CRACKING THE CODE

Influencing Millennial Science Engagement

KQED

PRELIMINARY REPORT | DIGITAL VIDEO

A *Deep Look* at Gender Disparity

A Scientific Investigation into the Science Communication of KQED's Nature Series

REPORT AUTHOR

Dan M. Kahan, Yale Law School

November 20, 2019

FUNDING

National Science Foundation
DRL-1810990

IN PARTNERSHIP WITH

TEXAS TECH UNIVERSITY

Yale Law School

NSF

NOTE

This report describes the results of an ongoing research project, "Cracking the Code: Influencing Millennial Science Engagement (CTC)," aimed at promoting science-informed reporting of nature videos, science news and insights. Project sponsors include the National Science Foundation, the John Templeton Foundation, and the Annenberg Public Policy Center at the University of Pennsylvania. The report was prepared for internal use by the project team members, who include both professional science communicators affiliated with KQED, the NPR and PBS affiliate in San Francisco, California; and empirical researchers affiliated with Texas Tech University College of Media & Communication and the Cultural Cognition Project at Yale Law School. The report is being publicly disseminated not only to share knowledge generated by the team's initial research but also to sharpen apprehension of how collaborative exchange between researchers and practitioners can promote genuine evidence-based methods of science communication.

The research team members who conducted the study featured in this report include:

Meredith Berger, Cultural Cognition Project at Yale Law School

Josh Cassidy, KQED

Daniel Chapman, Cultural Cognition Project at Annenberg Public Policy Center

Sevda Eris, KQED

Kristina Janét, Texas Tech University College of Media & Communication

Dan M. Kahan, Cultural Cognition Project at Yale Law School

Asheley Landrum, Texas Tech University College of Media & Communication

Sue Ellen McCann, KQED

Matthew Motta, Cultural Cognition Project at Annenberg Public Policy Center

Jenny Oh, KQED

Gabriela Quirós, KQED

Othello Richards, Texas Tech University College of Media & Communication

Craig Rosa, KQED

Lauren Sommer, KQED

FOR MORE INFORMATION, please visit *KQED's Cracking the Code: Influencing Millennial Science Media Engagement* web page. We also welcome inquiries from researchers, science media professionals and related science-communication professionals, and other curious and reflective people. If you have questions and/or would like more details about this research please contact Sue Ellen McCann, KQED's co-principal investigator at smccann@kqed.org.

CONTENTS

FIGURES

REPORT

A. Overview and summary conclusions

This report summarizes the third in a series of studies conducted to examine how the methods of the emerging *science of science communication* can be integrated into the craft of science media creation. The specific focus of the study featured in this report is a pronounced gender disparity in the YouTube audience for *Deep Look*, a popular online series of short life-science documentaries produced by the public media outlet KQED in San Francisco.

Major findings include:

1. The *Deep Look* audience gender disparity can be reproduced experimentally (pp. 5-10). This finding implies, among other things, that the disparity is not a consequence of content-propagation dynamics distinctive of YouTube or related on-line and social media platforms.

2. The disparity occurs because high-science-curiosity women are less likely to choose to view certain *Deep Look* episodes than are high-science-curiosity men (p. 6). The disparity is not a natural or inevitable consequence of any difference in the satisfaction that men and women take, respectively, in being exposed to scientific insights into the workings of nature. Indeed, when high science-curiosity women view *Deep Look* episodes, they are just as *engaged* by them as high-science-curiosity men. Rather, some other influence thus appears to impede women from electing to view episodes that they would enjoy were they to watch them (p. 9).

3. The difference in viewing rates of men and women is concentrated in high-science-curiosity women with modest levels of science comprehension — and disappears among high-science-curiosity women with higher levels of science comprehension (pp. 10-12). This surprising finding suggests that some unobserved disposition associated with science comprehension inhibits women (but not men) from availing themselves of an opportunity to satisfy their interest in science by choosing to view certain *Deep Look* videos.

4. Aversions to the subject matter associated with disgust do not appear to be responsible for the *Deep Look* gender disparity (pp. 12-15). Study measures that predict aversion to disgusting stimuli, such as a spider eating a roly poly or lice found on human hair, were not correlated with viewing decisions for men or women. Self-reported disgust aversions did vary among men and women but not in patterns that corresponded to gender differentials in viewing decisions.

These findings warrant additional investigation with more finely calibrated measures to determine the influences that discourage women, particularly those high in science curiosity but modest in science comprehension, from viewing *Deep Look* videos that they would likely find engaging.

B. Study background and motivations

1. Applied Science of Science Communication

This report describes interim findings in an ongoing National Science Foundation project, *Cracking the Code: Influencing Millennial Science Engagement,* (CTC) in the practical application of the science of science communication.

The science of science communication is an emerging field, the object of which is to use empirical methods to understand and promote public comprehension, engagement with, and enjoyment of the fruits of scientific inquiry (Jamieson, Kahan & Scheufele 2017). The field has been characterized by the steady and rapid accumulation of findings, the most conspicuous of which relate to the processes that guide enlightened public decision-making, and the forces that disrupt those processes (Kahan 2015).

But the field has also been constrained by a widely remarked deficiency. Lab studies, the mainstay of the science of science communication to date, generate insight through forms of controlled experimentation that quiet the cacophony of real-world influences that disrupt reliable identification of cause and effect. But for precisely that reason, findings derived from the tamed conditions of the lab cannot be readily translated into the much more unruly conditions in which real-world communicators operate. Only *additional* empirical study, *conducted in the field*, can adjudicate among competing plausible conjectures on how to reproduce real-world experiences of the noisy, chaotic environment of the science communication strategies shown to work in pristine lab conditions.

This *operational validity deficit* (Kahan & Carpenter 2017; Schellenberger 1974) was a major theme of a recent U.S. National Academy of Sciences expert

consensus report (2017) on the status of the science of science communication. That report recommended practitioner-researcher collaborations as a remedy. In such collaborations, practitioners work with researchers to identify problems that are suited for empirical examination; to formulate plausible competing hypotheses on the nature of those problems and their potential solutions; and to craft study designs calculated to support valid, realistic inferences on the relative strength of those hypotheses.

The current study is the product of such a collaborative research project. Supported by the National Science Foundation, the project — *Cracking the Code: Influencing Millennial Science Engagement* — contemplates a set of progressive studies to be conducted by a team of science-communication researchers and science-communication professionals from the San Francisco public-media producer KQED. The unifying theme of the underlying research agenda is to enable KQED to anticipate and harness dynamics distinctive of emerging social-media channels of programming dissemination, particularly among Millennials, the age cohort that is expected to become increasingly important to public media in coming years.

2. *Deep Look's* gender disparity and the "missing audience hypothesis"

The specific focus of the current study was a viewership anomaly relating to the KQED series *Deep Look*. *Deep Look* consists of short (typically less than five-minute) features that use innovative methods to showcase surprising dynamics, often occurring at the microscopic or near-microscopic level, essential to enabling living creatures to thrive in their natural environments. The series is highly popular. Nevertheless, its viewership reflects a pronounced disparity in gender: males make up a substantially higher proportion of the audience for the vast majority of *Deep Look* episodes on YouTube, the primary distribution source for the series.

Why is both a matter of concern and perplexity? Although aimed primarily at Millennials (the age group thought most likely to be consuming science programming via social media), the series is otherwise intended to be equally appealing to all individuals interested in science and nature, regardless of their socially defining characteristics. The resulting failure of the program to attract as many female as male viewers is thus a matter of consternation for the show producers.

On the basis of their professional judgment and experience, the producers have formed multiple conjectures on the possible source of this disparity. One possibility, of course, is that the viewership imbalance reflects a more deep-seated societal disparity in interest among men and women in science and nature generally. But another possibility is that the specific content of the features, many of which profile insects, could trigger sensibilities of disgust or fear that (as a result of social conditioning) are disproportionately common among women. Still another possibility is that social-media propagation dynamics, particularly ones associated with YouTube's statistically driven search-result and recommendation algorithms, bias viewership of *Deep Look* episodes toward male viewers.

More generally, the puzzle over the *Deep Look* gender disparity can be viewed as an instance of a general debate among popular-science communicators — particularly ones whose work is distributed primarily through public broadcasting outlets — over demographic imbalances in the audience for their programming (Kahan, Carpenter & Landrum 2015). The *natural audience hypothesis* attributes such imbalances — ones relating, typically, to age, religion, and politics, as well as gender — to the uneven distribution of interest in science programming within the population at large. In contrast, the *missing audience hypothesis* posits some hidden influence that discourages members of discrete, socially identifiable groups from engaging with high-quality science programming that would in fact satisfy their desire to know what science knows.

The principal aim of the current study is to generate information that would help KQED to weigh the relative plausibility of its producers' conjectures on the source of the gender disparity relating to *Deep Look*. But in so doing, the study is also expected to help adjudicate the dispute between proponents of the natural-audience and missing-audience hypotheses.

C. Study design, objectives, and creation

1. Generally

The study used an experimental design. In it, study subjects — consisting of a large (N = 2500), nationally representative sample of U.S. adults — were assigned to four conditions, each of which presented its members with the option to watch a different *Deep Look* episode.[1] The episode was described as "a video

[1] Subjects were recruited by the survey firm YouGov, which also administered the study online.

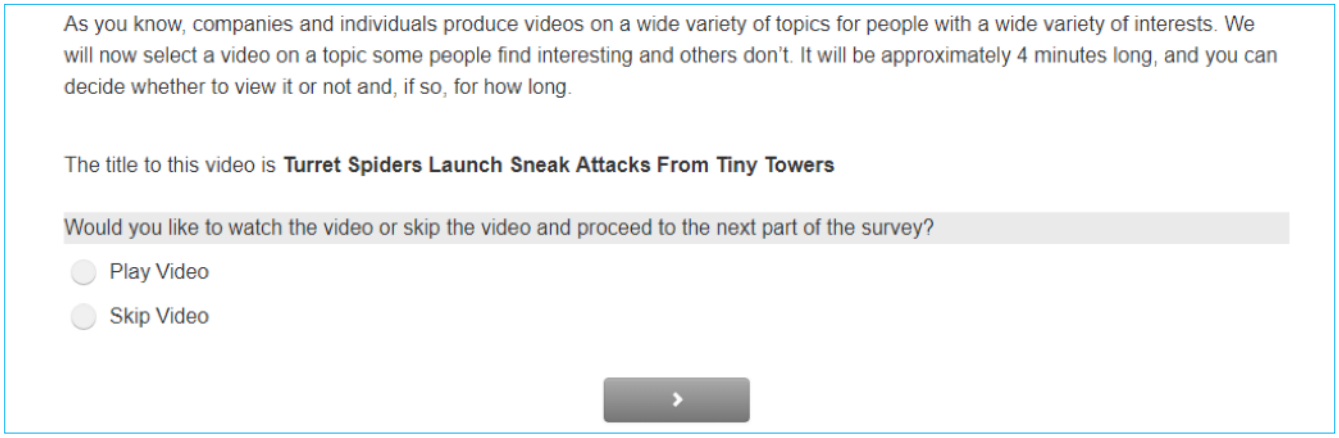


Figure 1. Decision-to-view option. Subjects elected to view the assigned Deep Look episode in response to this study prompt.

feature" of "approximately 4 minutes" in length, "on a topic some people find interesting and others don't." Subjects were presented with the option "to view it or not" and to view as much or as little of the feature as they chose (Figure 1).[2]

Identified by title, the four episodes were:

(1) "These Hairworms Eat a Cricket Alive and Control Its Mind" ("Hairworms");

(2) "How Your Dog's Nose Knows So Much" ("Dog Noses");

(3) "How Lice Turn Your Hair Into Their Jungle Gym" ("Lice"); and

(4) "Turret Spiders Launch Sneak Attacks From Tiny Towers" ("Spiders").

Spiders (86% male, 14% female) and Lice (52% male, 48% female) were included because they were the episodes that, in the period relatively close to the study, displayed the largest and smallest audience gender disparities, respectively. Hairworms (4.6 million YouTube views) and Dog Noses (227,000 views) were selected because they were the episodes that (in the relevant time frame) were the most and least popular *Deep Look* episodes overall.

The opportunity to view the assigned episode occurred about 15 minutes into what the subjects would have experienced as a consumer-marketing survey. In fact, embedded in the survey were items forming a standardized *Science Curiosity Scale* (SCS). Consisting of self-report and behavioral measures, SCS has been shown to furnish a highly discriminating measure of the propensity of individuals to voluntarily consume science-related material for personal satisfaction (Kahan *et al.* 2017). Embedding the SCS items in a larger collection of ones on a variety of other topics disguises the nature of the instrument as a measure of science curiosity, a feature essential to avoid triggering the incentive of subjects to overstate their interest in science.

The study instrument also measured the engagement level of subjects who chose to view the assigned *Deep Look* episode. This state was measured in part with items that solicited subjects' subjective appraisals (e.g., intensity of interest and surprise). But the study also employed behavioral indicators of engagement as well. These consisted of the subjects' decision to supply (or not) information necessary to receive alerts of future episodes of *Deep Look* and to "share" the viewed episode with others via email. Even more important, the study timed how long subjects viewed the assigned video. These various measures were appropriately combined to form an "engagement" scale.[3]

2. Individual characteristics

The study also collected a variety of additional information relating to the subjects' individual characteristics. These included standard demographics, including subject ages and genders. The study also included measures of the subjects' political outlooks and their level of religiosity. The subjects also completed a "risk perception" battery. A measure found to be valuable

[2] The study instrument is reproduced in the report Appendix.

[3] The scale was based on an Item Response Theory graded-response model. The Cronbach's for the 10-item measure was 0.79; the variable item-response reliability coefficient exceeded 0.70 across nearly the entire range of the measured latent variable (from the 5th to 95th percentiles). The scale thus furnished a highly reliable and discriminating measure of subject engagement with the featured *Deep Look* episodes (see generally DeMars 2010). Additional information on the scale formation and properties are available upon request.

in profiling subjects' outlooks and attitudes in a variety of other contexts (Cultural Cognition Project 2014), the battery directs subjects to indicate on an 8-point scale "how much risk" they believe each of a variety of identified behaviors or states of affairs "poses to human health, safety, or prosperity" (Kahan 2015). Finally, the subjects' level of science comprehension was measured with a standardized assessment, the Ordinary Science Intelligence ("OSI") test (Kahan 2017).

3. Hypotheses and inference strategy

It was anticipated that we would be able to use the study data to probe a variety of hypotheses relevant to the *Deep Look* gender disparity. As already adverted to, the **natural audience hypothesis**, for example, implies that the gender disparity in viewership is attributable to gender-based differences in interest in science. If this is correct, then one would expect any gender-based difference in the decision to view and be engaged by the various *Deep Look* features to be explained by the difference (if any) in the science curiosity levels of men and women generally.

The **missing audience hypothesis**, in contrast, asserts that there is some unaccounted for influence that suppresses enjoyment of certain forms of popular science communication among particular socially identifiable groups. The *Deep Look* gender disparity would fit this pattern if a gender difference in viewership or engagement persists even among women and men of *equal* levels of science curiosity.

In the event of such a finding, study data were geared to enable probing — if not definitive testing — of what the unobserved influence that accounts for a "missing audience" effect might be. The risk-perception, science-comprehension, religion, and political affiliation covariates, for example, can all be used to explore relevant surmises.

Alternatively, we anticipated that we might observe no discernible gender differences in the study subjects' decisions to view and be engaged by the study's *Deep Look* features. Such a finding would be consistent with the suspicion that the *Deep Look* gender disparity observed in the real world is being driven by social-media or platform-based mechanisms of content propagation, a surmise that we will call the **"algorithm hypothesis."**

4. Study generation

As indicated, a principal motivation for the research program of which this study is a part, is to remedy an operational-validity deficit (p. 2) in the science of science communication. Doing so demands collaborative investigations by social science researchers and professional science communicators (National Academy of Sciences 2017). The former are situated to contribute their familiarity with existing science-of-science communication research and their facility with the field's methods. But only the latter, equipped as they are with experienced-formed judgment, can identify consequential hypotheses and can vouch for the power of study designs to generate forms of information that practitioners can validly rely on in assessing such conjectures (Kahan & Carpenter 2017; Kahan 2014).

The *process* by which the current study was designed was vitally informed by this understanding. It was initiated by a series of face-to-face meetings designed to immerse the study team's empirical researchers in the outlooks and concerns of the team's professional communicators. On the basis of these meetings, hypotheses and the forms of evidence sufficient to help test them were jointly formulated. The researchers then devised a draft study instrument suited, in their view, for collecting such evidence, and shared it with the professional communicators, whose feedback was thereafter used to revise the instrument — a process that was repeated multiple times. The resulting instrument, in the team members' views, reflected not only the sum but the mutually educated synthesis of their respective forms of expertise.

D. Study results

The study generated a number of principal findings pertinent to the central study hypotheses, as well as a set of secondary findings of interest. We summarize the two sets of findings separately.

1. Principal findings

a. Experimental reproduction of the *Deep Look* gender disparity. SCS has proven a powerful predictive tool for identifying individuals interested in consuming science-related content for pleasure (Kahan et al. 2017). Not surprisingly, it was an extremely strong predictor of who watched the *Deep Look* episodes featured in this study (Figure 2). Averaged across all the episodes, the experimental data suggested that a modestly science curious individual (someone one standard deviation above the population mean — or

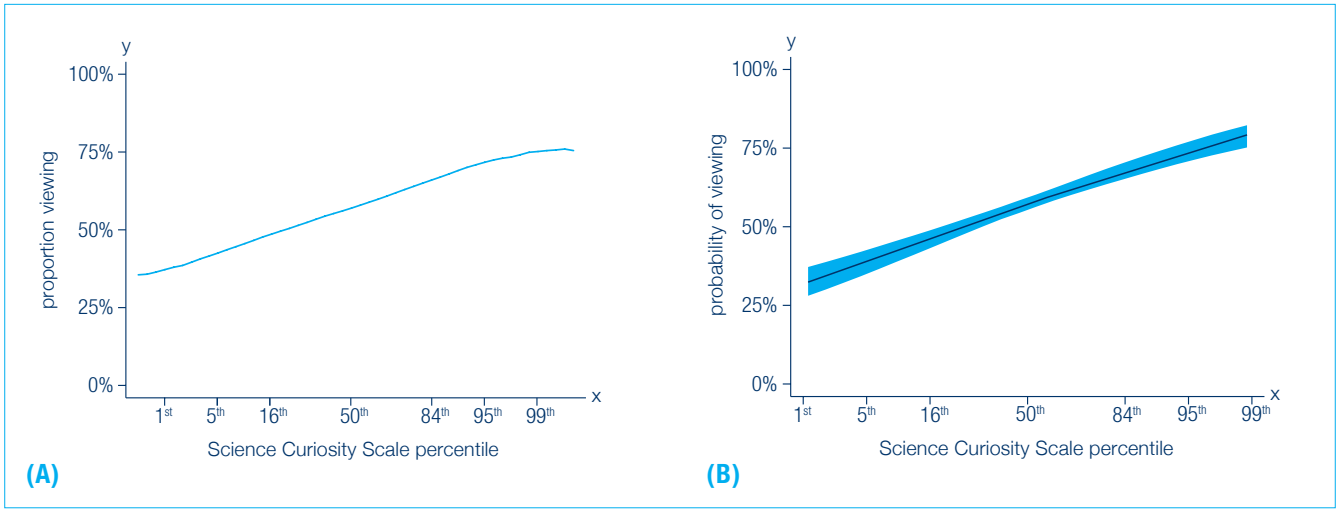


Figure 2. Impact of science curiosity on subject viewing decisions (all videos). N = 2500. Panel (A) uses locally weighted regression to plot the trend of the raw data. In effect, this data-reporting technique plots the average value of data points observed along a continuum of finely-graded intervals on the x-axis, generating a visual representation of any data trend (Cleveland 1979). Panel (B) fits a logistic regression model to the data, in which the score on SCS is the predictor and viewing decision the outcome variable. That form of analysis uses a more formal statistical model or formula to estimate how changes in the x-axis value (here SCS) influences the probability of a binary outcome (here to view or not view the assigned video) (Cohen et al. 2003). The shaded region reflects the 0.95 confidence interval of the predicted probability of viewing.

at approximately the 84th percentile in this disposition) would be 26 percentage points (± 6)[4] more likely to view the assigned episode than would a modestly incurious individual (someone one standard deviation below the population mean — or at approximately the 16th percentile in this disposition).

Nevertheless, SCS predicted that a high-science-curiosity individual's decision to view a *Deep Look* episode depended on that person's gender. It is reasonable to assume that only persons who are particularly high in science curiosity are likely to seek out and consume a YouTube video as specialized in content as a *Deep Look* episode. Accordingly, we focus here on the likely viewing decision of an individual who scored at the 95th percentile of SCS. Such an individual, the data suggest, was 18 percentage points (± 8) more likely to choose to view the assigned episode if the viewer was male than if the viewer was female.[5]

The disparity, however, varied significantly across the various episodes featured in the study (Figure 3). Indeed, among two of the videos — Lice, the video feature that women were most likely to view in the real world, and Dog Noses — the disparity among high-science-curiosity subjects was not statistically or practically different from zero. In contrast, a man scoring at the 95th percentile on SCS was 25 percentage points (± 16) more likely to view Spiders, the video with the highest real-world gender disparity, than a woman with the same SCS score. For Hairworms, the predicted gender differential was similarly substantial: 26 percentage points (± 17).[6]

While large, the gap observed in Spiders and Hairworms is smaller than the gender-viewing disparity observed in the real world. But there are additional real-world factors that need to be taken into account to assess the practical import of the experimentally observed effect of gender.

One is the gender composition of the audience for YouTube videos generally. If the composition is weighted toward males, than the experimentally observed disparity in viewing choice will be magnified in the real world in proportion to how many more men are in the effective pool of potential viewers. Available consumer data suggest that in general the YouTube audience is 60% male (Merch-Dope 2019).

Another factor is the relative distribution of science curiosity. If there are more science curious men than women (a premise of the natural audience hypothesis),

4 All "±" designations reflect 0.95 confidence intervals.

5 This estimate was formed on the basis of a logistic regression model (N = 2500) in which the decision to view the assigned episode was the outcome variable and the SCS score, gender, and appropriate cross-product interaction terms were the predictors.

6 When responses for the Spider and Hairworm conditions were pooled, the difference in the likelihood of viewing for a man versus a woman scoring at the 95th percentile on SCS was 27 percentage points, with a 6 percentage point confidence interval at 0.95.

Figure 3. Viewing decisions for individual videos in relation to science curiosity, male and female. Plotted lines based on locally weighted regression (see Figure 2 for explanation). Male-female differentials at the 95th percentile of SCS based on logistic regression model (N = 2500) in which outcome variable is decision to view the video and predictors include science curiosity, gender, experimental condition, and appropriate cross-product interactions. Predicted differences in male and female probabilities based on Monte Carlo simulation (King, Tomz & Wittenberg 2000).

then for that reason, too, the proportion of men in the potential viewing pool will be larger, and the impact of the experimentally observed disparity among *equally* science curious men and women will be larger than the disparity observed in the data.

Consistent with previous studies (Kahan, Carpenter & Landrum 2016), we found a modest but not inconsequential ($r = 0.15$, $p < 0.01$) gender difference in SCS scores. In our data, men were about twice as likely as women to score at or above the 95th percentile on the scale (Figure 4).

These factors — the likely gender imbalance in YouTube viewership and the modest difference in male and female SCS scores — can be combined with the experimentally observed gender differential in viewing decisions to form an estimate of the overall *Deep Look* gender disparity. Bayes's Theorem (posterior odds = prior odds x likelihood ratio), a formula for determining the probability of outcomes given the probability of one or more additional conditions, supplies the appropriate logical framework for doing so (Abelson 2012). Starting with a gender-balanced potential viewing pool, and assigning the appropriate gender-based likelihood ratios to the various factors relevant to determining the probability that a person scoring at or above the 95th percentile will elect to view Spiders (the *Deep Look* episode with the greatest real-world gender disparity), one would estimate an

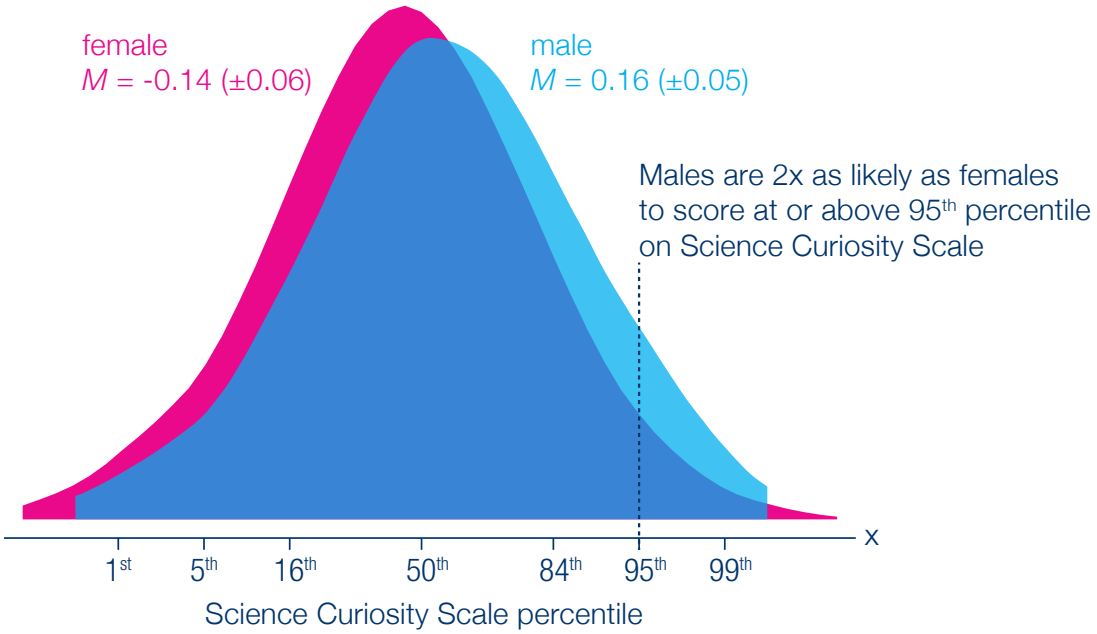


Figure 4. Relationship of science curiosity to gender. Density distribution plots based on raw data. Reported mean scores and 0.95 confidence intervals based on centered SCS scale (mean = 0).

audience that is 81% male and 19% female.[7]

This is *extremely* close to the gender disparity observed for that episode in the real-world: 86% to 14%. One is warranted, then, in concluding that the experimental procedure succeeded in essentially reproducing the Deep Look gender disparity observed outside the lab.[8]

b. Engagement parity. In addition to viewing decisions, SCS strongly predicted the level of "engagement" with the study videos (as measured with the composite scale discussed at p. 4 and note 3, above) (Figure 5). The impact of science curiosity was observed among both the scale's subjective elements (self-report satisfaction) and its behavioral ones (including viewing time and manifested intentions to engage in future viewing).

Interestingly, however, there was no *gender disparity* in this relationship. For men and women, as science curiosity increased, so did engagement.

This conclusion suggests that the *Deep Look* gender disparity is *not* based on some element of the content of the relevant episodes but rather on a viewing decision necessarily made before women view them. Indeed, *were they to view the episode*, women

7 The relevant likelihood ratios in this calculation were 1.5 for the YouTube audience gender disparity (60% male, 40% female); 2 for the science-curiosity gender disparity (men 2x as likely as women to score at or above the 95th percentile on SCS); and 1.4 for the experimentally observed gender disparity for Spiders (85% predicted male viewership, 60% female, for subjects at 95th percentile on SCS), the episode with the largest real-world disparity. For those who are (admirably) curious, the nature of this computation can be explicated in straightforward terms using natural frequencies (Sedlmeier & Gigerenzer 2001; Spiegelhalter, Person & Short 2011). We do not know what fraction of a gender-balanced potential viewer pool subscribe to YouTube. But we know that the YouTube audience is skewed 60% to 40% toward males. Thus, among 100,000 potential *Deep Look* viewers, we would (based on the YouTube gender imbalance), expect 60,000 to be male, 40,000 female. Overall, of course, we would expect 500 of every 10,000 potential viewers to score at or above the 95th percentile on SCS. But because twice as many men as women achieve that score, we would expect approximately 667 of every 10,000 men and 333 of every 10,000 women, to achieve that score. Thus, from our pool of 60,000 potential male YouTube viewers and 40,000 female ones, we would expect to find approximately 4002 men (6 x 667), and approximately 1332 women (4 x 333) scoring at or above the 95th percentile on SCS. Based on the experimental data, we would then expect about 3402 of those men (0.85 x 4002) and about 799 of those women (0.60 x 1332) to choose to view Spiders, the feature with the highest gender imbalance. Based on these calculations, one would estimate that 81% of the viewers would be male and 19% female.

8 We performed two additional sets of analyses to check for possible sources of bias. In one, we constructed and substituted a "wildlife interest" scale for the Science Curiosity Scale based on concern that the SCS might be too focused on "hard science" issues and insufficiently on nature. We found that the former scale was highly correlated with the latter and generated results that did not differ materially from those generated by SCS in our key analyses. We also identified subjects who reported either subscribing to *Deep Look* on YouTube or having been exposed to at least one episode of it. The percentage responding affirmatively to the relevant queries were small (5% and 7%, respectively), and exclusion of them did not materially change any analysis results.

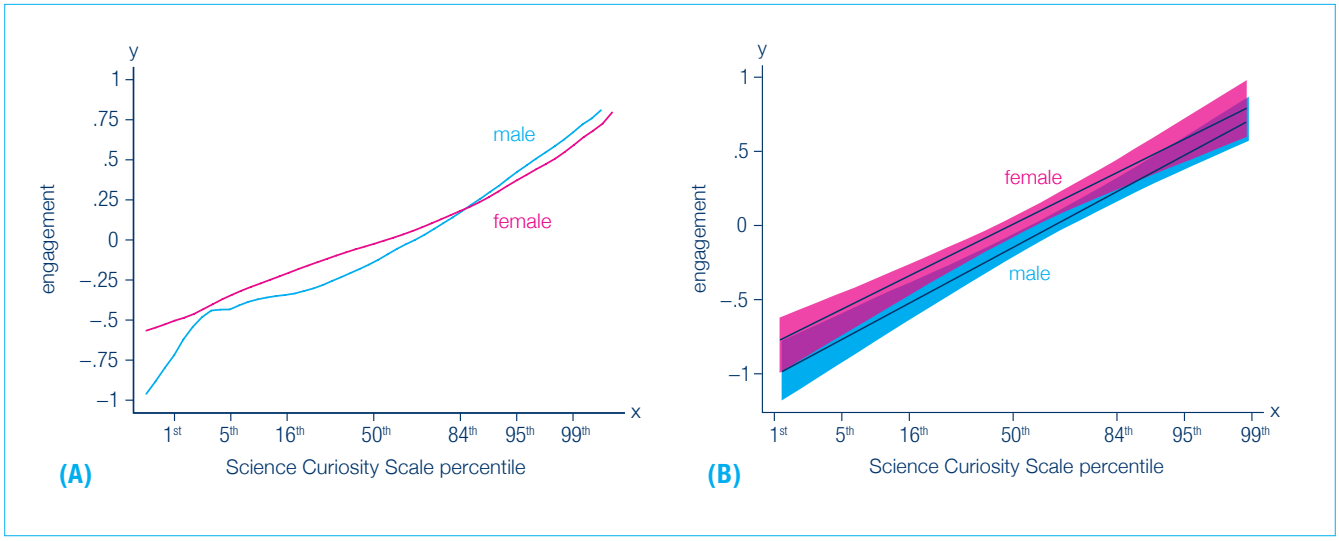


Figure 5. Engagement in relation to science curiosity, male and female (all videos). N = 2500. "Engagement" refers to the study composite engagement scale (see p. 4 and note 3, above). The scale is scored in standard deviations and centered on 0. Plotted lines in panel (A) are based on application of locally weighted regression to raw data (see Figure 2 for explanation). Panel (B) fits a linear regression model to the data, in which the score on SCS, gender, and an appropriate cross-product interaction term are the predictors and engagement the outcome variable. A linear regression uses a mathematical model or formula to estimate the relationship between values on the x-axis (here SCS), conditional on other specified values of interest, to values on the y-axis (here engagement level). The shaded region reflects the 0.95 confidence interval of the estimated engagement level.

of high levels of science curiosity would derive just as much satisfaction from them as men of comparable science curiosity.

This conclusion, however, must be viewed with some caution. We did not randomly assign subjects to view the assigned features. Necessarily, then, we cannot be fully certain that if the study design had obliged high-science curiosity women to view the assigned episode, those who otherwise would have elected not to do so would have been just as engaged as the women who would have chosen voluntarily to view it.

c. The gender-specific interaction of science curiosity and science comprehension. The data also identified an influence of special significance for the gender-viewing disparity: science comprehension.

In general, there is only a modest correlation ($r = 0.19$, $p < 0.01$) between science comprehension (as measured by Ordinary Science Intelligence (OSI)) and science curiosity (as measured by SCS). Plenty of individuals of only moderate science comprehension still take immense pleasure in knowing what science knows, while many of high science comprehension get comparatively little satisfaction from contemplating the discoveries of scientific investigation. Consistent with these relationships, we found that SCS predicted viewing decisions and engagement more strongly than science comprehension as measured by OSI (Figure 6).

We also examined the interaction of science curiosity and science comprehension. Among study subjects assigned to the conditions featuring Dog Noses and Lice, SCS scores uniformly predicted the viewing decisions of study subjects, male and female, of all levels of science comprehension. But this was not the case for those assigned to the conditions featuring Hairworms and Spiders — the two conditions that evinced gender disparities in viewing decisions. Among those subjects, SCS scores uniformly predicted the decision of men and women independently of OSI scores *only* for subjects scoring relatively high of OSI (in the top 50th percentile). Among subjects of relatively modest OSI scores (the bottom 50th percentile), we observed the signature gender gap among the highest science curiosity individuals (Figure 7).

In other words, the experimentally observed gender disparity for Hairworms and Spiders was attributable largely to high-science-curiosity women of modest science comprehension. Those study subjects were substantially less inclined to view Spiders and Hairworms than were high-science-curiosity men whose science comprehension levels were comparably modest.

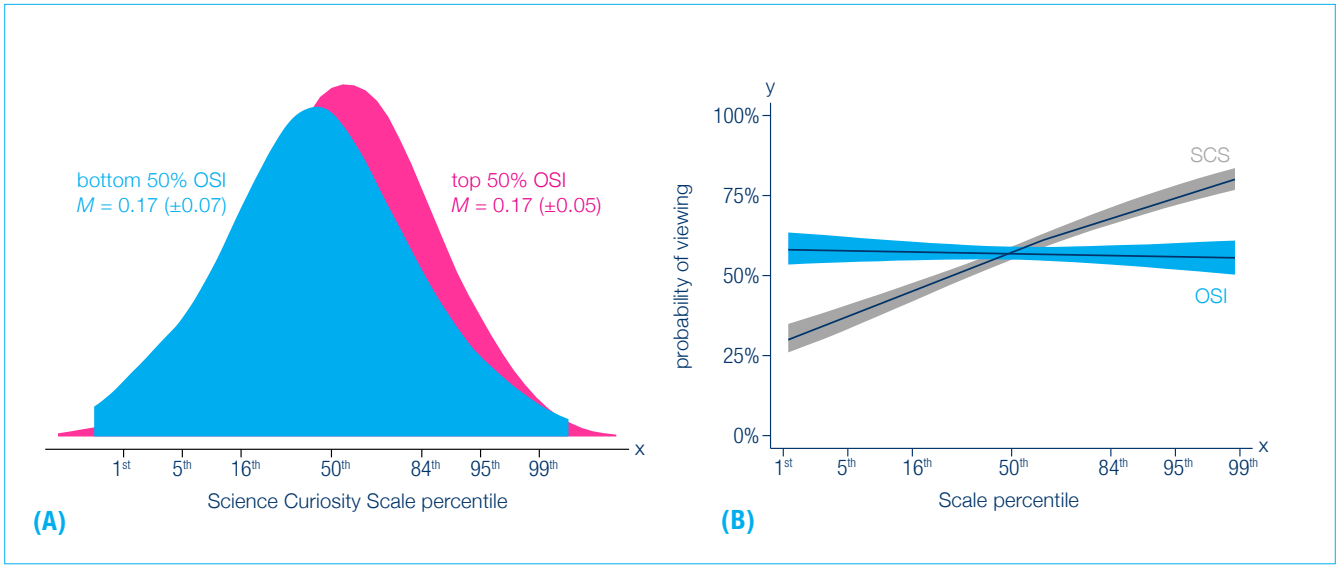


Figure 6. Relationship of science curiosity (SCS) and science comprehension (OSI), and respective impact on predicted viewing decisions (all videos). Density distribution plots in Panel (A) based on raw data. Reported mean scores and 0.95 confidence intervals based on centered SCS scale (mean = 0). Panel (B) fits a logistic regression model to the data, in which the score on SCS, OSI, and appropriate cross-product interaction term are the predictors and viewing decision the outcome variable. The x-axis reflects the percentile for the scale associated with the plotted relationship. In each plotted relationship, the impact of the indicated scale is predicted controlling for the other scale, the value of which is set to its mean. The shaded regions reflect the 0.95 confidence interval of the predicted probability of viewing the assigned *Deep Look* episode.

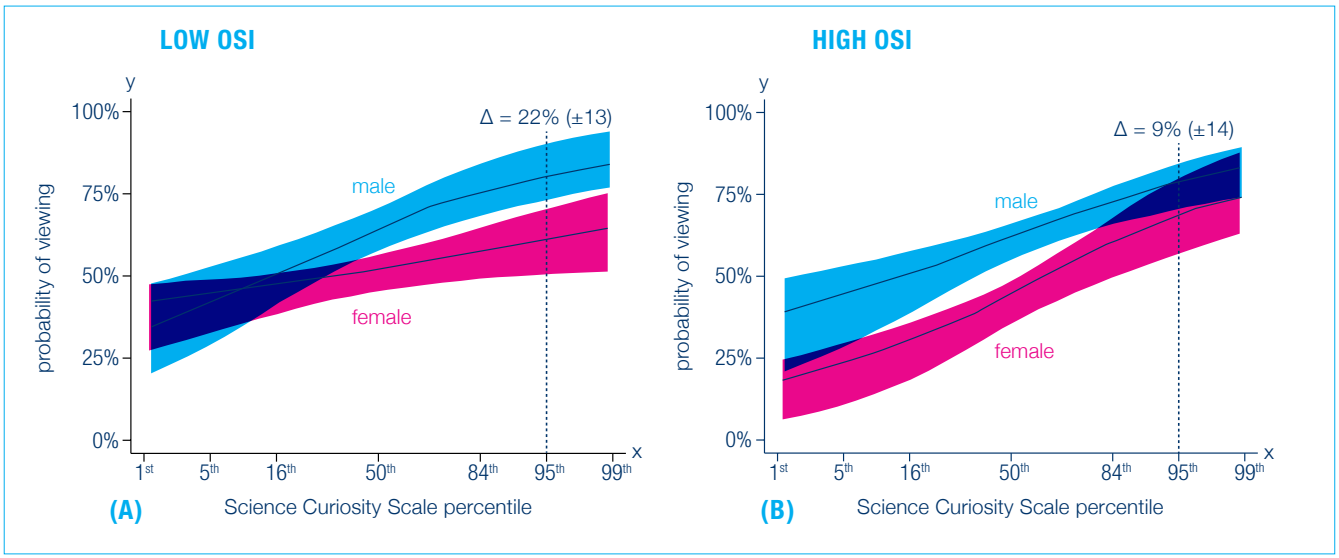


Figure 7. Relative impact of science curiosity on viewing decisions for men and women of varying levels of science comprehension (Hairworms and Spiders only). N =1235. Samples of Hairworms and Spiders conditions combined. Based on a logistic regression model, in which the score on SCS, OSI, gender, and appropriate cross-product interaction term are the predictors and viewing decision is the outcome variable. Scores on OSI are set at +1 standard below the mean Panel (A), and -1 standard deviation below in Panel (B). The shaded regions reflect the 0.95 confidence interval of the predicted probability of viewing the assigned Deep Look episode.

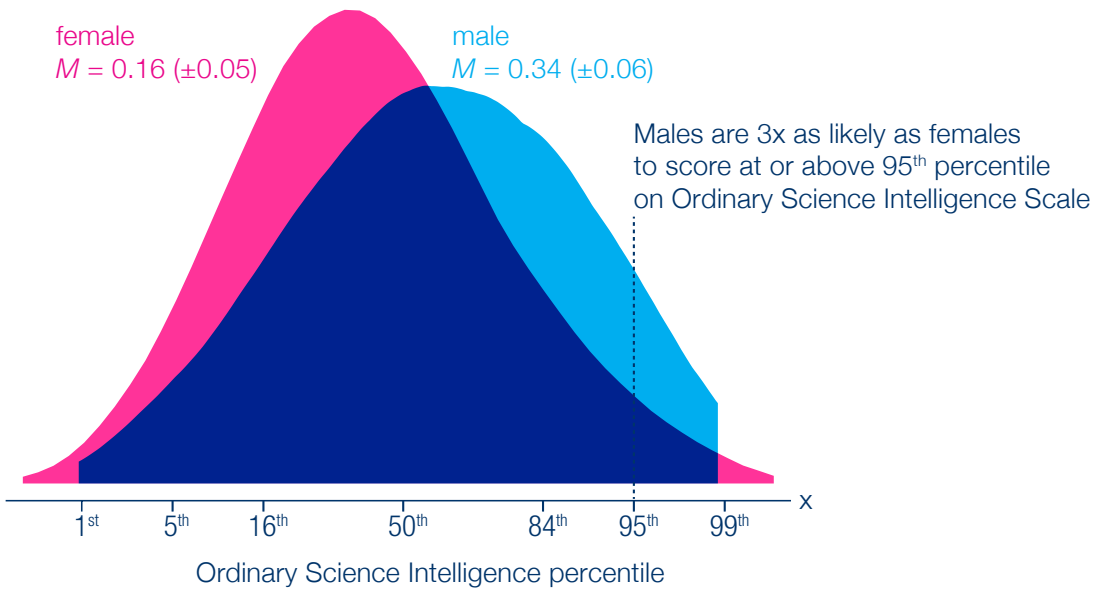


Figure 8. Gender distribution of OSI scores. Density distribution plots based on raw data. Reported mean scores and 0.95 confidence intervals based on centered OSI scale (mean = 0).

It is also important to note that OSI scores display a fairly marked gender disparity. In a result consistent with standardized assessments of science and math aptitude, men are approximately three times more likely to score at or above the 95th percentile on OSI than are women (Figure 8).

This pattern can be expected in the viewing-decision gender disparity associated with the interaction between science comprehension and science curiosity among women. Again, it makes sense to assume that in the real-world the decision to regularly view episodes of a YouTube series as specialized as *Deep Look* will be confined to individuals who possess a very high level of science curiosity. The gender disparity in OSI scores implies that among highly science curious individuals, the proportion who are modest in science comprehension will be greater among women than men. The greater propensity of such women to forgo viewing certain *Deep Look* episodes, then, will necessarily aggravate the gender imbalance associated with factors such as the male bias in YouTube usage and the (slight) gender gap in science curiosity.

d. The attenuated role of "disgust." "Disgust" is a species of emotion that consists in revulsion toward potential contaminants (Nussbaum 2004; Miller 1997). The contaminants may be bodily ones — like toxic substances or bodily excretions. But they can also include forms of moral deviancy, which are understood to expose individuals to experiences and tastes that degrade their character and make them more vulnerable to engaging in harmful behavior (Chapman & Anderson 2013; Gutierrez & Giner-Sorrolla 2012; Miller 1997; Douglas 1966).

A character disposition related to experiencing disgust was identified by study subjects' responses to the study's risk-perception battery (p. 4; see also Appendix). Consistent with other studies (Cultural Cognition Project 2014), those responses cohered along two factors or dimensions (Figure 9). One of these displayed concern over "physical harms," such as those associated with widespread private gun possession, with global warming, with fracking, and with air pollution. The other — clustered on fear of GMOs, fluoridated water, x-rays, sex education, prostitution, and drug use — evinced worry over substances and behaviors that (in some persons, at least) evoke disgust.

There was no meaningful relationship, however, between scoring high in this form of disgust-risk concern and reluctance to view *Deep Look* features (Figure 10). Nor did scoring high on this factor predict less engagement with these features.

In addition, this relationship was uniform among men and women. Accordingly, there was no link between the experimentally observed gender disparity in viewing decisions and this measure of disgust sensitivity.

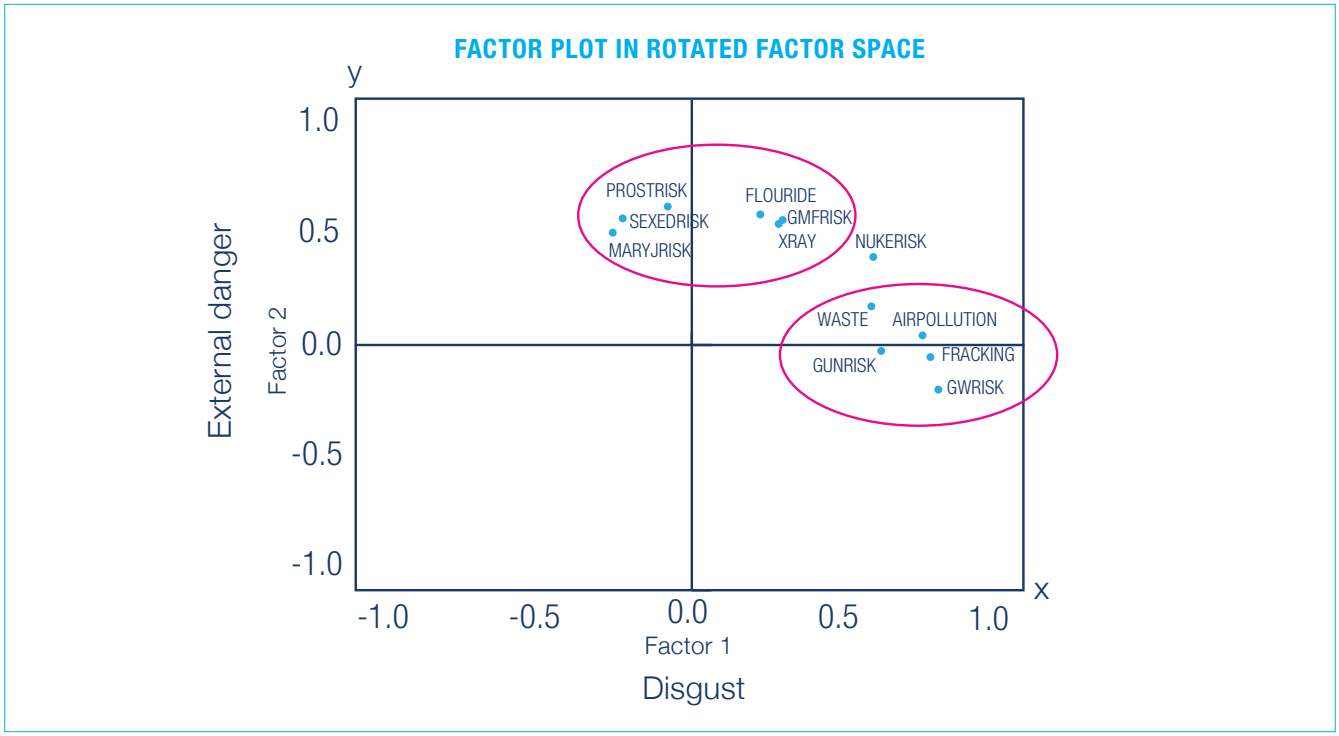


Figure 9. Risk perception sensibilities. Factor analysis results of responses to risk-perception battery items.

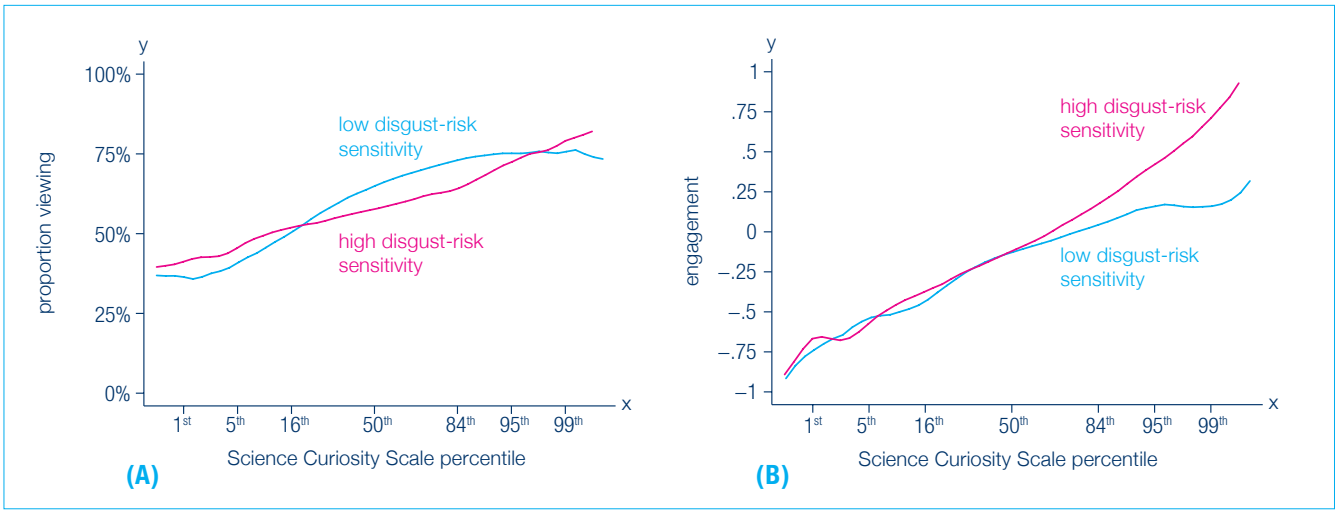


Figure 10. Impact of disgust-risk concern on viewing decision and engagement (Spiders only, all subjects). Ns = 618 (Panel (A)) and 374 (Panel (B)). "High" and "low disgust-risk sensitivity" reflect division of subjects in relation to median score on "Disgust" factor. Locally weighted regression plots.

The conclusion that disgust was not a source of the viewing-decision gender disparity was reinforced by subjects' explanations for declining to view the assigned *Deep Look* episode. All subjects who elected not to view the episode were solicited to respond to an open-ended item inviting them to "tell us why [they] chose not to watch." Responses were coded into various categories (e.g., short-age of time, lack of interest in the topic, resentment or suspicion of "clickbait" material) including the perceived disgustingness or "grossness" of the assigned video.

Disgust-related considerations played a relatively large role in subjects' proffered explanations for declining to view Lice. Some 40% of the female subjects and about 17% of men who responded to this item identified grossness or disgustingness as a reason not to view the episode (Figure 11).

As noted (p. 6 and Figure 3), however, there was no gender disparity in viewership for Lice, either in the real-world or in the experiment. The marked gender disparity in offering disgust-related explanations for declining to view this feature, then, was necessarily unrelated to any actual gender disparity in viewing decisions.

The episode that did generate the highest gender-based viewership disparity in the real world was Spiders. As discussed, that disparity was reproduced experimentally (pp. 5-10). Significantly, however, only a trivial percentage of women (5%) offered disgustingness or grossness as their reason for choosing not to view Spiders. Neither the volume nor the gender-disparity in proffered disgust explanations seems large enough to contribute meaningfully to the observed gender disparity in viewing decisions for this episode.

Hairworms did generate an experimentally observed gender disparity in viewing decisions (p. 6 and Figure 3) and a gender disparity in proffered disgust-related explanations among nonviewers. Nevertheless, the ratio of female to male disgust-related references did not vary meaningfully in that condition from the ratio of female to male disgust-related references in Lice, where, again, no meaningful gender disparity was observed in viewing decisions. So here, too, the disparity in disgust references does not support an inference that disgust or the anticipation of it was a genuine source of viewing-decision gender disparities.

2. Secondary findings

a. Relative popularity of features. Beyond the *Deep Look* gender disparity, a secondary focus of the study was the relative popularity of *Deep Look* episodes. Again, Dog Noses and Hairworms were included in the study because they were, respectively, the *Deep Look* episodes least and most viewed on YouTube over the time period selected for review.

The real-world popularity of these two features was not mirrored in the study results. On the contrary, some 68% of the subjects assigned to Dog Noses chose to watch some or all of the video, making it the episode in the study most likely to be viewed. Only

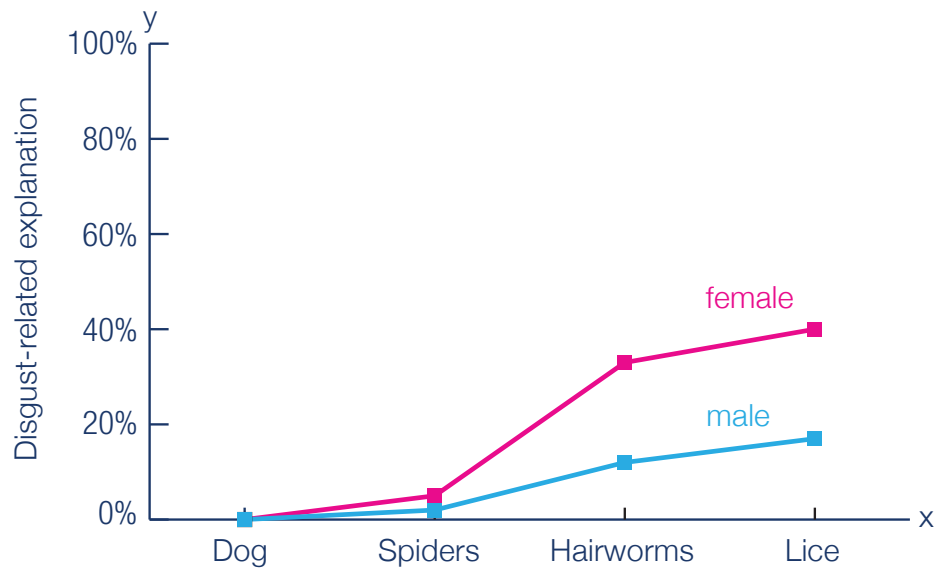


Figure 11. Percentage volunteering disgust-related explanation for decision not to view, by gender.

51% of the subjects afforded the chance to view Hairworms did so.

This differential in popularity, however, paints a highly misleading picture. What it ignores is the relative impact of *science curiosity* on viewing decisions for the indicated episodes (Figure 12).

Science curiosity played a relatively small role in the decision of subjects to view Dog Noses. Even decidedly incurious subjects — ones who scored at the 10th percentile of SCS — were likely to choose to view some or all of that feature. Subjects of average curiosity (50th percentile at SCS) were highly likely — over 70% — to do so.

Science curiosity was much more consequential for subjects' decisions to view Hairworms. Indeed, at the 10th percentile of SCS, there was less than a 30% chance they would view it, in whole or in part. Even at the 50th percentile of SCS, subjects were decidedly less likely than not to view it. In order for them to be 70% likely to view Hairworms, study subjects had to score at or above the 95th percentile on the Science Curiosity Scale (Figure 12).

These results suggest a plausible reconciliation of the study and real-world results — and a resulting insight into the relative popularity of *Deep Look* features generally. Dog Noses is a program the expected appeal of which does not depend particularly on science curiosity: lots of people, including many with low science interest, are interested in consuming information about dogs. Teeming as YouTube is, though, with videos on dogs, one would not expect anyone, including *Deep Look* viewers, to tune into *Deep Look* to get their "fix" on dog-related content.

What science curious individuals *do* tune into the *Deep Look* series to acquire is scientific insights unlikely to be widely available elsewhere in the media universe. Whether as regular subscribers or as individuals exposed to particular episodes by virtue of their higher rate of propagation among science curious individuals, the science curious are more likely to choose to view *Deep Look* episodes in proportion to how uniquely those episodes satisfy their desire to consume interesting (or indeed surprising and awe-inspiring) science content. A feature that examines the power of hairworms to "control the minds" of crickets after "eating them alive" more clearly fits into that category.

In sum, the relative popularity of *Deep Look* features is likely to turn on how close the content of the feature comes to the distinctive niche of *Deep Look* in the social media environment. *Deep Look* caters to highly science curious individuals. Thus, even when informed by scientific methods, episodes that examine subjects of wider interest — ones that are likely to appeal even to individuals of only modest science interest — are unlikely to be as popular as the episodes that hone in on more unusual phenomena, the appeal of which consists much more decisively on the power of science to illuminate the surprising workings of nature.

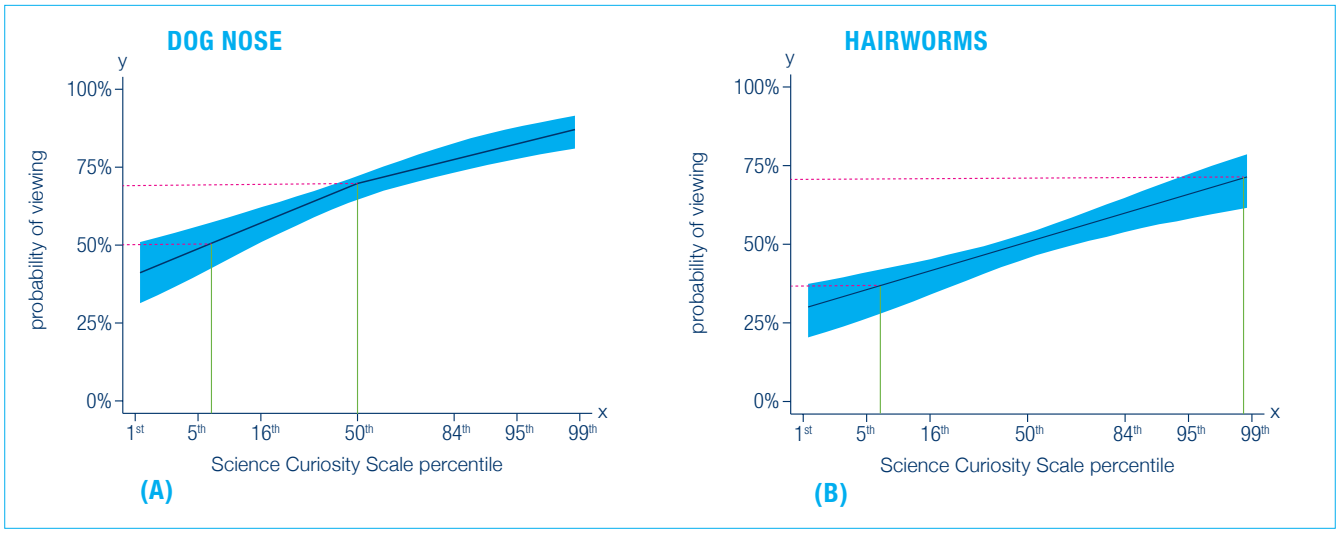


Figure 12. Relative impact of science curiosity on decisions to view Dog Noses and Hairworms. Ns 623 (Panel (A)) and 617 (Panel (B)). Based on logistic regression models, in which the score on SCS is the predictor and viewing decision for the indicated episode the outcome variable. The shaded regions reflect the 0.95 confidence interval of the predicted probability of viewing. The points denoted by intersections of solid vertical and dashed horizontal lines illustrate how much more consequential science curiosity was for viewing of Hairworms.

b. Millennials. Consistent with previous data collections (McCann et al 2019), the data collected in this analyses found that Millennials are more science curious (Figure 13). But the margin of superiority was very small, particularly in relation to members of Generation X. Baby boomers, too, displayed science-curiosity levels only slightly lower than those of Millennials. Members of Generation Z and the Silent Generation, in this analysis, evinced less science curiosity.[9]

Arguably more interestingly, science curiosity did not influence Millennials in a manner different from members of other age cohorts. That is, Millennials were no more or less likely than others to choose to view *Deep Look* features, or to be engaged by them, in relation to their levels of science curiosity. Nor did the gender disparity observed in the sample as a whole differ among Millennials. These findings imply that, in this setting at least, Millennials are not interacting with *Deep Look* material differently from members of other age cohorts, notwithstanding that *Deep Look* features are characteristic of science content disseminated on social media rather than by more traditional forms of content distribution.

The only observed difference between Millennials and members of other cohorts relate to self-reported subscriptions to the *Deep Look* series and to prior exposure to individual series episodes. The proportions were modest in size but still higher for Millennials (Figure 14).

c. Other individual characteristics. We also examined the influence of other individual characteristics such as religiosity, political outlooks, education, and income. In addition to gender, many of these characteristics are also likely to feature in variants of the missing audience hypothesis. However, none of these influenced either the decision to view or engagement with the *Deep Look* episodes featured in the study.

E. Analysis and interpretation

1. Some provisional answers

Earlier (pp. 4-5), we identified a number of hypotheses that motivated the current study design. What can we infer about these hypotheses based on the study results?

To begin, these results are more consistent

[9] For purposes of analyses comparing age cohorts, we used the age cohort definitions identified by the Pew Research Center.

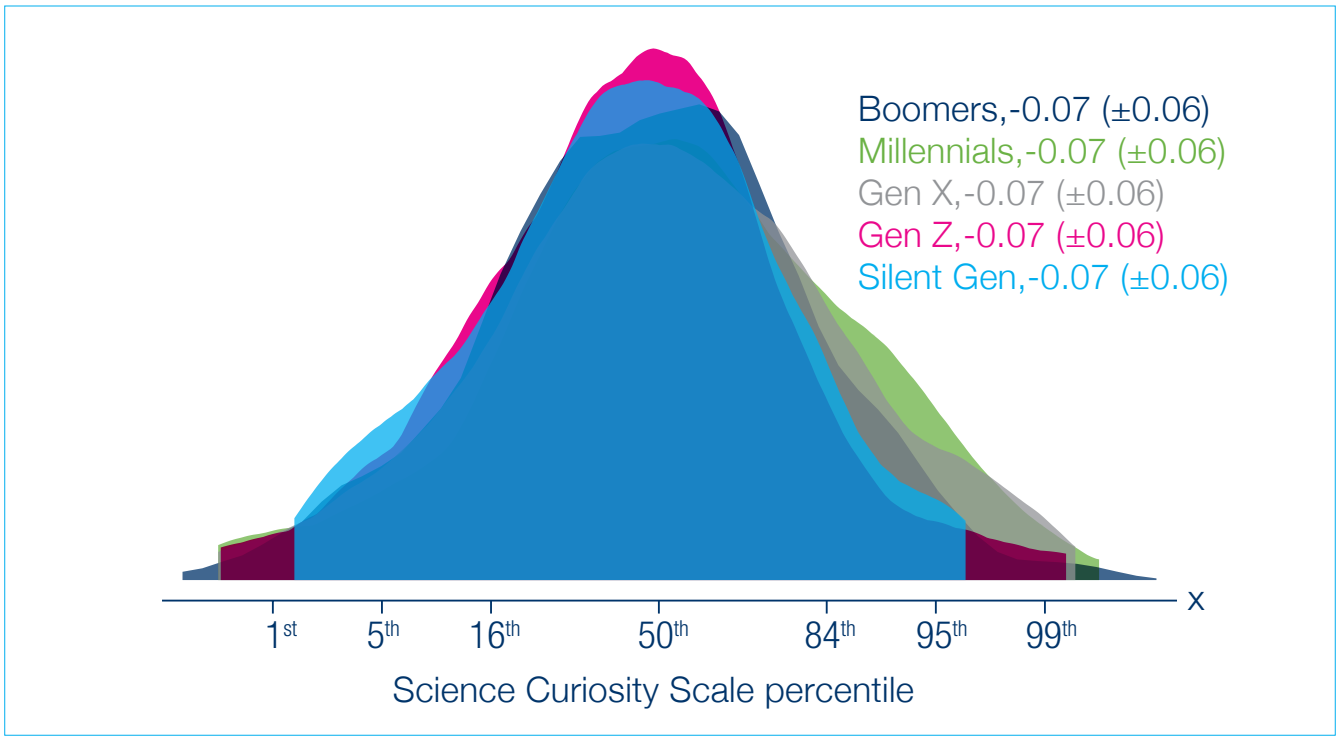


Figure 13. Distribution of science curiosity among age cohorts. Density distribution plots based on raw data. Reported mean scores and 0.95 confidence intervals based on centered SCS scale (mean = 0).

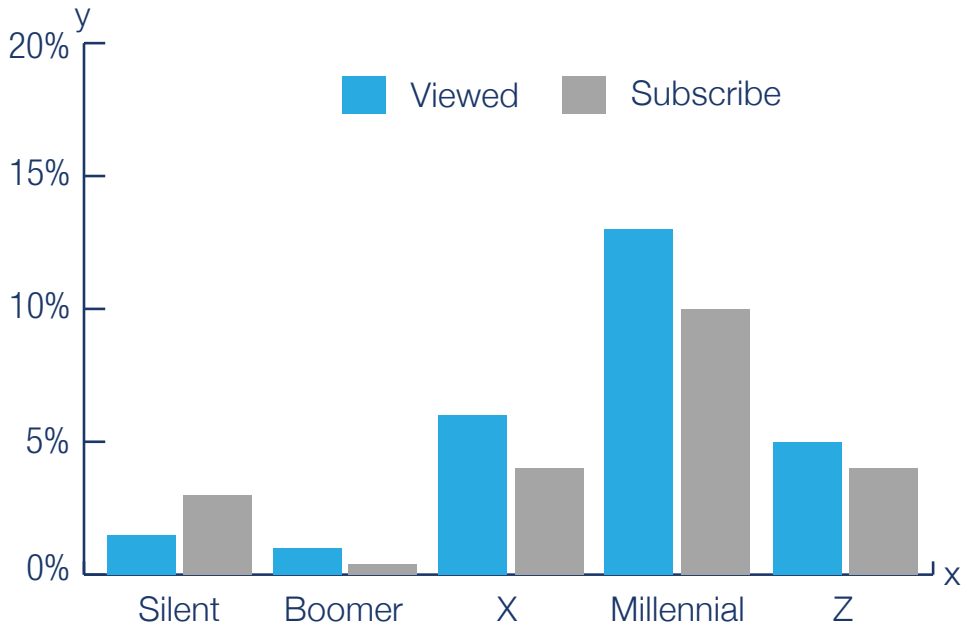


Figure 14. Self-reported subscription to Deep Look series or prior exposure to any Deep Look episode by age cohort.

with the missing audience hypothesis than with the natural audience hypothesis. Based on SCS scores, one would say there is a small but significant disparity in science curiosity among men and women. But as the data show, even among men and women of equally high levels of science curiosity, men are more likely than women to view *Deep Look* features.

Moreover, the gender difference in SCS scores does not predict a viewership disparity nearly as dramatic as that observed in the real world. To reproduce that level of disparity, one must take the uneven viewing decisions of high-science-curiosity men and high-science-curiosity women into account. Thus, the data are consistent with the conclusion that the real-world disparity consists of a "missing audience" of women who, based on their high-science-curiosity scores, would be expected to get just as much pleasure as their male counterparts from viewing features such as Hairworms and Spiders.

Similarly, the data suggest reason to doubt what we have called the "algorithm hypothesis." There is a gender-based viewing disparity among high-science-curiosity individuals. Moreover, taking into account additional factors themselves not associated with content-propagation (namely, the YouTube audience gender disparity and the small gender disparity in science curiosity), the experimentally observed gender disparity in viewing decisions can reproduce the disparity observed in the real world. Since the experiment was designed to remove any influences associated with propagation dynamics distinctive of social media, there is reason to believe the observed disparities in the real world can be accounted for based on factors — including, decisively, the relative reluctance of high-science-curiosity women to view certain *Deep Look* episodes — that are themselves unrelated to YouTube's search or user-recommendation algorithms.

The study also generated evidence that the *Deep Look* gender disparity is being driven by differences in the viewing decisions — but not the engagement levels—of high-science-curiosity women who are modest in science comprehension. Women of high-science-curiosity but modest science comprehension are less likely than men of high-science-curiosity but modest science comprehension to view the *Deep Look* features in which a gender-based viewing disparity is observed. **It can thus be inferred that some unobserved influence that accounts for *Deep Look's* missing audience problem is itself related to gender-specific effects of science comprehension on the decision of women to consume potentially engaging science content.**

The study furnished little support, however, for thinking disgust sensibilities are related to the *Deep Look* gender disparity. We can think of no particular reason why any gender-based disgust sensibility is related to differences in science comprehension. Moreover, the study's measure of disgust-related risk concerns was uncorrelated with decisions to view or engagement with *Deep Look* videos, including those that displayed a gender bias.

Finally, there was no coherent relationship between disgust and gender in the open-ended responses of subjects who elected not to view the *Deep Look* features. Indeed, it is more plausible to think that the decision to view was driven by some other unobserved factor, and that the invoking of disgust was a confabulation that varied by gender for reasons related to conventional social scripts.

2. Some additional questions

All science research, including that relating to human behavior, tends to be characterized by the conservation of perplexity. Evidence that supplies more confident answers to particular questions invariably begets additional questions for which confident answers are lacking. This study is no exception.

a. Why does science comprehension matter? As indicated, we found, unexpectedly, that the *Deep Look* gender disparity was being driven by women who are high in science curiosity but modest in science comprehension. Again, that pattern suggests that some influence connected to science comprehension is part of *Deep Look's* distinctive missing audience problem.

But what exactly *is* that influence? Research suggests that as a result of internalized stereotypes, women often perform less well in science tasks and evince a lower science interest level than one might anticipate based on their science knowledge (Appel & Kronberger 2011; Stout et al. 2011). Could some form of inhibition of this sort explain the interaction we observed between science comprehension and science curiosity among women? If so, is there a way to counteract that inhibition that could readily be used by *Deep Look* producers to offset the gender disparity in episode viewing decisions?

These are matters that the research team anticipates examining in follow up studies.

b. Viewing and engagement. Also worthy of additional investigation is the marked difference in impact that gender had on *viewing decisions*, on the one hand, and *engagement*, on the other. One inference that could plausibly be drawn from the data is that high-science-curiosity women who choose not to view certain *Deep Look* episodes would be just as engaged as high-science-curiosity men were the former to actually watch the episodes. If so, what is the cue or preconception that discourages such women from choosing to watch a feature they would actually find to be satisfying? Informed conjectures are in order, as are appropriate experimental methods for testing them.

c. The persistence of the disgust hypothesis. Finally, while we have emphasized that data from this study suggests reason to doubt the surmise that potential revulsion or disgust plays a significant role in the *Deep Look* gender gap, we do not believe that hypothesis should be conclusively discarded. Better measures might still reveal that disgust, possibly in conjunction with other influences, plays a role. Examining this possibility as a rival hypothesis can also help sharpen study designs that examine alternative explanations, including ones related to internalized stereotype inhibitions or other influences that preempt the viewing decisions of science-curious women who would in fact derive considerable personal satisfaction from viewing *Deep Look* features.

F. Conclusion

This study had two broad objectives.

One was to examine, specifically, the sources of the *Deep Look* gender disparity. As described in this report, definitive answers are not yet in hand. Nevertheless, the right questions to be asked about the sources of that disparity are even now more readily apparent and ripe for additional study.

The second objective was to advance the science of science communication by doing the study in a real-world science communication setting. The unique researcher/professional communicator collaboration featured in this study program, we believe, is essential to the full maturation of the science of science communication and to any genuine contribution this emerging field can make to societal well-being (National Academy of Sciences 2017).

In our experience in pursuing this form of research described in this report, we have encountered intense excitement but also understandable uncertainty among science-media producers and related science-communication professionals. The source of both is the intriguing question of whether a craft — science communication — that is necessarily dependent on a highly particular, finely tuned species

of professional intuition can be enhanced by the use of highly general and highly analytical methods such as those featured in the science of science communication.

There is, of course, only one valid way to determine whether a marriage of such outlooks and methods can be a harmonious one. And that is to try it and see.

The very initiative to test whether the tools and concepts of the science of science communication can be profitably integrated into the practice of science communication is thus itself a meta-experiment. Whether the results corroborate or discredit the conjecture that such a synthesis can be worthwhile is a determination that those involved in both research and professional communication must make — and make by acting on the basis of their assessments of the type of findings described in a report such as this one.

We are far from the end of the CTC research project of which the current study was a part. But we believe that we are already in a position, by detailing our preliminary findings and our process for deriving them, to contribute to the body of knowledge that will ultimately help researchers and practitioners judge the utility of the science of science communication for science communication as a profession. We offer this preliminary report in that spirit.

REFERENCES

Abelson, R.P. *Statistics as principled argument* (L. Erlbaum Associates, Hillsdale, N.J., 1995).

Appel, M., Kronberger, N. & Aronson, J. Stereotype threat impairs ability building: Effects on test preparation among women in science and technology. *European Journal of Social Psychology* **41**, 904-913 (2011).

Chapman, H.A. & Anderson, A.K. Things rank and gross in nature: A review and synthesis of moral disgust. *Psychological Bulletin* **139**, 300-327 (2013).

Cleveland, W.S. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association* **74**, 829-836 (1979).

Cohen, J., Cohen, P., West, S.G. & Aiken, L.S. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (L. Erlbaum Associates, Mahwah, N.J., 2003).

Cultural Cognition Project. Vaccine Risk Perceptions and Ad Hoc Risk Communication: An Experimental Investigation. *Cultural Cognition Risk Perception Studies* Rep. No. 17 (Jan. 27, 2014), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2386034.

DeMars, C. *Item response theory* (Oxford University Press, Oxford, New York, 2010).

Douglas, M. *Purity and danger; an analysis of concepts of pollution and taboo* (Praeger, New York, 1966).

Gutierrez, R. & Giner-Sorolla, R. Anger, disgust, and presumption of harm as reactions to taboo-breaking behaviors. *Emotion* **7**, 853-868 (2007).

Jamieson, K.H., Kahan, D.M. & Scheufele, D.A. *The Oxford Handbook of the Science of Science Communication* (Oxford University Press, 2017).

Kahan, D.M., Carpenter K., & Landrum, A. (2015). Evidence-based science filmmaking initiative. Study No. 1. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2713563.

Kahan, D.M. & Carpenter, K. Out of the lab and into the field. *Nature Climate Change* **7**, 309 (2017).

Kahan, D.M. 'Ordinary science intelligence': a science-comprehension measure for study of risk and science communication, with notes on evolution and climate change. *J Risk Res* **20**, 995-1016 (2017).

Kahan, D.M. Climate-Science Communication and the Measurement Problem. *Advances in Political Psychology* **36**, 1-43 (2015).

Kahan, D.M. Making Climate-Science Communication Evidence-Based—All the Way Down. *Culture, Politics and Climate Change* (ed. M. Boykoff & D. Crow) 203-220 (Routledge Press, New York, 2014).

Kahan, D.M. What is the "science of science communication"? *J. Sci. Comm*, **14**, 1-12 (2015).

Kahan, D.M., Landrum, A., Carpenter, K., Helft, L. & Hall Jamieson, K. Science Curiosity and Political Information Processing. *Political Psychology* **38**, 179-199 (2017).

King, G., Tomz, M. & Wittenberg., J. Making the Most of Statistical Analyses: Improving Interpretation and Presentation. *Am. J. Pol. Sci.* **44**, 347-361 (2000).

McCann, S. Cracking the Code: Survey Results on Millennials and Their Science Curiosity. *KQED* (May 1, 2019), available at https://ww2.kqed.org/about/2019/05/01/cracking-the-code-survey-results-on-millennials-and-their-science-curiosity/.

MerchDope. 37 Mind Blowing YouTube Facts, Figures and Statistics – 2019, (July 29, 2019), available at https://merchdope.com/youtube-stats/.

REFERENCES, CONTINUED

Miller, W.I. *The Anatomy of Disgust* (1997).

National Academies of Sciences, E.M., Education, D.B.S.S. & Agenda, C.S.S.C.A.R. *Communicating Science Effectively: A Research Agenda* (National Academies Press, 2017).

Nussbaum, M.C. *Hiding from humanity: Disgust, Shame, and the Law* (Princeton University Press, Princeton, N.J., 2004).

Schellenberger, R.E. Criteria for Assessing Model Validity for Managerial Purposes. *Decision Sciences* **5**, 644-653 (1974).

Sedlmeier, P. & Gigerenzer, G. Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General* **130**, 380-400 (2001).

Spiegelhalter, D., Pearson, M. & Short, I. Visualizing Uncertainty About the Future. *Science* **333**, 1393-1400 (2011).

Stout, J.G., Dasgupta, N., Hunsinger, M. & McManus, M.A. STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology* **100**, 255-270 (2011).

APPENDIX: STUDY INSTRUMENT

Intro. We would like to include you as a participant in a research study. The purpose of the study is to learn about your interests in a wide variety of topics and activities. Participation is voluntary, and you may decline to answer any questions that you do not want to answer. The survey will take about 25 minutes to finish.

___ Agree to take survey

___ No thanks

Science Curiosity Scale (SCS)

1. News Interest

There are a lot of issues in the news and it is hard to keep up with every area. We will list some topics that get covered in the media. Please indicate **how closely you follow the news relating to each topic** either in the newspaper, on television, on radio, or on the internet. [1. Not at all; 2. A little but not closely; 3. Closely but not very closely; 4. Very closely.]

[randomize, separate page]

1. NICRIME. Crime
2. NIED. Education
3. NIPOLGOV. Government or politics
4. NISPORT. Sports
5. NIRELIGION. Religion
6. NIINTL. International affairs
7. NIBUSINESS. Business or finance
8. NISCIENCE. Scientific research or discoveries
9. NITECHNOLOGY. New technologies
10. NIENTERTAIN. Entertainment or celebrities
11. NIWEATHER. Local weather

1. Leisure Activity

We'd also like to know about some of the activities you engage in. We will list some activities. Please indicate **how many times you engaged in the activity in the last year.**

[randomize except PUBLICLECTUREb. immediately after PUB LECTUREa; separate page;]

1. LASCIMUSEUM. visited a science or technology museum
2. LASPORTS EVENT. attended a live sports event
3. LAART MUSEUM. visited an art museum
4. LAMUSICAL. attended a musical performance or concert
5. LAZOO. went to a zoo or aquarium
6. LALIBRARY. went to a public library
7. LAGUNSHOW. went to a gun show
8. LATHEMEPARK. visited a theme park or amusement part
9. LAHUNT. went hunting or fishing
10. LAHIKE. went hiking
11. LAPUB LECTUREa. Attended a public lecture
12. [if yes on PUBLICTUREa] PUBLECTUREb. What was the topic of the lecture you attended? If you attended more than one, please indicate all that apply.
    a. history
    b. science or technology
    c. public affairs or politics

d. religion
e. economics
f. other

11. LAPOLEVENT. participated in a political rally or attended a political event

1. Books

We'd also like to know whether you read books in your spare time. We will list some book topics. Please indicate **whether you have read a book on that topic in the previous year.**

[randomize, separate page]

1. BCRIME. Real life crime
2. BSCIFI. Science fiction
3. BMYSTERY. A mystery novel
4. BED. Education
5. BPOLGOV. Government or politics
6. BSPORT. Sports
7. BRELIGION. Religion (other than holy scripture text)
8. BINTL. International affairs
9. BBUSINESS. Business or finance
10. BSCIENCE. Scientific research or discoveries
11. BHISTORY. History
12. BWILDLIFE. Nature or wildlife

1. Conversation

We are also interested in knowing about the sorts of topics you discuss with family members, friends, or co-workers. We will list some conversation topics. Please indicate about **how often you discuss these topics with either friends, family members, or co-workers.** [Never; rarely; more than rarely but not often; often]

[randomize, separate page]

1. CCRIME. Crime
2. CED. Education
3. CPOLGOV. Government or politics
4. CSPORT. Sports
5. CRELIGION. Religion
6. CINTL. International affairs
7. CBUSINESS. Business or finance
8. CSCIENCE. Scientific research or discoveries
9. CTECHNOLOGY. New technologies
10. CENTERTAIN. Entertainment or celebrities

1. Social media

Some people share news stories with each other on social media (Facebook, Twitter, etc.).

SMa. How often do you share a news story on social media? [a. Never; b. rarely; c. more than rarely but not often; d. often]

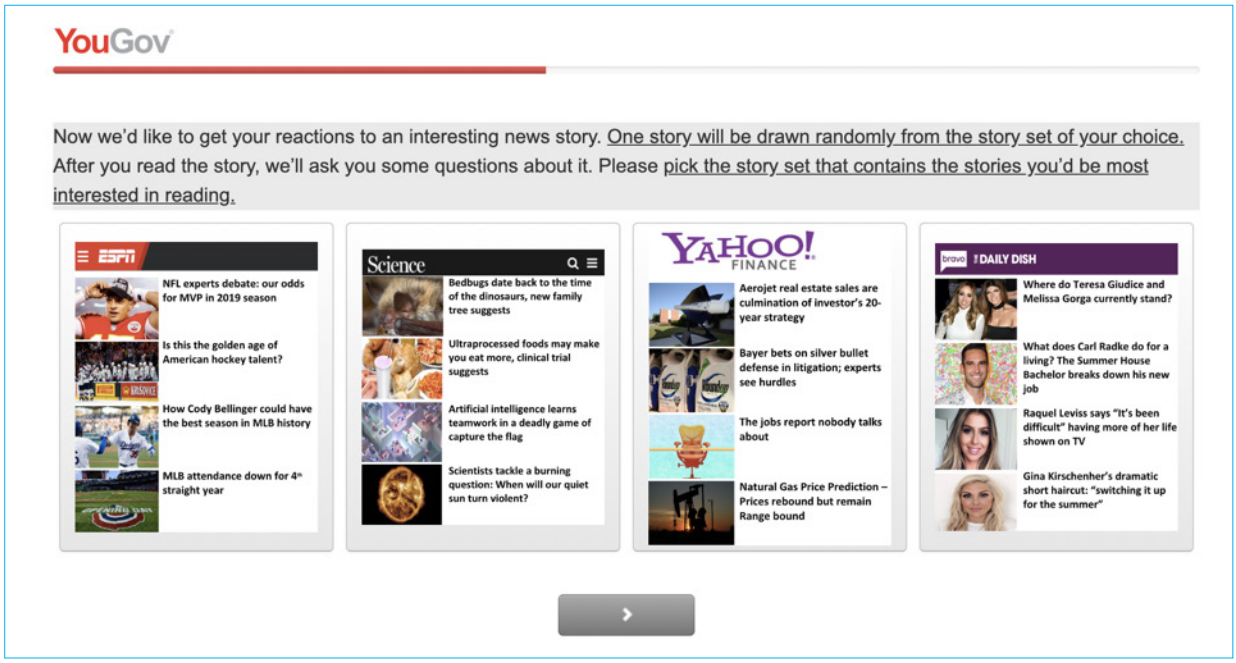
SMb [if b, c, or d on SMa.] We would like to know the kinds of news stories you are most likely to share on social media. Here is a list of topics. Please **rank the topics in order of how likely you are to share or post them on social media from most likely on top to least likely at the bottom. If the topic is one you would never bother to share, please leave it off your list.**

[randomize items on left; drag&drop or equivalent to rank in a numbered list on right]

1. smCRIME. Crime
2. smED. Education
3. smPOLGOV. Government or politics
4. smSPORT. Sports
5. smRELIGION. Religion
6. smINTL. International affairs
7. smBUSINESS. Business or finance
8. smSCIENCENEWS. Scientific research or discoveries
9. smTECHNOLOGY. New technologies
10. smENTERTAIN. Entertainment or celebrities

1. Reading Selection

Now we'd like to get your reactions to an interesting news story. **One story will be drawn randomly from the story set of your choice.** After you read the story, we'll ask you some questions about it. Please **pick the story set that contains the stories you'd be most interested in reading.**



[filler stories & related questions omitted]

1. Self-report

We will now list some topics that some people are interested in, and some people are not interested in. For each topic, please indicate **how interested** you are in that topic. [1 Not at all interested; 2 slightly interested; 3 more than slightly interested but not very interested; 4 very interested]

1. IPOLGOV. Government or politics
2. ISPORT. Sports
3. IRELIGION. Religion
4. ITRAVEL. Foreign travel
5. ISCIENCE. Scientific research or discoveries

6. ITECHNOLOGY. New technologies
7. IENTERTAIN. Entertainment or celebrities
8. INATURE. Nature
9. IMUSIC. Music

Deep Look Video Treatment

IDWATCH. As you know, companies and individuals produce video features on a wide variety of topics for people with a wide variety of interests. We will now select a video feature on a topic some people find interesting and others don't. It will be approximately 4 minutes long, and you can decide whether to view it or not and, if so, for how long.

[next page:]

The title to this video feature is [INSERT TITLE OF RANDOMLY SELECTED DL EPISODE].

To **skip** the video and proceed to the next part of the survey, click here.

To view the video, push the **"start"** button.

You can push the **"stop"** button at any time. If you wish to watch the entire 5-minutes, the video will turn itself "off."

[IF **VIDWATCH = 1**, RANDOMLY PRESENT RESPONDENTS WITH ONE OF THE FOUR ATTACHED LINKS. PLEASE COLLECT RESPONDENT-LEVEL METADATA ON TIME SPENT VIEWING THE VIDEO]

<Present Video Treatment >

[IF **VIDWATCH =0**]

We would like to hear more about why you chose not to watch this video. In the text box below, please tell us why you chose not to watch this video.

<Open-ended Essay Text Field>

Subjective Engagement Indicators

[IF **VIDWATCH = 1**]

Interesting. Overall how interesting did you find the feature? [[0 "Not at all interesting"; 1 "only slightly interesting"; 2 "Between slightly interesting and moderately interesting"; 3 "moderately interesting"; 4 "Between moderately interesting and very interesting"; 5 "Very interesting"]

Informative. Overall, how informative did you find the feature? [[0 "Not at all informative"; 1 "only slightly informative"; 2 "Between slightly informative and moderately informative"; 3 "moderately informative"; 4 "Between moderately informative and very informative"; 5 "Very informative"]

Surprising. Did you find the information in the feature surprising? [[0 "Not at all surprising"; 1 "only slightly surprising"; 2 "Between slightly surprising and moderately surprising"; 3 "moderately surprising"; 4 "Between moderately surprising and very surprising"; 5 "Very surprising"]

Attention. Overall how closely did you pay attention? [1. Not at all; 2. A little but not closely; 3. Closely but not very closely; 4. Very closely.]

documentary? [___ Yes ___ No]

Morequestions2. Like what? Please list as few or as many questions as you want. [open ended]

[only if yes to "Morequestions1"]

Morequestions3 [only if yes to "Morequestions1"].

How big a likelihood is there that you will try to find out the answers to any question or questions you had about the information in the documentary? [1. No likelihood at all. 2. Very small likelihood; 3. Small likelihood; 4. Modest likelihood; 5. High likelihood; 6. Very high likelihood]

Behavioral Engagement Indicators

[IF **VIDWATCH = 1**]

NOTE: PLEASE PRESENT ALL ITEMS IN THIS BATTERY IN RANDOM ORDER EXCEPT FOR "COMMENT" AND "COMMENT_TXT," WHICH SHOULD COME LAST.

SHARE. Would you be interested in **sharing** the video you just watched with friends and family?

If you answer YES to this question, we will provide you with an internet link to this video that you can share **after completing the survey.**

1. Yes
2. No,

EMAIL_SUBSCRIBE. The feature you watched came from a series called "*Deep Look*." The producers of the feature create and distribute videos like it on a regular basis. Would you be interested in receiving email updates when new features in the *Deep Look* series are available?

If you answer YES to this question, YouGov will provide *Deep Look* with your e-mail address. But be assured that your personal contact information, and your responses to the rest of this survey, are kept strictly confidential.

1. Yes
2. No

SM_SUBSCRIBE. *Deep Look* also posts updates on Facebook, Twitter, Instagram, and other social media platforms. Would you be interested in receiving such updates?

If you answer YES to this question, we will provide you with links you can use to subscribe to these updates after completing this survey.

1. Yes
2. No

YT_SUBSCRIBE. Additionally, *Deep Look* posts links to new videos on its YouTube page. Would you be interested in receiving such updates?

If you answer YES to this question, we will provide you with a link you can use to subscribe to these updates after completing this survey.

1. Yes
2. No

BEHIND_SCENES. *Deep Look* regularly produces short "behind the scenes" videos, featuring additional information about the creation of its episodes. Would you be interested in receiving a link to a "behind the scenes" video to watch on your own time, after completing this study?

1. Yes
2. No

COMMENT. As you know, producers of video features on the internet often give viewers an opportunity to comment on such features. Some viewers comment, others don't. The comments are often read by other people who have viewed or are thinking of viewing the feature.

We would like to know what comments, if any, you might make about the feature you just watched.

COMMENT_TXT. Please type any comments you have about this video in the text box below.

<Open-ended essay text entry>

Prior Exposure

[IF **VIDWATCH = 0 or 1**]

PRIOR_FAM. To the best of your knowledge, have you ever watched a video created by *Deep Look* in the past?
1. Yes
2. No

IF PRIOR_FAM = 1

PRIOR_FAM_CERT. How certain are you that you have watched a video created by *Deep Look* in the past?
1. Very certain
2. Just somewhat certain
3. Not very certain

PRIOR_SUB. To the best of your knowledge, do you already subscribe to e-mails, social media, or other updates from *Deep Look*?
1. Yes
2. No

IF PRIOR_SUB = 1

PRIOR_SUB_CERT. How certain are you that you already subscribe to e-mails, social media, or other updates from *Deep Look*?
1. Very certain
2. Just somewhat certain
3. Not very certain

Risk Perception Battery

INTRO. Next, we'd like to know your opinion on certain types of risks.

As individuals and as a society, we face a number of possible hazards. Some threaten people's health, safety, or financial well-being directly. Others threaten health, safety, or financial well-being indirectly through the damage they can impose on the environment or the economy. How much risk do you believe each of the following poses to human health, safety, or prosperity? [0 "no risk at all"; 1 "Very low risk"; 2 "Low risk"; 3 "Between low and moderate risk"; 4 "Moderate risk"; 5 "Between moderate and high risk"; 6 "High risk"; 7 "Very high risk"]

[RANDOMIZE ORDER; each separate page; horizontal radio buttons]

1. GWRISK. Global Warming
2. GUNRISK. Private Gun Ownership
3. NUKERISK. Nuclear Power
4. GMFRISK. Genetically modified food
5. AIRPOLLUTION. Air pollution
6. WASTE. Disposal of hazardous wastes in landfill sites
7. XRAY. Medical x-rays
8. FLOURIDE. Fluoridation of drinking water
9. CIGRISK. Exposure to second-hand cigarette smoke
10. MARYJRISK. Legalization of marijuana
11. PROSTRISK. Legalization of prostitution
12. FRACKING. "Fracking" (extraction of natural gas by hydraulic fracturing)
13. TAXRATE. Excessive government taxes on businesses
14. ENVREGRISK. Environmental regulations
15. SEXEDRISK. Teaching high school students about birth control
16. RADON. Airborne radon.

Ordinary Science Intelligence 2.0

[RANDOMIZE BLUE; Keep brown after immediately preceding blue]

INTRO. Finally, are also interested in what members of the public believe scientists have shown in other areas. Please respond to these statements and questions:

1. PEWGAS1. What gas do most scientists believe causes temperatures in the atmosphere to rise? Is it [hydrogen, helium, carbon dioxide, radon]
2. RADIOACTIVE. All radioactivity is man-made. (True/False)
3. LASERS. Lasers work by focusing sound waves. (True/False)
4. ELECTRONS. Electrons are smaller than atoms. (True/False)
5. PEWGAS2. Which gas makes up most of the Earth's atmosphere? [Hydrogen, Nitrogen, Carbon Dioxide, Oxygen]
6. EVOLUTION. Human beings, as we know them today, developed from earlier species of animals. (True/False)
7. COPERNICUS1. Does the Earth go around the Sun, or does the Sun go around the Earth? (Earth around Sun/Sun around the earth)
8. COPERNICUS2. [only if "earth/around sun" for Conpernicus1]: How long does it take for the Earth to go around the Sun? (1 day, 1 month, 1 year)
9. DADGENDER. It is the father's gene that decides whether the baby is a boy or a girl. (True/False)
10. ANTIBIOTICS. Antibiotics kill viruses as well as bacteria. (True/False)
11. VALID. Two scientists want to know if a certain drug is effective against high blood pressure. The first scientist wants to give the drug to 1,000 people with high blood pressure and see how many of them experience lower blood pressure levels. The second scientist wants to give the drug to 500 people with high blood pressure and not give the drug to another 500 people with high blood pressure, and see how many in both groups experience lower blood pressure levels. Which is the better way to test this drug? Why is it better to test the drug this way? [The first way/The second way]
12. A doctor tells a couple that their genetic makeup means that they've got one in four chances of having a child with an inherited illness.
    a. PROBABILITYa. Does this mean that if their first child has the illness, the next three will not? (Yes/No)
    b. PROBABILITYb. Does this mean that each of the couple's children will have the same risk of suffering from the illness? (Yes/No)

[In order:]

13. WARMER. From what you've read and heard, is there solid evidence that the average temperature on earth has been getting warmer over the past few decades, or not? [YES/NO]
14. WHYWARMER. [only if WARMER = YES]. Do you believe that the earth is getting warmer (a) mostly be cause of human activity such as burning fossil fuels or (b) mostly because of natural patterns in the earth's environment?

Intro. Here are some word problems that vary in difficulty. Please answer as many as you can.

15. DIE. Imagine that we roll a fair, six-sided die 1,000 times. Out of 1,000 rolls, how many times do you think the die would come up as an even number?
16. BUCKS. In the BIG BUCKS LOTTERY, the chances of winning a $10.00 prize are 1%. What is your best guess about how many people would win a $10.00 prize if 1,000 people each buy a single ticket from BIG BUCKS?
17. SWEEP. In the ACME PUBLISHING SWEEPSTAKES, the chance of winning a car is 1 in 1,000. What percent of tickets of ACME PUBLISHING SWEEPSTAKES win a car?
18. DISEASE1. If the chance of getting a disease is 20 out of 100, this would be the same as having a _____% chance of getting the disease.

19. DISEASE2. If the chance of getting a disease is 10%, how many people would be expected to get the disease out of 1000?
20. CONDPROB

Suppose your friend just had a mammogram. The doctor knows from previous studies that, of 100 women like her, 10 have tumors and 90 do not. Of the 10 who do have tumors, the mammogram correctly finds 9 with tumors and incorrectly says that 1 does not have a tumor. Of the 90 women without tumors, the mammogram correctly finds 80 without tumors and incorrectly says that 10 have tumors. The table below summarizes this information. Imagine that your friend tests positive (as if she had a tumor), what is the likelihood that she actually has a tumor?

	Tested positive	Tested negative	Totals
Actually has a tumor	9	1	10
Does not have a tumor	10	80	90
Totals	19	81	100

Answer: _____ out of _____

21. How much does the ball cost? _____cents
22. LILLYPAD. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? _____days
23. WIDGET. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? _____minutes

Post-Interview Script

Thank you for completing this survey!

IF SHARE=1
OR EMAIL_SUBSCRIBE=1
OR SM_SUBSCRIBE=1
OR BEHIND_SCENES=1

DISPLAY:

Below, please find the additional information you requested, based on responses you gave to questions asked earlier in this survey:

IF **SM_SUBSCRIBE=1**

Here are links to *Deep Look's* Facebook <link>, Twitter <link>, and Instagram <link> accounts.

IF **YT_SUBSCRIBE=1**

Here is a link to *Deep Look's* YouTube page <link>.

IF **BEHIND_SCENES=1**

Here is a link <link> to bonus "behind the scenes" footage about the making of *Deep Look* videos.

APPENDIX: SURVEY PARTICIPANT INFORMATION

All metrics reported below are based on 2,500 responses. Demographics provided are based on the categorizations YouGov collects (with the exception of age cohort, which we constructed based on reported birth years). Median and Mean age = 48 years old, Age range: 19-91.

DEMOGRAPHIC	CATEGORY	PERCENTAGE (%)	AGE IN 2019	FREQUENCY
Age Cohort	Silent	6	74+	156
	Boomers	33	55-73	833
	Generation X	26	39-54	645
	Millennials	30	23-38	742
	Generation Z	5	4-22	124
Gender	Male	47		1,168
	Female	53		1,332
Race	White	69		1,721
	Black	11		268
	Hispanic	13		313
	Asian	3		69
	Native American	1		27
	Mixed	2		61
	Other	2		39
	Middle Eastern	< 1		2
Education	No High School	5		118
	High School Graduate	31		763
	Some college	19		487
	2-year Associates degree	13		320
	4-year Bachelors degree	21		515
	Post-graduate training	12		297
Family Income	< 10K	6		154
	10-19.9K	8		210
	20-29.9K	10		254
	30-39.9K	11		275
	40-49.9K	8		207
	50-59.9K	9		217
	60-69.9K	7		168
	70-79.9K	6		150
	80-99.9K	8		190
	100-119.9K	5		134
	120-149.9K	4		100
	150-199.9K	3		79
	200-249.9K	1		19
	250-349.9K	1		18
	350-499.9K	< 1		4
	500K or more	< 1		12
	Prefer not to say	12		307