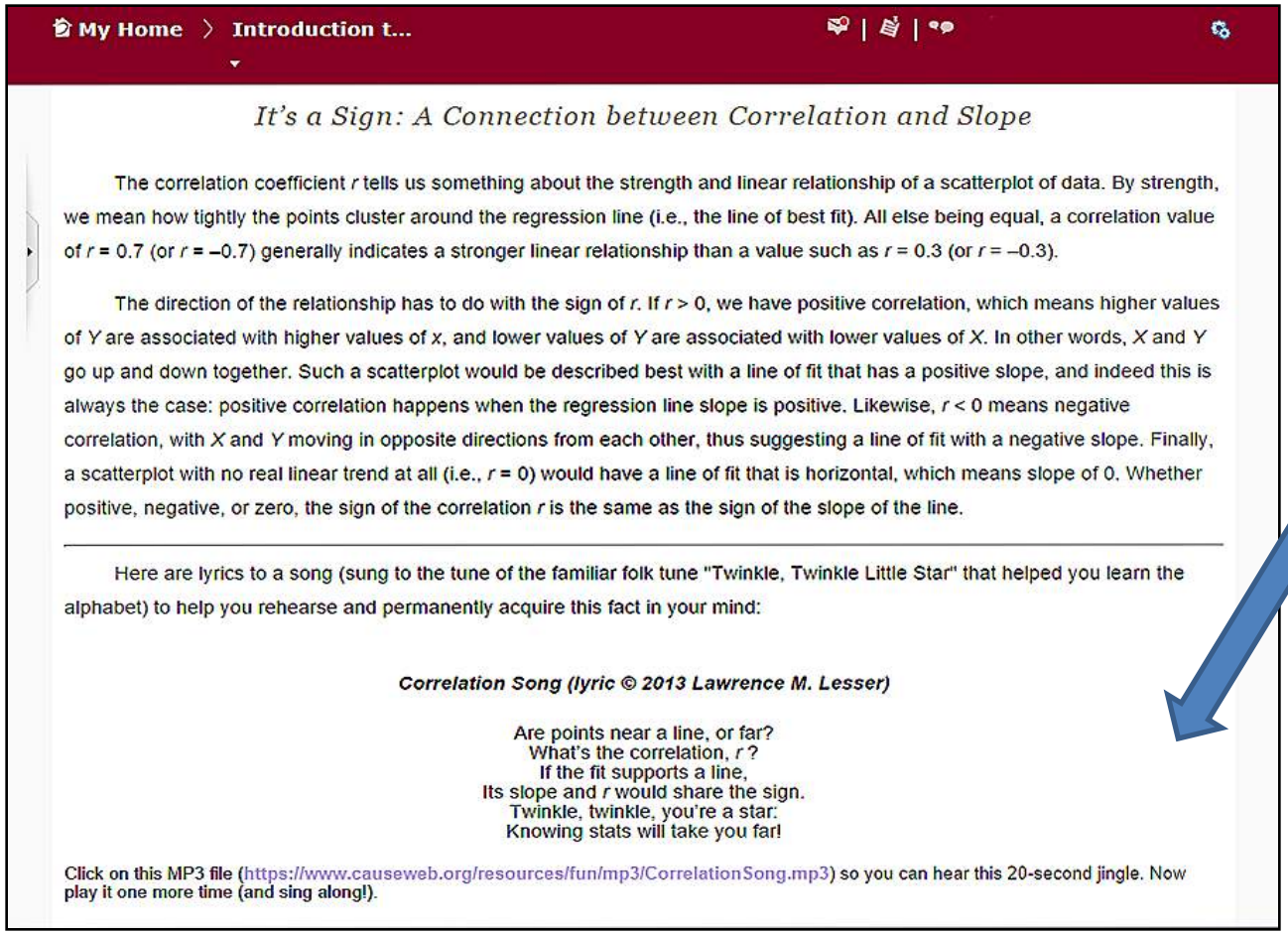
	Two-Year College	University (medium-size)
Region of U.S.	Southeast	Southwest
Student population	mostly Black	mostly Hispanic
Sample size (*n*)	53	194
Main audience	General education	Pre-service teachers
Text	Sullivan (2014) *Fundamentals of Statistics: Informed Decisions Using Data*	Utts (2005) *Seeing Through Statistics*
LMS (Learning Management System)	Desire2Learn	Blackboard

student-randomized experiment

- All students told their exams would have (12-14) embedded (MC) items related to online (LMS) content readings.

- Half the students randomized to always have “fun inserts” (song, cartoon, etc.) in those readings

mini-reading with insert



My Home > Introduction t...

It's a Sign: A Connection between Correlation and Slope

The correlation coefficient *r* tells us something about the strength and linear relationship of a scatterplot of data. By strength, we mean how tightly the points cluster around the regression line (i.e., the line of best fit). All else being equal, a correlation value of $r = 0.7$ (or $r = -0.7$) generally indicates a stronger linear relationship than a value such as $r = 0.3$ (or $r = -0.3$).

The direction of the relationship has to do with the sign of *r*. If $r > 0$, we have positive correlation, which means higher values of *Y* are associated with higher values of *x*, and lower values of *Y* are associated with lower values of *X*. In other words, *X* and *Y* go up and down together. Such a scatterplot would be described best with a line of fit that has a positive slope, and indeed this is always the case: positive correlation happens when the regression line slope is positive. Likewise, $r < 0$ means negative correlation, with *X* and *Y* moving in opposite directions from each other, thus suggesting a line of fit with a negative slope. Finally, a scatterplot with no real linear trend at all (i.e., $r = 0$) would have a line of fit that is horizontal, which means slope of 0. Whether positive, negative, or zero, the sign of the correlation *r* is the same as the sign of the slope of the line.

Here are lyrics to a song (sung to the tune of the familiar folk tune "Twinkle, Twinkle Little Star" that helped you learn the alphabet) to help you rehearse and permanently acquire this fact in your mind:

Correlation Song (lyric © 2013 Lawrence M. Lesser)

Are points near a line, or far?
What's the correlation, *r* ?
If the fit supports a line,
Its slope and *r* would share the sign.
Twinkle, twinkle, you're a star.
Knowing stats will take you far!

Click on this MP3 file (https://www.causeweb.org/resources/fun/mp3/CorrelationSong.mp3) so you can hear this 20-second jingle. Now play it one more time (and sing along!).

% Correct without and with Song Inserts

Topic	Without song *n* = 88	With Song *n* = 80	Difference
Margin of error: down with n down by √n	 57.3% 9.1%	 61.3% 10.0%	 4.0% 0.9%
Standard score	62.5%	75.0%	12.5%
Correlation & slope	60.2%	73.8%	13.6%
Equiprobability bias	40.9%	50.0%	9.1%
Multiplicity	36.1%	37.0%	0.9% (medium university)
p-value	44.4%	50.0%	5.6% (2-yr. college)
OVERALL	**42.3%**	**50.0%**	**7.7%**
			2-tailed *p*-value ≈ 0.04

we wondered……

why did songs have significant difference
but not cartoons?

and if it's because songs are more interactive,
how could songs get still *more* interactive?

Note: **Active learning** is a principle for good practice in undergraduate education (Chickering & Gamson, 1987); active learning in STEM is supported by meta-analysis of 225 studies (Freeman et al., 2014)

continuum of interactiveness of song
(see my 2017 VOICES talk:
https://www.causeweb.org/voices/2017/panel/1-3)

For example: having a student…

hear song <
provide inputs <
write song

Quick! Write down:

An exclamation

An adverb

A verb ending in ‘ing’

A plural noun

now read your “Mad Libs result”:

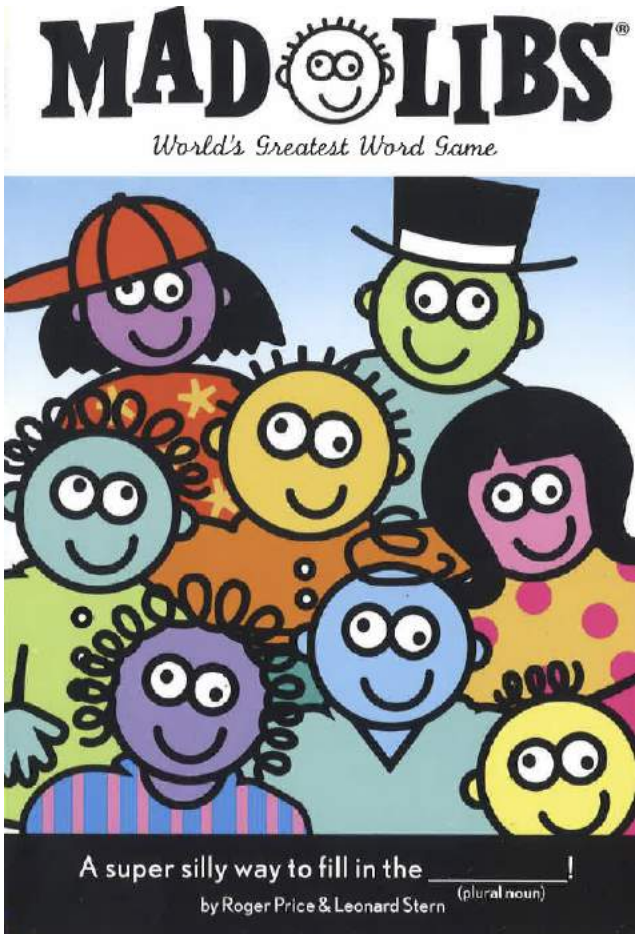
“_____!” she said _____. This talk is like_____ _____!”

[exclamation] *[adverb]* *[verb ending in ‘ing’]* *[plural noun]*

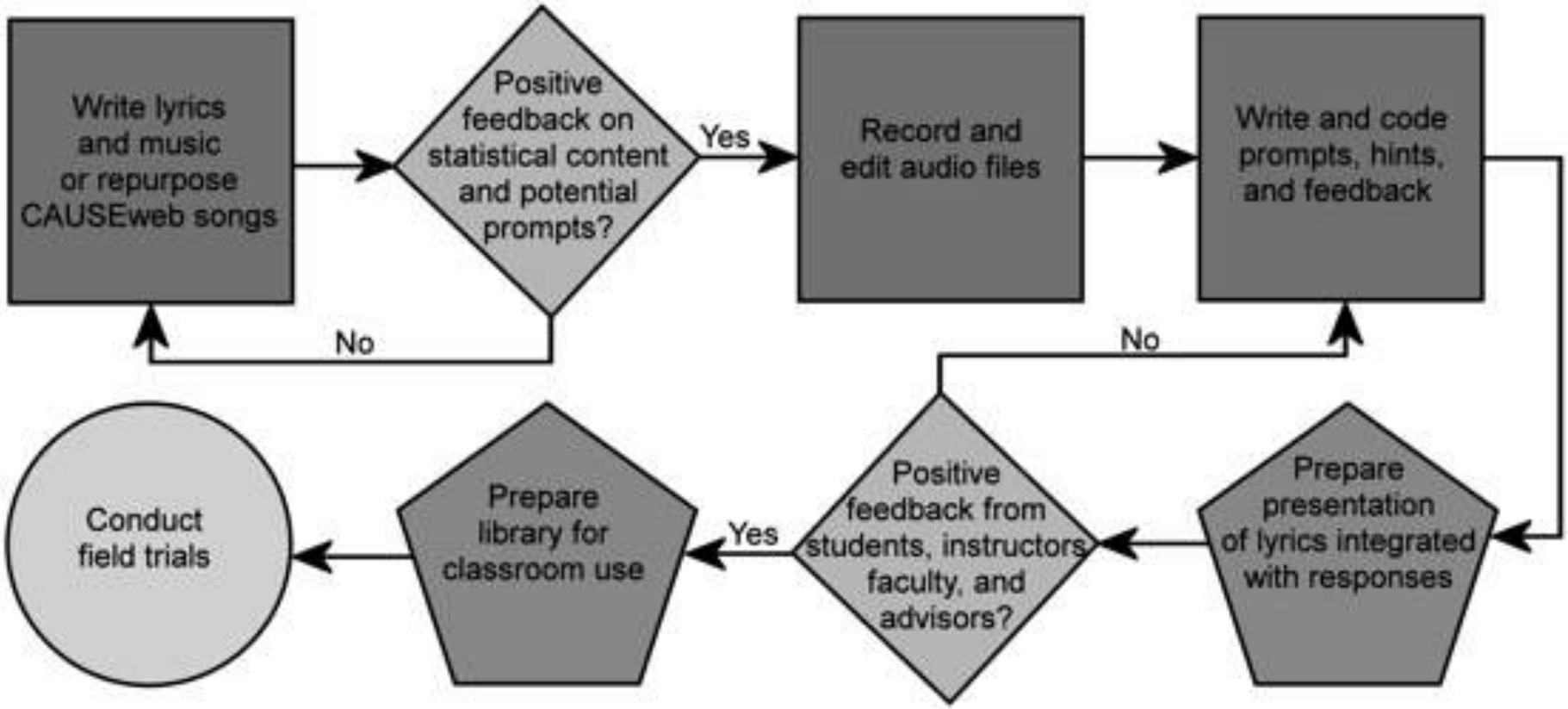
a Mad Libs approach…

has been adapted for educational use, including in teaching statistics (e.g., Trumpower, 2010)

A researcher uses a 2 *height (short, tall)* × 2 *relationship status (unmarried, married)* between-subjects, factorial design to study the effects on *maximum bench press ability* (as measured on a 10-point scale, with 10 indicating the greatest and 1 indicating the least *maximum bench press ability*). After collecting data, the researcher calculates the mean *maximum bench press ability* of participants in each condition of the study, as summarized in



process


Write lyrics and music or repurpose CAUSEweb songs
Positive feedback on statistical content and potential prompts?
Yes
No
Record and edit audio files
Write and code prompts, hints, and feedback
Prepare presentation of lyrics integrated with responses
Positive feedback from students, instructors, faculty, and advisors?
No
Yes
Prepare library for classroom use
Conduct field trials

www.causeweb.org/smiles/songs

Smiles Home About Library Our Team

Home / Song Library

Song Library

1. A Fitting Conclusion – 0:42

   Apply relationships among alpha level, *p*-value, and the decision of a hypothesis test.

   Build a Song | Studio Demo | View Reading

2. A Radical Approach – 0:35

   Understand that standard error changes with the square root of the sample size.

   Build a Song | Studio Demo | View Reading

3. ANOVA – 2:46

   Recognize the conceptual idea of ANOVA as comparing *within* to *between* variance.

   Build a Song | Studio Demo | View Reading

4. Central Limit Theorem – 1:12

   Recognize when the Central Limit Theorem applies.

   Build a Song | Studio Demo | View Reading

5. Chi-Squared Dance – 1:35

song topics

- Levels of measurement
- Mean vs. median
- Convenience vs. random sampling
- Correlation vs. causation
- Patterns of correlation
- Correlation and slope
- Statistic vs. parameter
- Estimator bias
- Margin of error in poll
- Probability rules
- Effects on width of CI
- Framework of testing H_o
- *p*-values
- Reporting test conclusion
- Effect of *n* on significance
- Concepts of regression model
- Observed/fitted/residuals
- Concepts of ANOVA test
- Variances (not SDs) add
- Bayesian reasoning (most tests for rare traits yield false positives)
- Central Limit Theorem
- Simpson's Paradox
- Ethics in statistics
- Inferential vs Descriptive

Pre-song prompts

- About 5 per song
- More than a quarter contain hints
- 43% involve free response with synthetic voice on playback, the rest are forced-choice answers highlighted on playback but sung with human voice
- 96% require statistical knowledge

Prompts vary in….

format

- Drop down from menu
- Drag-and-drop matching
- Fill-in (numerical)
- Fill-in (words)

Some not revealed before previous questions done, if that would “give away” an answer.

purpose

- Solicit context, example, or variable
- Apply procedure
- Make conceptual connection
- Connections across questions
- Playfulness

Some questions have > 1 reasonable answer

Throw That Out?

1. A possible **bad** reason for excluding a point from a data set is just to:

Select

- to fix a typo
- reach significance
- get a round n
- the measurement was done wrong
- remove an inconvenient value
- see no difference

2. excluding a point from a data set is becau

3. organization that establishes rules for the

prompts for “Super Bowl Poll”

Super Bowl Poll

1. Pick your favorite NFL team; if you don't have a favorite, just pick a team you think might be good: Select .

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

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** by a factor of _____ .

   - increase
   - decrease

Continue

Super Boll Poll

Will the **Texans** win next season's Super Bowl?

We asked about 1,000 fans in a scientific poll.

The margin of error was **3** %

That's roughly the reciprocal of the square root of *n*.

17% answered 'YES' in the poll,

But what could it be for the population as a whole?

At the 95% level of confidence

The interval goes from **14** to **20** %.

If we multiply the sample size by a factor of 9,

The new margin of error that we could find

Would be a ***third*** as large as what we had before.

Thanks to the formula, you know the score.

Go **Texans** !

The margin of error was 3 %

STUDIO VERSION

00:18

open-ended prompts solicit context

Regression Rumba

1. Insert the name of a visual graph you could view to decide if there is a linear relationship between the heights and weights of students in your class. Hint
2. For a relationship that interests you, insert the name of a quantitative variable (*Y*) that might play the role of a **response** (i.e., dependent) variable. Please do not use height or weight.
3. For a relationship that interests you, insert the name of the quantitative variable (*X*) that might play the role of an **explanatory** (i.e., independent) variable. Please do not use height or weight.

SONG UNDER CONSTRUCTION

Correlation Illustration Song

1. Think of two specific real-world variables that are quantitative (i.e., a variable whose values are numerical, not something like "eye color") and that should have a (strong or moderately strong) **positive correlation**.

   Give the name of one of the variables:
   Give the name of the other variable:

Correlation Does Not Imply Causation

For questions 1 – 3, consider this sentence:
"She likes to _____ in order to get _____."

1. For the first blank, please give a one-syllable **action verb** that is an activity someone does. Hint

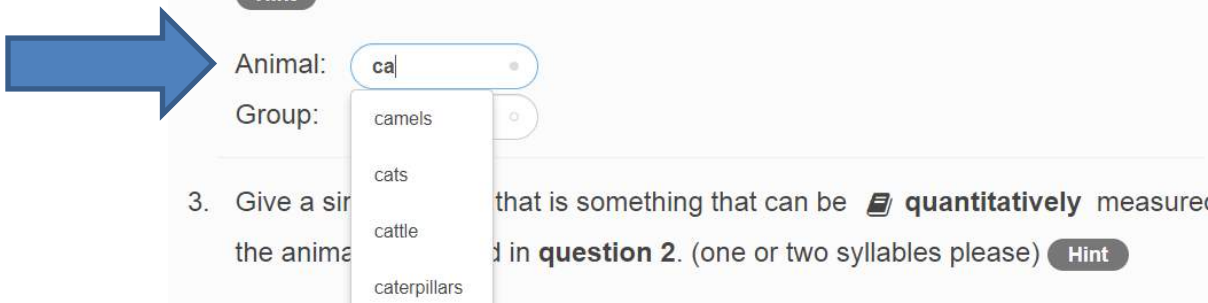
Simpson's Paradox

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

Checks on open-ended inputs

- Auto-corrects close spellings & grammar
- Allows British spelling
- Screens for profanity
- Checks if too many syllables
- Check for values out of range (e.g., $r > 1$) or inconsistent with other answer (sign of *r* & *b*)
- Accepts synonyms (scatterplot, scattergram, XY plot; normal, Gaussian, bell-shaped; bigger, larger, greater)
- Suggestions from first letters:

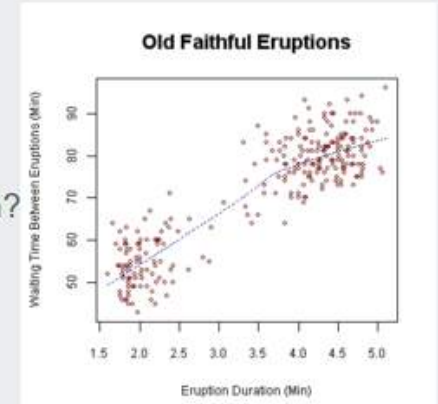


Some reasons for hints:

- not getting an answer could leave a student unduly “stuck” from continuing,
- a definition or symbol is used that a student might not know,
- academic wording of a question might not be clear to all,
- a word might be unfamiliar to someone new to the English language and/or American society,
- we want to teach the student something along the way by giving them a way to deduce the answer rather than repeatedly guess, or
- seeing example or visual may help understand a definition

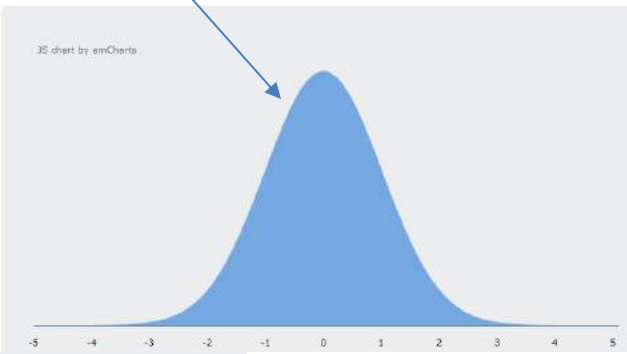
Insert the name of a visual graph you could view to decide if there is a linear relationship between the heights and weights of students in your class. Hint

What is a name for this type of graph?



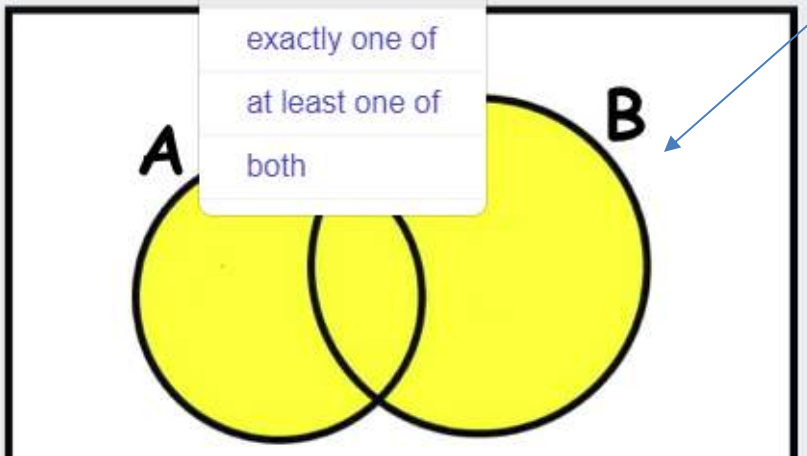
Central Limit Theorem

1. The sampling distribution of means or proportions for large random samples has what shape? Hint



The union of events A or B is often described using the words "A or B" (A ∪ B). This means that Select events A or B happened. Hint

- exactly one of
- at least one of
- both



For a given scatterplot the correlation coefficient and the slope of the regression line would Select share the same sign. Hint

always
sometimes
never

Try pl applet at http://illuminations.nctm.org/Activity.aspx?id=4186 with the "show line of bes ked and watch how r and the slope behave (the slope is the number that x is multip

If we think of the null hypothesis as the person (defendant) on trial, then "failing to reject the null hypothesis" would be analogous to a decision to Select the defendant on trial. Hint

To **convict** means to decide that there **is** a sufficient level of evidence that someone accused of a crime is guilty. To **acquit** means to decide that there **is NOT** a sufficient level of evidence that someone accused of a crime is guilty.

Three Songs for You to Try

- Central Limit Theorem
- Correlation Does not Imply Causation
- Super Bowl Poll

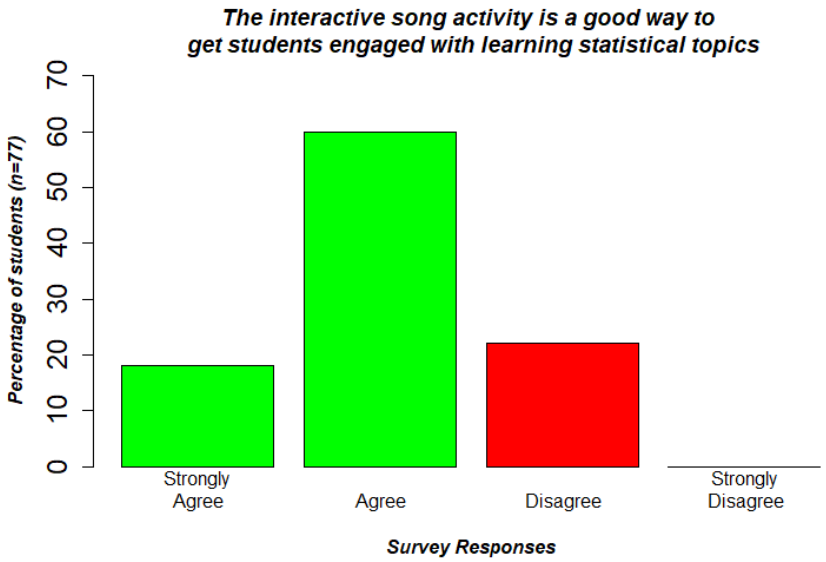
Critique and Discuss

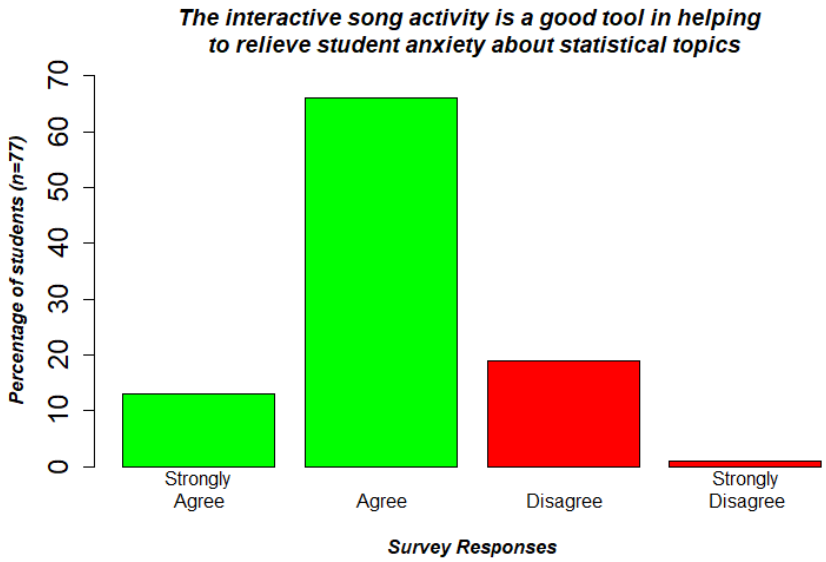
- Clarity of Interface (poll)
- Usefulness for Learning (poll)

- Strengths? Suggestions for Improvement?
- Pedagogical considerations
- Further comments

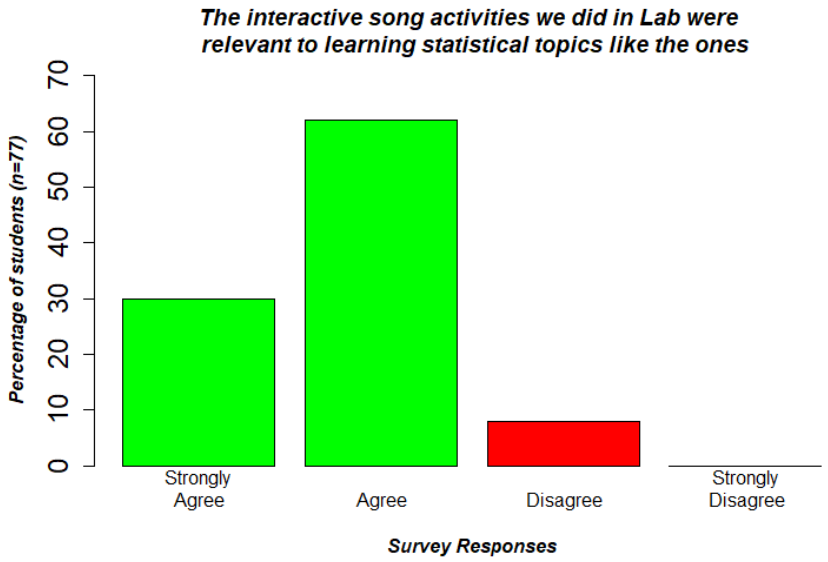
(Use chat window to enter your comments)

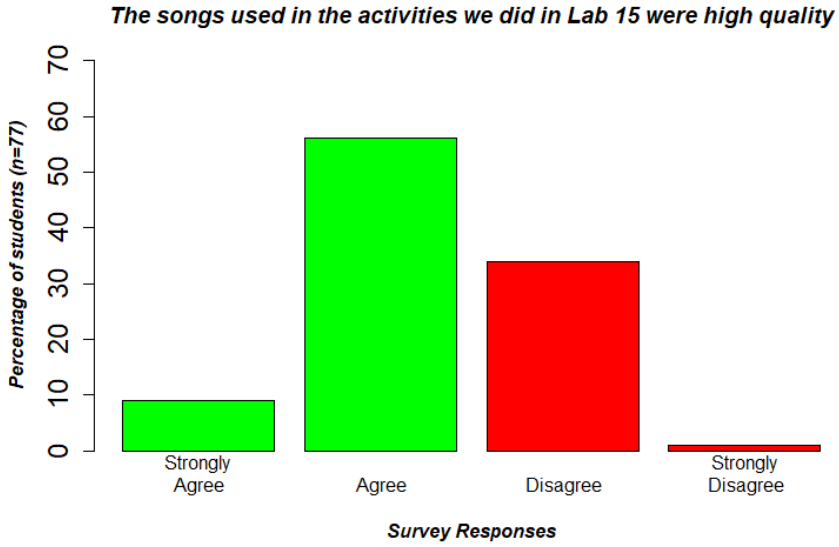
Student Reponses to Likert Items in Student Feedback Study



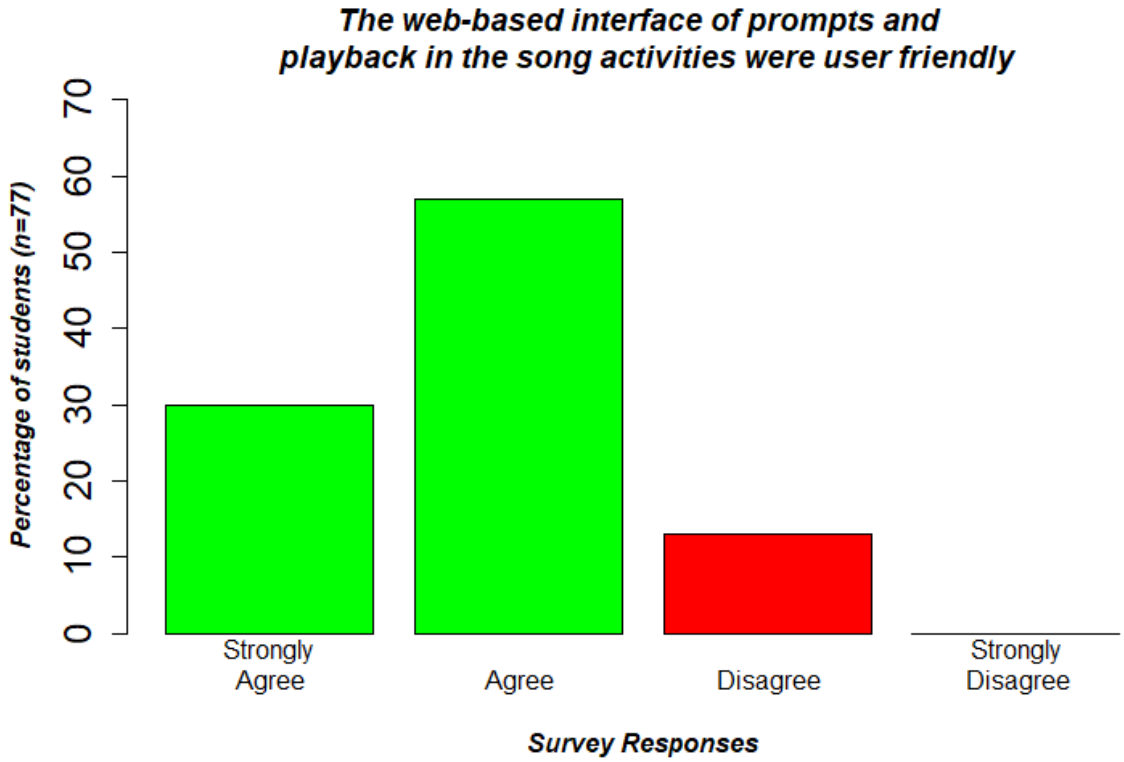


Student Reponses to Likert Items in Student Feedback Study





Student Reponses to Likert Items in Student Feedback Study



Pre/Post Knowledge

- Question 2 of **Levels of Measurement** asks students to find real examples of nominal/ordinal/interval/ratio
  34% got this right the first time.

  In the follow-up assessment, students were asked too identify level of measurement for allergy rating of pet dog (low, medium, or high); weight of pet dog; body temperature of pet dog (degrees Centigrade); and breed of pet dog

  82% got all four correct

Pre/Post Knowledge

- Questions 1 and 3 of **Height of Confidence** asked students what happens to CI with a larger n (question 1) and with a higher level of confidence (question 3)
  64% got question 1 correct on first try
  62% got question 3 correct on first try

  In the follow-up assessment, students were asked to order CIs from widest to narrowest amongst
  a. n = 500, with 80% confidence,
  b. n = 100, with 99% confidence,
  c. n = 500, with 95% confidence, and
  d. n = 100, with 95% confidence

  62% got this right (b then d then c then a)

Pre/Post Knowledge

- Question 4a and b of **Super Bowl Poll** asked students if n goes up by a factor of 9 then the MOE would go (up or down = question 4a) by a factor of (fill in 3)

  97% got 4a correct (MOE would decrease) but only 15% got the factor right on the first try. Note that this was after Height of Confidence that might have helped with 4a.

  In the follow-up assessment, students were asked to identify the factor by which MOE would change when sample size decreased by a factor of 4.
- 58% got this correct (bigger MOE by double)

and we're analyzing results of these student-randomized trials

	Fall 2017	Spring 2018
2-year college (mostly Black)	4 instructors, 12 sections	6 instructors, 15 sections
Research university		1 section (115 students)
Control Group	readings only	readings only
Treatment Group 1	SMILES platform, but not readings	SMILES platform, but not readings
Treatment Group 2		readings and (noninteractive) studio versions of songs

these ideas apply across STEM!

have STEM colleagues browse archived 2017 VOICES meeting and save the date for Sept. 26-27, 2018!

(STEM-focused, almost free, pedagogy/research/practitioner angles)
causeweb.org/voices/

Stay in Touch!

Contact us at smiles@causeweb.org

- The SMILES library is free to use (register to see assessments)
- Free registration to VOICES for SMILES workshop participants (email at address you used to register for eCOTS will come to you about this)