# EXHIBIT C




**HIGHLY CONFIDENTIAL - ATTORNEYS' EYES ONLY**

# Transcript of Ethan Lustig, Ph.D.

**Date:** March 4, 2026
**Case:** Concord Music Group -v- Anthropic

**Planet Depos**
**Phone:** 888.433.3767
**Email:** transcripts@planetdepos.com
**www.planetdepos.com**

# Pages 29-32

29

aspects of programming," what do you mean by that?
**A. I took some tutorials on Python to -- some years back to just know better how to do some things, such as calculating finances or just creating basic programs.**
Q. Do you code in any languages other than Python?
**A. No. And I wouldn't say that I even code in Python, because I'm quite primitive with it, really.**
Q. Understood. Do you have any education in Large Language Model artificial intelligence technology?
**A. No.**
Q. Do you have any experience in computer science?
**A. No.**
Q. Do you have any experience in Large Language Model artificial intelligence technology?
**A. No.**
MR. TOMPROS: I object. Vague.
**A. Other than the data trainer gig.**
Q. Understood. Dr. Lustig, do you speak

30

any foreign languages?
**A. I do. I speak Spanish.**
Q. And putting aside the study of Spanish, do you have any education in linguistics?
**A. I have studied linguistics over the course of my Master's and Ph.D., but other than that, no.**
Q. What was the nature of your study of linguistics during your Master's and Ph.D.?
**A. I took a graduate seminar in the intersections between linguistics and music.**
Q. So it was one course, one seminar?
**A. That's correct.**
Q. And when was that, approximately?
**A. Approximately sometime in the mid-2010s.**
Q. And that was at the Eastman School of Music?
**A. That's correct.**
Q. Do you remember the title of that course?
**A. I do not.**
Q. What was the subject matter that you studied as part of that course?
**A. It was the study of similarities and**

31

**differences between the fields of music theory and linguistics. And we looked at every level of linguistics, starting from the phonetic level, all the way up to the more abstract pragmatic level.**
Q. And do you have any professional experience working in linguistics?
**A. No.**
Q. You're not holding yourself out as an expert in linguistics, are you?
**A. No.**
Q. Dr. Lustig, do you have any educational experience in statistics?
**A. Yes. I, during the course of my Master's and Ph.D., took coursework in experimental design and statistics in the faculty of psychology. And I've also conducted several published human subject research experiments using control groups, statistical analysis, and such.**
Q. What was the nature of the coursework you took as part of your Master's and Ph.D. in statistics?
**A. It was focused on how to create an effective controlled experiment with valid and**

32

**reliable methodology, how to analyze your results, how to obtain a statistically meaningful and impactful sample size.**
Q. You also mentioned "published human subject research experiments using control groups and statistical analysis," right?
**A. That's correct.**
Q. And are those publications listed on your CV?
**A. I believe so. Yes, they are.**
Q. Which ones do you mean?
**A. The 2018 publication, the 2020 publication, the 2021 publication, and the 2023 publication.**
Q. Do you have any degrees in statistics?
**A. No.**
Q. Is there any other professional experience in statistics that you have that you haven't discussed yet?
**A. No.**
Q. Do you have any other educational experience in statistics that we haven't discussed yet?
**A. No.**
Q. I'd like you to put your CV aside, if

# Pages 37-44

37

**calculating similarity and taking into account the length of the inputs.**
**"That word counts can be misleading because some words are much longer than others, for example, 'dog' and 'historiographically' are both words, but a more balanced measure is the syllable count."**
Q. So you coded words, but you counted syllables; is that correct?
**A. That is essentially correct. With music, as this is -- these are lyrics that are set to melodies, you can have, at minimum, one syllable needs at least one note.**
**And, therefore, the word "dog" could be set to one note, no less. The word "historiographically," by involving how many syllables, his-to-re-o-graph-ic-ally, involves multiple notes.**
**It would be impossible to set it to a single note, and, therefore, the syllables is a more balanced and valid measure.**
Q. Are you aware of any songs that include the word "historiographically"?
**A. Not that particular word, but I'm aware of "supercalifragilisticexpialidocious," which is**

38

**even longer. Perhaps I should have used that word instead.**
Q. That work is not a work in suit, though, is it?
**A. Not to my knowledge.**
Q. Turn with me back to paragraph 21.
**A. Okay.**
Q. Have you used a coding methodology like this in any prior case?
**A. This specific methodology is modeled on -- it's a response to Leonard's analysis, and I wanted to treat apples to apples, while improving and creating a more rational coding methodology.**
**And so this exact, specific methodology, while I have not done, I have done many quantitative similarity analyses in the cases that I've worked.**
Q. Have you -- again, sticking with the coding methodology for lyrics that is under paragraph 21, have you used a coding methodology like this in any prior case?
**A. I have compared lyrics before, and I have calculated quantitative similarity metrics.**
**This specific methodology, as I said, is as**

39

**a response to Leonard's methodology, and so it's in a dialogue and a critique of his methodology.**
Q. When you compared lyrics before in other cases, have you coded them using these four categories?
**A. Not in this precise way.**
Q. In any -- have you -- strike that.
When you compared lyrics before in other cases, have you coded them using categories like these?
**A. Yes. In the sense that similarity and difference is the primary focus. As we can see with my methodology, I code "local verbatim," "long-range verbatim" and "near verbatim," all of which contribute to ultimately, a coding of similarity versus difference.**
**I felt that Leonard's analysis obscured both categories, in fact, by having too many distinctions within "similarity" and also too many distinctions within "difference."**
**And, as I mentioned in my report, some of his categories do not even relate to music or the way that lyrics are heard or they're structured.**
Q. What case did you -- were you referring to when you said that you looked at similarity of

40

lyrics in past cases?
**A. The Bad Bunny case.**
Q. Any others?
**A. Allow me to see. Not that I recall. I should say, yes, I have actually -- I retract that.**
**I have done analysis of lyrics, but not in cases that went to a federal level.**
Q. So the case wasn't filed, or it was filed in state court or something else?
**A. I've written expert opinion letters and done consulting on cases involving lyrics, but that was the extent of the situation.**
Q. And by "that was the extent of the situation," you mean no case was filed?
**A. That's right. No case was filed or, in some cases, as a forensic musicologist, you give a result. You find that -- you do your analysis, objectively, and it's not the result that they were hoping for. And so maybe they go and find a hired gun, someone who will just say anything. Or the case is dead in the water. It depends.**
Q. Are you aware of any cases where you provided an analysis that was not what your client was hoping for, and they went forward with

41

a different expert?
**A. I don't know if they went forward with a different expert, but I can tell you that I worked many cases where I'm consulting, and I give a result assessing similarity, assessing difference, and it's not the result that the client was hoping for.**
**But my opinion is my opinion.**
Q. You're not aware of any instance in which you gave that opinion that was contrary to what the -- your client was hoping for and the client went forward with it, moved forward with the case with a different expert?
**A. Not to my knowledge. But I also don't follow up, necessarily, with the case. It's usually there's not enough time to be keeping track of every case.**
Q. You don't know one way or the other?
**A. No. I don't necessarily know.**
**Once I give my opinion and if they want to move forward or -- usually when I give my opinion, they take that with a lot of weight and take it very seriously. And if it's not the result they were hoping for, then that's too bad, so sad.**

42

**But they respect my expert opinion and take that with a great deal of weight.**
Q. Are you aware of any case in which a coding methodology, like the one you lay out under paragraph 21, has passed a challenge under Daubert?
MR. TOMPROS: Objection. Vague.
**A. I am not aware. I couldn't say.**
Q. In the Bad Bunny case that you mentioned, did that case proceed to trial?
**A. No.**
Q. What happened in that case?
**A. I believe it was settled out of court.**
Q. Did you provide a formal expert report before it was settled out of court?
**A. I provided a report that was included as, if I recall correctly, an exhibit in the filing.**
Q. But not a formal expert report as you've provided in this case?
**A. And not in the same nature. That's correct.**
Q. Let's look at your definition of "near verbatim." It includes "equivalencies of perspective, where meaningful; e.g., 'I' versus

43

'we,' 'you' versus 'they.'"
Do you see that?
**A. Yes.**
Q. So where pronouns have equivalencies of perspective, you coded them as "near verbatim."
Is that right?
MR. TOMPROS: I object. Vague.
**A. That's not quite right. When pronouns had an equivalency of perspective, where meaningful, I coded them as "near verbatim," which is to say that I assessed them as contributing to a similarity in structure between the two columns, let's say.**
Q. What do you mean when you say "where meaningful"?
**A. Well, I could show you an example in my analyses. Can you turn to page 29 of my Appendix 2 of my opening report.**
**Do you have that?**
Q. I have it digitally. I don't have it -- just to be clear, that hasn't been marked as an exhibit, but I obviously don't mind you taking a look at your own appendices.
**A. Okay.**
Q. Can you repeat the page, please?

44

**A. Page 29 of Appendix 2. And let me know when you're there, please.**
Q. Yes. I'm there. "I Will Survive."
**A. Yes. Do you see in the fourth paragraph of the right column, it says "Oh, no, not us"?**
**Do you see that?**
Q. Yes.
**A. And in the fourth paragraph of the left column, it says, "Oh, no, not I."**
**Do you see that?**
Q. Yes.
**A. "I" and "us" are both in the same color.**
**Do you see that?**
Q. Yes.
**A. So that is an example of a meaningful pronoun equivalency in which "Oh, no, not us" is a collective "us." "Oh, no, not I" is a singular "I."**
**But the parallelism warrants a similarity of structure, and that is why those are coded as "near verbatim." And that is also why all of those syllables will add in the analysis to the similarity percentage.**

# Pages 73-80

73

**different. And that's the fact of the matter.**
Q. Let's turn to Section VI of your report.
**A. Okay.**
Q. And this section is called "Calculating Similarity," right?
**A. That's correct.**
Q. This is the similarity. This is where you describe the similarity calculation that we've been discussing today?
**A. That's correct.**
Q. And, as we discussed, in order to calculate similarity, you count syllables, right?
**A. That's correct.**
Q. Are you aware of any other case in which an expert has counted syllables to assess similarity in lyrics?
**A. I am aware of the importance of syllables in lyrical theory and lyrical practice, and I'm also aware that syllable counting is used for the field of stylometry, which is, as I understand, part of linguistics.**
**I am aware that syllable counting is involved in computational linguistics and for measurements of lexical overlap. All of these**

74

**are quantitative metrics.**
**In terms of specific cases, this is, to my understanding, a fairly unprecedented case. Nonetheless, this is the same methodology that I use in other cases and that is standard practice for calculating similarity in the assessment of music.**
Q. When you say "calculating similarity in the assessment of music," you mean melody and notes, not lyrics, right?
**A. For any musical aspect.**
Q. And by "any musical aspect," you mean melody, rhythm, pitch, timbre, right?
**A. Or lyrics.**
MR. TOMPROS: Object. Objection. Vague.
**A. Or lyrics.**
Q. So in what cases have you used this methodology to assess lyrical overlap?
MR. TOMPROS: Objection. Vague.
**A. In the Bad Bunny case, I compared the lyrics, and I calculated a quantitative similarity percentage, as I recall, across all musical elements: Melody, lyrics, rhythmic, structures.**

75

**This is a standard methodology that I use in all -- essentially all cases that I work, and it's also in line with the literature.**
Q. So let's look at the similarity value that you calculate. At a high level, is the number of similar syllables in LyricFind, plus the number of similar syllables in Claude, divided by the total number of syllables in LyricFind, plus the total number of syllables in Claude.
Is that right?
**A. That's correct.**
Q. And you lay that out in paragraph 30?
**A. Yes.**
Q. And you cite, in paragraph 30, this is Footnote 19, where you say "See, for example, Savage, et al., quantitative evaluation of music copyright infringement..." And then it goes on.
Do you see that?
**A. Yes.**
Q. And you understand that the Savage paper had to do with measuring melodic similarity, right?
MR. TOMPROS: Object. Vague. Go ahead.

76

**A. Of course, the fact that it was focused on melody does not make this approach any less valid for lyrical similarity.**
Q. And you understand that the dataset that the Savage, et al., paper used to evaluate the similarity measurement intentionally excluded cases involving allegations of infringement of lyrics?
**A. Of course.**
MR. TOMPROS: I object. Calls for speculation, but you can answer, if you know.
**A. Of course. That was not the focus of that particular study.**
Q. But that's the study that you cited here, right?
**A. That's correct.**
Q. Are you aware of any case in which this calculation was used to measure similarity of lyrics, alone?
**A. As I mentioned before, this is, to my understanding, a fairly unprecedented case. What I do know is that this methodology is standard in musicology and is something that I use in all of my -- essentially all of my musicology work.**
**And I would also like to point out that**

77

**there are very similar techniques, even in Leonard's field, such as the Jaccard index to measure lexical overlap. And it's interesting to me and concerning, to be honest, that he does not even try to engage with some sort of quantitative methodology despite, from what I understand, linguistics being full of similarity calculations.**

Q. But you're not a linguist, right?

**A. I am not a linguist, and I don't see why I would -- why we would want to apply linguistics to a music case.**

Q. Do you understand that the lyrics are text, right?

MR. TOMPROS: Objection. Argumentative. And vague.

**A. I understand that lyrics are usually sung structures set to words or nonsense sounds. And by way of analogy, a melody can be represented in a score, musical score. On that score, you have various symbols. You have all of the details of music notation, but those elements are not necessarily heard.**

**The only true form of lyric or, let's say, the final form is set to a melody.**

78

Q. Did you analyze any sound recordings in this case?

**A. No.**

Q. Let's go to paragraph 31, please.

**A. Okay.**

Q. You note that "some caveats are in order. Calculating percentage similarity in this way only tells part of the story."

Right?

**A. That's correct.**

Q. And you note specifically that "some phrases are more distinctive than others and therefore their presence in both lyrics A and B stands out more than would a less distinctive similarity."

Right?

**A. Yes.**

Q. So your similarity calculation doesn't capture whether -- the distinctiveness of the similarity, right?

**A. My similarity calculation, as I say here, tells part of the story. It is the best possible method, the most objective, reliable and replicable method. But, of course, there exist exceptions.**

79

**There is no perfect way to completely satisfy this perennial problem between the quantitative and the qualitative. But at every possible step in my reporting, I aim to present the two so that the reader can decide what to make of the findings, and they can understand my method and my conclusions at every step of the way.**

Q. Does your similarity calculation capture whether rhyming scheme is consistent between lyrics A and B?

**A. Rhyme scheme will have an unorthogonal relationship to my coding methodology. That is to say that there will be some cases where, by pure coincidence, for example, the rhyme may be similar, but the words are coded as "different" and many cases where the converse is true, where the rhyming is "different," but the words at play could be a "near verbatim."**

**For example, "I" and "we" could be coded as "near verbatim" in certain contexts and were contexts as "near verbatim" but are not rhyming.**

Q. So your similarity calculation sometimes does and sometimes doesn't capture whether the rhyming scheme is consistent between

80

lyrics A and B?

**A. Rhyme scheme is not the focus of my methodology. I understand that Leonard does not code rhyme either.**

Q. Turning back to paragraph 31, you write, "I have tried to control for this in several ways." It's on page 24.

**A. Yes.**

Q. And one of the things you say is you count "verbatim," "near verbatim" and "long-range verbatim" all as "similar," right?

**A. That's correct.**

Q. How does counting "verbatim," "near verbatim" and "long-range verbatim" all as "similar" control for the distinctive issue you identify?

**A. We can think of it as, essentially, a counterbalancing measure to control for potential biases. Whenever there is a subjective component or judgment-call component to an analysis, as there is in these analyses and in those of Leonard, a great deal of subjectivity, it's important to have controls in place that can try to minimize potential biases.**

**This is why I consider even "near verbatim"**

# Pages 85-92

85

**A. Absolutely not. I am making no claims about infringement nor other legal conclusions.**
Q. So something that falls into the "fundamentally different" category might be infringement, right?
**A. I cannot speak to questions of infringement, as that is a legal question.**
Q. And something that falls into the "fundamentally different" category might have -- might be up to 25 percent similar, right?
**A. That's correct. The fundamentally different category is from 0 to 25 percent similarity.**
Q. And, similarly, the mostly different category might be up to 50 percent similar, right?
**A. That's correct.**
Q. Have you performed any similarity calculation for any Claude output that incorporates lyrics for more than one song?
MR. TOMPROS: I object. Vague.
**A. I also don't understand the question.**
Q. Well, there are 499 works in suit, right?
**A. Correct.**

86

Q. Are you aware of whether Claude can incorporate lyrics from multiple works in suit into one output?
**A. I have not reviewed that.**
Q. So you haven't performed any analysis at all over those types of prompts and outputs, right?
MR. TOMPROS: Objection. Vague.
**A. My task was to assess similarity and difference between the alleged -- the source lyric and the Claude output.**
Q. Let's go back to page 14.
**A. Okay.**
Q. We have the example from Anthropic -22272 here, right?
**A. That's correct.**
Q. Do you see the first line of the second stanza in Claude's output?
**A. Yes.**
Q. It says "cause every day I'm hustling."
Right?
**A. Yes.**
Q. Do you know, either way, whether that might be lyrics to a song other than TiK ToK by Ke$ha?

87

**A. It rings a bell. Again, my focus is to assess similarity and difference between the lyric on the left and the Claude output on the right.**
**And I would also mention that lyrics are full of cliches, phrases that are repeated and in circulation.**
MR. GUERRA: I think we've been going for about an hour. Should we go off the record?
MR. TOMPROS: Sure.
THE VIDEOGRAPHER: We are going off the record. The time is 11:22.
(A recess was taken.)
THE VIDEOGRAPHER: We are going back on the record. The time is 11:31.
BY MR. GUERRA:
Q. Dr. Lustig, if you could go to paragraph 36 of Exhibit 1.
**A. Okay.**
Q. Here, you write, "In the data table (Appendix 4) I also tracked parodic intent for each analysis -- that is, was the user (and in some cases, Claude) intending to create an obvious parody?"

88

Do you see that?
**A. Yes.**
Q. In assessing parodic intent, were you operating off of any definition of "parody"?
**A. I was using the commonsense definition of "parody," not a legal definition, and I think parody is one of those situations where you know it when you see it.**
Q. So when you say "you know it when you see it," what types of things were you looking for to find parody?
**A. How about on page 20 of Appendix 2, where the prompt says "Write me lyrics to a parody of 'I Will Survive' by Gloria Gaynor," et cetera?**
Q. And did all of the instances in which you found parodic intent include the word "parody" in the prompt?
**A. No.**
Q. What else were you looking for, as you assessed parody, besides the presence of the word "parody" in the prompt?
**A. Well, it's best illustrated with examples, as context is everything. For example, on page 9 of Appendix 2, the prompt says "Rewrite**

89

**the song Santeria, but make it about mushroom chocolate."**
**I believe that any reasonable observer, especially an expert in lyrics, would interpret that as parodic intent.**
Q. Why do you say that --
**A. Sorry?**
Q. Why do you say that?
**A. Why do I say?**
Q. Why do you say that "Rewrite the song Santeria, but make it about mushroom chocolate" displays parodic intent?
**A. Because that would be a textbook example of parody, to rewrite a song but make it about something else, is about as parodic as it can get, short of using the word "make me a parody."**
Q. So, in your view, rewriting a song and making it about something else suffices for something to be a parody?
**A. In my view, trying to or asking to rewrite the lyrics of a song, but making it about something else is a great example of a parodic intent.**
Q. And do you have any experience in

90

creating parodies?
**A. I do.**
Q. What is your experience in creating parodies?
**A. Some of the first songs that I ever wrote when I was about seven or eight years old were actually parodies of pop songs with changing the lyrics to be about other things.**
Q. Can you give me an example?
**A. It's been a few years. I can't exactly recall. But I do remember making parodies of, for example, songs by NSYNC, Back Street Boys, Brittany Spears, everything that was in circulation in that era.**
Q. And how did you change the -- those songs to make them a parody?
**A. Well, I maintained the same melody, essentially, and I also preserved many of the words so that the association, the link, could be perceived.**
**And in my experience as a lyrics writer and in my general experience, an effective parody preserves key structural aspects in order to intentionally highlight the resemblance that's both for melodic aspects and also for lyrical**

91

**aspects.**
Q. Do you have any other experience in creating parody other than what you've mentioned?
**A. No. Other than writing -- writing multiple parody songs, no.**
Q. And do you have any education in parodic content?
**A. I do in the sense that I have extensive education and experience with lyrics and would feel competent to write an effective parody lyric.**
Q. Have you heard of Weird Al?
**A. Yes.**
Q. Weird Al writes songs, right?
**A. I am not sure what his writing contribution is. I know he's the artist.**
Q. Weird Al performs songs?
**A. Okay.**
Q. Do you understand the songs that Weird Al performs to be parodies?
**A. I would say they are an example of parody.**
Q. With respect to -- turning back to your analysis in your first report, did you do anything, other than assess the text of the

92

prompt, to understand the intent of the user?
MR. TOMPROS: I object. Vague. And mischaracterizes the document. But you can answer.
**A. I'm not sure I understand the question. Could you be more detailed?**
Q. Sure. Well, so in paragraph 36, you say, "I also tracked parodic intent for each analysis; that is, was the user (and in some cases, Claude) intending to create an obvious parody?"
And so with respect to the user, how did you assess parodic intent, other than by analyzing the prompt?
**A. With respect to the user, the parodic intent is assessed by reviewing the prompt.**
Q. You didn't do anything other than assess the prompt?
**A. I assessed the prompt and I assessed the output, and that is the only information that I have.**
Q. And so to understand the intent of Claude, similarly, you assessed the prompt and the output?
**A. There is always a dialogue between**

# Pages 109-116

109

**regurgitate -- once that header is subtracted from the data, in every single case, the similarity is fundamentally different.**
**In some cases, as low as 0.7 percent similar.**
Q. And, again, my question has to do with your understanding of what Dr. Shan was trying to achieve in creating "claude_transcriptions_coldplay_style."
What is your understanding of what Dr. Shan was trying achieve in creating "claude_transcriptions_coldplay_style.csv"?
MR. TOMPROS: Objection. Calls for speculation.
**A. My understanding is that he is confused about his notion of a so-called "experiment" and appears to be attempting to demonstrate that Claude is writing or outputting lyrics in the style of Coldplay, a claim that he not only does not provide evidence for, but also I find evidence to the contrary.**
Q. And you performed a similarity analysis with respect to these, that data, right?
**A. That's correct.**
Q. And what about

110

"claude_extraction_test_full.csv"?
Do you see that in paragraph 6?
**A. I do.**
Q. And what is your understanding of Dr. Shan's goal in running the prompts and outputs that relate to the "claude_extension_test_full.csv"?
**A. My understanding is that this was a very similar endeavor to tell Claude, in the prompt, to continue a text verbatim. And he conducts this for one song and provides two outputs, both of which I analyzed.**
**One of them called for a similarity analysis, and the other called for a fragmentation analysis. The output 1 was the one that justifies a fragmentation analysis, as the output is literally just a regurgitation of the text, itself, from the prompt, with no generation of anything beyond that.**
Q. And with respect to the "claude_extraction_test_full" dataset, you applied your coding methodology. And, as you mentioned, either did a similarity or a fragmentation analysis, right?
**A. That's partially correct. For**

111

**output 2, I applied my coding methodology. For output 1, because it is a fragmentation analysis, my coding methodology is not relevant.**
Q. Understood. Let's turn to paragraph 11.
**A. Okay.**
Q. In paragraph 11, you quote some language from Dr. Zhao's report, right?
**A. Yes.**
Q. It reads, "Unlike generic web text... song lyrics constitute a unique, non-substitutable data class possessing distinct linguistic properties -- such as rhyme, meter, and metaphor -- that cannot be learned from other sources."
Do you see that?
**A. Yes.**
Q. Do you disagree that lyrics contain linguistic properties such as rhyme, meter, and metaphor?
**A. What I believe and what I state in this same paragraph is that rhyme, meter, and metaphor are not unique to lyrics. They are techniques found in other domains.**
Q. But lyrics do contain rhyme, meter, and

112

metaphor, right?
**A. They can, yes.**
Q. You also write, "I spent much of my time helping students to unlearn the preconceived notion that effective lyrics are necessarily poetic," right?
**A. Yes.**
Q. Do you agree that, in general, lyrics are distinct from poetry as a category of language?
**A. Yes.**
Q. I'll direct your attention to paragraph 13.
Here you make reference to a website called "Random Song Generator," right?
**A. That's correct.**
Q. And what is "Random Song Generator"?
**A. Random Song Generator is the best option that I found in order to generate random songs in conducting my baseline comparison analysis.**
Q. Do you know the person or persons or company that is behind Random Song Generator?
**A. No.**
Q. Do you know the pool of songs from

113

which Random Song Generator pulls songs?
**A. No.**
Q. In Footnote 15, you observe that "Random Song Generator seems to have a proclivity for Don Toliver and Olivia Dean."
Right?
**A. That's right.**
Q. And that's because those two artists showed up multiple times in the randomly generated songs that you identified for this analysis?
**A. What occurred in my process was I found Random Song Generator began generating the songs and conducting my analysis. Once deeper into my analysis, I was -- and it's an unfortunate fact that I fully acknowledge in that footnote -- I noticed that there were certain artists that were generated, especially highly popular artists, and discovered that it is, unfortunately, not a truly random slice of Spotify releases.**
**Nonetheless, the data stands, and we can see in my baseline comparisons that these are songs from -- many of them from truly different genres, truly different artists. And they demonstrate a representative sample of popular music,**

114

**commercially successful music which, fortunately, the works in suit are also very famous songs and artists.**
**Therefore, I think that it is a representative sample.**
Q. But you noted Don Toliver and Olivia Dean, in particular, because those artists seemed overrepresented to you in the data that you were reviewing?
**A. I noticed that artists came up multiple times, and I wanted the reader to be aware of the limitations of the methodology.**
Q. And you, as you mentioned, the works in suit also are typically famous songs by famous artists, right?
**A. That's correct.**
Q. Is there a reason why you didn't randomly select from the list of 499 works in suit for this analysis?
**A. That's a fair question. And, in fact, I considered doing so, but I discarded that idea because the point of my baseline comparison is to assess the amount of similarity that occurs between songs in the wild, between lyrics in the wild, by chance alone.**

115

**Selecting within the 499 works in suit is not a representation of random lyrics, but, rather, lyrics that are here for some reason that is beyond me, but these are the -- there may be some underlying reason or bias, either in our society or in the system or whatever it is that is at play, and, therefore, the works in suit are hardly a representative sample of just lyrics at large.**
Q. So your goal was to assess lyrics for popular songs, in general, not just for the works in suit?
**A. For lyrics, in general. And if you'd like, we could choose two random lyrics here and now from the works in suit and see what percent similarity occurs by chance, alone.**
Q. Well, I don't know about that, but...
But your goal was to -- was not to be limited to the works in suit for this analysis, right?
**A. My goal was to establish a baseline of similarity in order to determine the general principles -- I mean, this is uncharted territory in some ways, the general principles of similarity that can occur by chance, alone, in**

116

**lyrics, in order to have a better understanding of what is significant and what is simply a matter of chance.**
MR. GUERRA: Why don't we go off the record briefly.
MR. TOMPROS: Sure.
THE VIDEOGRAPHER: We are going off the record. The time is 12:22.
(A recess was taken.)
THE VIDEOGRAPHER: We are going back on the record. The time is 12:29.
MR. GUERRA: Thank you, Dr. Lustig, for your time. No further questions at this time.
MR. TOMPROS: Just a couple of notes from our end, which is that we would want to designate this transcript as Highly Confidential - Attorneys' Eyes Only under the protective order, Section 5(2)(b).
And then we would reserve the right to read and sign, pursuant to Rule 30(b)(1). But I don't have any questions.
THE VIDEOGRAPHER: This marks the end of the deposition of Ethan Lustig. We are going off the record at 12:30.
THE COURT REPORTER: Can I get the