Citation for published version:
Gheorghiu, AI, Callan, MJ & Skylark, WJ 2020, 'A thin slice of science communication: Are people's evaluations of TED talks predicted by superficial impressions of the speakers?', *Social Psychological and Personality Science*, vol. 11, no. 1, pp. 117-125. https://doi.org/10.1177/1948550618810896

DOI:
10.1177/1948550618810896

Publication date:
2020

Document Version
Peer reviewed version

Link to publication

Gheorghiu, Ana I. ; Callan, Mitchell J. ; Skylark, William J./ A thin slice of science communication: Are people's evaluations of TED talks predicted by superficial impressions of the speakers?. In: Social Psychological and Personality Science. 2018 (Forthcoming). Copyright (C) 2018 Copyright Holder. Reprinted by permission of SAGE Publications.

University of Bath

Alternative formats

If you require this document in an alternative format, please contact:
openaccess@bath.ac.uk

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 18. Jan. 2021

A thin slice of science communication: Are people's evaluations of TED talks predicted by superficial impressions of the speakers?

Ana I. Gheorghiu*
Department of Psychology, University of Portsmouth, Portsmouth, UK
Department of Psychology, University of Essex, Colchester, UK

Mitchell J. Callan
Department of Psychology, University of Bath, Bath, UK

William J. Skylark
Department of Psychology, University of Cambridge, Cambridge, UK

In press, *Social Psychological and Personality Science*

Word Count: 4989

***Corresponding author:**
Ana I. Gheorghiu
Department of Psychology, University of Portsmouth
King Henry Building, King Henry I Street, Portsmouth, UK, PO1 2DY
ana.gherorghiu@port.ac.uk

Author Bios:

Ana I. Gheorghiu received her PhD from the University of Essex. She is currently a Lecturer in Psychology in the Department of Psychology at the University of Portsmouth.

Mitchell J. Callan received his PhD from the University of Calgary. He is a Professor in Psychology at the University of Bath.

William J. Skylark received his PhD from the University of Cambridge. He is a Senior Lecturer in Psychology at the University of Cambridge.

Note:

This work was supported by studentship ES/J500045/1 from the Economic and Social Research Council.

Abstract

First impressions based on physical characteristics and superficial information predict a wide variety of social judgments and outcomes. We build on recent work examining the effects of such impressions on the communication of scientific research and ideas to the general public. A large diverse sample viewed and evaluated scientific TED talks, while a separate group viewed short, silent excerpts of each video and judged the speakers on three core socio-cognitive traits: competence, morality, and sociability. Neither the perceived scientific quality nor the entertainment value of the talks was meaningfully predicted by the thin-slice judgments; likewise, they were independent of the speakers' age, gender, ethnicity, and attractiveness. We propose that these null results arise because the influence of superficial visual cues was overwhelmed by the wealth of more diagnostic information, and by our participants' attentiveness to this information. Our results suggest limits to the predictive power of superficial impressions.

Keywords: Science communication; Impression formation; Social cognition; Thin slices

A thin slice of science communication: Are people's evaluations of TED talks predicted by superficial impressions of the speakers?

Impressions formed from a person's face or brief glimpses of their expressive behaviour predict social outcomes in a variety of domains. In addition to a long line of research demonstrating that visible characteristics such as gender, ethnicity, and attractiveness bias many social judgments (Boring, 2017; Johnson & King, 2017; Zebrowitz & Mcdonald, 1991), researchers have examined the accuracy with which people can infer social traits and relationships from limited visual cues (Ambady & Gray, 2002; Ambady, Hallahan, & Conner, 1999; Carney, Colvin, & Hall, 2007), and how such inferences predict social outcomes. In a pioneering study, Ambady and Rosenthal (1993) found that assessments of teacher effectiveness based on 10-s silent video clips of instructors delivering their lectures positively predicted actual students' end-of-semester evaluations of the instructors. Similarly, Todorov and colleagues (2005) demonstrated that competence judgments based on campaign photographs predicted the outcomes of US congressional elections. These and other studies suggest that impressions based on facial appearance or "thin slices" of behavioural data are powerful predictors of social outcomes, although these impressions are often invalid (for reviews, see Ambady, Bernieri, & Richeson, 2000; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015; Uleman & Kressel, 2013).

The study of impression formation is important for several reasons. First, it helps to identify the fundamental dimensions on which people evaluate others and the structural basis for such judgments (e.g., Oosterhof & Todorov, 2008). In particular, appraisals based on superficial visual cues commonly map onto two or three putative core dimensions of social appraisal (e.g., Olivola & Todorov, 2010; Rule et al., 2010; Todorov, Said, Engell, & Oosterhof, 2008), which can conveniently be labelled "competence" (indicating efficacy, ability, and potency) and "warmth" (indicating likeability, kindness, and gregariousness) (Fiske, 2018), with recent evidence that the latter comprises distinct dimensions of "sociability" (an enthusiasm for social interaction) and "morality" (adherence to social norms and rules) (Brambilla & Leach, 2014; Brambilla, Rusconi, Sacchi, & Cherubini, 2011; Gheorghiu, Callan, & Skylark, 2017; Goodwin, 2015). Second, studies of superficial impressions establish the extent to which social outcomes are based on accessible but often invalid visual cues rather than careful consideration of relevant evidence (Todorov et al., 2015; Todorov & Porter, 2014); this may lead to the development of interventions that minimize bias in important decisions (see Todorov et al., 2015 for examples). Third, the study of thin-slice judgments illuminates the traits that are perceived as most relevant in particular domains by particular populations. For example, Rule et al. (2010) found that electoral outcomes were predicted by face-based judgments of "power" in the US, but by assessments of "warmth" in Japan.

The current work extends recent studies of impression formation in a new domain: science communication. Scientists are increasingly encouraged and required to communicate their research to non-expert audiences, including politicians, businesses, and the general public (Scheufele, 2014). Such communication encompasses traditional media and new, web-based formats (e.g., vlogs). The outcomes of such communication shape people's beliefs about the physical and social world; they also contribute to funding decisions and career success, thereby influencing what science "gets done", and by whom (Lok, 2010). However, although there is extensive research into scientist stereotypes (e.g., Schinske, Cardenas, & Kaliangara, 2015), and into the possibility of gender/ethnicity bias in publication and hiring (Ford, Brick, Blaufuss, & Dekens, 2018; Moss-Racusin, Dovidio, Brescoll, Graham, & Handelsman, 2012), there has been little study of the links between science communication outcomes and impressionistic assessments of core socio-cognitive traits.

One exploration of this issue was provided in a recent paper by Gheorghiu, Callan, and Skylark (2017). Using photos of real scientists, these authors found that interest in a scientist's work was positively predicted by their apparent morality, competence, and attractiveness; in contrast, the extent to which they looked like a "good scientist" (who does important, high-quality work) was positively related to apparent competence and morality, but negatively predicted by apparent sociability and attractiveness. When the faces of scientists were paired with real science news stories/titles, they biased the selection and evaluation of these communications. For example, research was judged more positively when putatively authored by a competent-looking researcher.

These results help illuminate the traits that shape different aspects of the communication process. However, they are limited in several ways. First, although Gheorghiu et al. (2017) used real scientists and real science news stories, the news stories came from third-party science websites and were experimentally paired with scientists' faces. Thus, we do not know whether thin-slice impressions predict the evaluations of ecologically-valid communications produced by the scientists themselves. Second, the studies used text-based communications, yet there is increasing emphasis on video-based science communication (for example, Vsauce, a science channel on www.youtube.com, had more than 13m subscribers in April 2018), which affords more cues from which to form an impression of the communicator. Notably, Gheorghiu et al. found that (static) facial appearance exerted a larger influence on people's selection of science news stories when they believed they would be watching a video rather than reading an article.

The present paper therefore investigates whether speaker characteristics and thin-slice judgments predict the evaluations of real-world video communications: TED talks (www.ted.com). TED talks are approximately 10-minute presentations of "ideas worth sharing", usually by a single speaker; many concern scientific topics. These videos are typically viewed several million times via the TED website and other platforms (e.g., YouTube). The number and popularity of TED talks, coupled with their relatively standardized format and production values, makes them ideal for an investigation of ecologically-valid science communication.

We showed one group of people short, silent video excerpts from TED talks and had them rate each speaker's apparent competence, morality, sociability, and attractiveness (Gheorghiu et al., 2017). These thin-slice stimuli are like those used in previous studies of communicator impression-formation (e.g., Ambady & Rosenthal, 1993) and capture the visual cues that a person might use when viewing and evaluating a talk; ratings based on these cues therefore capture people's superficial trait judgments, rather than their *first* impressions -- which might come from, for example, a photo of the speaker. A separate group of participants watched the full-length talks (complete with audio); each person saw five talks, and answered questions that probed their thoughts and feelings about each video. We then tested whether the communication outcomes (the evaluations based on the full videos) were predicted by the thin-slice-based impressions of the speakers and by basic speaker characteristics (age, gender, and ethnicity).

Based on the apparent ubiquity of impression effects, and the fact that people's reactions to science are often swayed by presentational factors (Eriksson, 2012; Harold, Lorenzoni, Shipley, & Coventry, 2016; Weisberg, Taylor, & Hopkins, 2015), we expected that the thin-slice judgments would predict the communication evaluations. However, we did not make strong predictions about the pattern of effects, for two reasons. First, the dimensional structure of the predictors and outcomes was itself a subject of inquiry: we were open-minded about whether morality and sociability constitute distinct traits, or whether they comprise a single "warmth" variable (and therefore correlate very highly); likewise, we were open about the factor structure of the communication-evaluation questions. It is hard to formulate firm predictions about the relations between constructs that are not fully specifiable in advance. The second reason why we did not make strong predictions is that prior work suggests a potentially complex pattern of relationships. For example, although it might be reasonable to suppose that apparent competence, sociability, and morality are positive traits that will lead to positive evaluations, these traits might differentially correlate with distinct outcomes. Sociability, for example, might positively predict whether a talk is seen as entertaining but be irrelevant to whether it is judged to contain high-quality scientific ideas (much as "warmth" is a relatively poor predictor of electoral success in the US; Rule et al, 2010; but see Olivola & Todorov, 2010). Indeed, sociability and attractiveness may conflict with the stereotypical conception of a scientist, and therefore have negative consequences for science communication (Gheorghiu et al., 2017). Similar ambiguity surrounds basic traits such as gender. On the one hand, female scientists are discriminated against in their working life (e.g., Ford et al., 2018);

on the other, women are stereotypically better communicators than men (e.g., Briton & Hall, 1995). The unique contribution of any trait (demographic or socio-cognitive) is even harder to anticipate with confidence when one considers that all such traits are usually correlated to varying degrees.

We therefore formulated a careful and pre-registered plan for data collection and analysis, but conducted the work in a spirit of open enquiry rather than seeking to test specific directional hypotheses. We asked: How are people's evaluations of science TED talks related to basic characteristics of the speakers (age, gender, ethnicity, attractiveness), and to judgments of core socio-cognitive traits (competence, morality, and sociability) made on the basis of short, silent video clips?

Method

The study was pre-registered on the Open Science Framework (osf.io/hgfap). In part 1, participants viewed short, silent extracts from TED talks (hereafter referred to as "thin-slice videos") and rated the speaker on core socio-cognitive dimensions; we refer to these judgments as trait ratings. In part 2, a separate sample of participants watched the original TED talk videos in their entirety and judged the quality of these scientific communications; we refer to these judgments as communication evaluations.

Sample size and exclusion rules

The number of videos was based on obtaining at least 80% power to detect a medium-sized correlation ($\rho = 0.3$) between the speaker's characteristics and evaluations of the TED talks (Faul, Erdfelder, Lang, & Buchner, 2007); the sample gave 80% power to detect a correlation of $r = .28$ or greater. (The availability of suitable TED videos and of participants to watch them meant that a larger sample of videos was not feasible.)

Participant samples sizes were chosen to ensure reliable estimates of the trait ratings for each speaker. For part 1, we aimed for a minimum of 16 participants per dimension. If the reliability (Cronbach's alpha) of the ratings for a given trait was below 0.7, we increased the sample size to 25. We replaced participants who had zero variance in their ratings ($n = 1$) or were not native English speakers ($n = 1$). For part 2, we aimed for a minimum of 15 participants per video, and replaced participants who recognized more than 40% of the videos ($n = 7$) or were under 18 years old ($n = 1$).

Participants

For part 1, the final sample comprised 73 students (47 women) from the University of Essex's psychology volunteers email list; each was paid £8. Ages ranged from 18 to 63 ($M_{age} = 23.3$,

SD_{age} = 8.5). Eighty-one percent of participants reported being British; all were native speakers of English. Their average science engagement score (on a scale from 1 to 7) was 4.74, SD = 0.88.

For part 2, the final sample comprised 300 students and staff members (204 women) recruited through the psychology volunteers and "small ads" staff email lists at the University of Essex, and community forums, message boards, and posters in the local town. Participants were paid £6 for their time, increasing to £8 when the sign-up rate decreased. Ages ranged from 18 to 72 (M_{age} = 24.3, SD_{age} = 12.4). 86% of participants reported being British, and all were native speakers of English. Their average science engagement score was 4.69, *SD* = 0.98.

Stimuli and Procedure

The stimuli were 100 TED talk videos from the science category of the TED talk website (https://www.ted.com/topics/science), edited to remove the "intro" and "outro" sequences (e.g., opening credits/TED logo), and any question-and-answer time with the speaker. We selected videos that were 5-14 minutes long (to minimise boredom), that contained no gruesome images, and whose speakers were unlikely to be well-known to a UK audience. We only chose videos which allowed us to obtain three 10-second video segments of the speaker, one from each third of the video. The gender (0 = male or 1 = female) and ethnicity (0 = white or 1 = non-white) of the speakers was coded by 2 independent raters with disagreements were resolved by a third rater (Table 1). The age of the TED speakers was coded by the same 2 independent raters, and averaged. We also recorded the length of each video (in seconds), and the age of the video (the time since the video was posted on the TED website, normalised to range from 0 for the newest video to 1 for the oldest).

Video Demographics	**Composition**	**Inter-rater agreement**
Gender	34 women, 66 men	1
Ethnicity	15 non-white, 85 white	0.93
Apparent age	*M* = 42.9, *SD* = 10.7	0.92

Table 1. Composition of the video sample in terms of gender, ethnicity and apparent age of the speaker. Inter-rater agreement is expressed by Kappa for gender and ethnicity, and the correlation (Pearson's *r*) between the two raters for age.

Part 1. To build thin-slice videos, we extracted three 10-second video segments, one from each third of the video; each clip comprised the first segment from that third in which the speaker appeared on their own, with no information or equipment in the background to suggest the topic of their talk. The segments were joined in chronological order, giving 30 seconds of silent footage of the speaker, with a one-second blank between each segment.

Participants were told "You will be asked to view short, silent videos of people giving a TED talk, and to indicate your assessment of the person in the video on a social dimension." (Full instructions are provided in the Supplementary Materials.) Each participant saw all 100 thin-slice videos, and rated each speaker on either competence, sociability, morality, or physical attractiveness (e.g., "How COMPETENT is this person?"; Gheorghiu et al., 2017). Participants were tested in individual cubicles with a break half-way through the session. Judgements were between 1 (Not at All) and 9 (Extremely); participants pressed "r" instead of providing a rating if they recognised the speaker. The video order was randomized for each participant, and the allocation of participants to traits was counterbalanced.

Part 2. Participants were told: "You will be asked to view videos of people giving a talk, and to indicate your assessment of the video on several dimensions." Each participant saw 5 full-length videos (excluding the intro, outro, and question time), and answered 7 questions after each video (Table 2). Participants were tested in individual cubicles, and listened to the audio through identical over-the-ear headphones. Responses were between 1 (Not at All) and 9 (Extremely) for Q1-Q6 and Yes/No for Q7. The allocation of participants to one of 20 pre-determined sets of five videos was counterbalanced; the order of the videos in the sets was always the same. Participants' reaction times were recorded in both parts.

	Question
Q1	How would you rate the overall quality of this scientist's research?
Q2	How good were this scientist's ideas?
Q3	How easy to follow/comprehend was this scientist's presentation?
Q4	How engaging/entertaining was this scientist's presentation?
Q5	How much would you recommend that we show this video to the next group of participants?
Q6	How likely would you be to share this video on any social media platform?
Q7	Have you heard/read about this research before (not just this general topic, but this specific piece of research) or seen this scientist before?

Table 2. The set of questions asked after each video.

At the beginning of both stages, participants indicated their age, gender, nationality, and first language, and completed a questionnaire probing their engagement with science (Gheorghiu et al., 2017).

The experiment was run using PsychoPy2 v1.84.2 (Peirce, 2007) on 21.5 inch LCD-screen computers (1920 x 1080 pixels). The videos were in mp4 format, 1280 x 720 pixels, 30 frames/second. All stimuli are available upon request.

Results

The data are available from the Open Science Framework (osf.io/a4jzr). Throughout, trials on which the participant recognised the video were excluded (135/7300 judgments in part 1; 116/10500 judgments in part 2).

Dimension	Sample size	Cronbach's alpha
Attractiveness	16	0.96
Competence	16	0.77
Sociability	16	0.91
Morality	25	0.69

Table 3. Cronbach's alpha values for the core social dimensions judged on the basis of thin-slice videos in part 1. After testing 16 participants, Morality had an alpha value below 0.7 (α = 0.61), so as per our pre-registered strategy the sample size was increased to 25; the final reliability is at a level conventionally labelled "questionable".

Table 3 lists the reliabilities of the trait ratings from part 1. To examine the structure of the communication evaluations made in part 2, the 6 evaluative judgments were submitted to an exploratory factor analysis (Supplementary Materials). (Here and throughout, analyses were "by item", meaning that we averaged across participants to get a single value of each variable for each video). A Principal Axis Factoring extraction with Direct Oblimin rotation suggested a 2-factor solution, which was supported by visual inspection of the scree plot. The items loading onto each factor were averaged to form two composite measures, "Research Quality" (questions 1 and 2, tapping overall quality and quality of ideas; Cronbach's alpha = 0.91), and "Entertainment Value" (questions 3-6, tapping quality of presentation and likelihood of sharing with others; Cronbach's alpha = 0.94). The mean Research Quality and Entertainment Value judgments are quite high, but the distributions of both measures have reasonable range and variance and were not bunched at the top of the range, with no indication of ceiling effects.

Table 4 gives the zero-order correlation matrix for the predictors and outcome variables. Competence, sociability and morality were significantly correlated, but with small-to-medium strength relationships; they were therefore treated as distinct variables. Neither video age nor video length correlated significantly with either research quality or entertainment value. In accordance with our registered analysis plan, these variables were therefore not considered further.

	Attract	**Comp**	**Mor**	**Soc**	**Age**	**Gen**	**Ethn**	**Vid Len**	**Vid Age**	**Res Qual**	**Ent Value**
Attract	3.71 (1.44)	.36*	.44*	.26*	-.64*	.46*	.21*	-.04	-.23*	.15	.01

Comp		5.80 (0.87)	.35*	.47*	-.01	.02	.12	.07	-.15	-.05	.02
Mor			5.87 (0.47)	.37*	-.26*	.25*	.48*	-.07	-.22*	.05	.03
Soc				5.64 (1.03)	-.26*	-.01	.18	-.03	-.05	-.06	-.04
Age					42.88 (10.67)	-.20*	-.32*	.07	.35*	-.18	-.11
Gen						0.34 (0.48)	.05	-.07	-.16	.11	.02
Ethn							0.15 (0.36)	-.07	-.12	.06	.14
Vid Len								603.94 (148.15)	-.25*	-.05	-.08
Vid Age									0.35 (0.23)	-.17	-.05
Res Qual										7.14 (0.77)	.55*
Ent Value											6.24 (0.95)

Table 4. Pearson's correlations between predictors and outcomes (* p< .05). The long diagonal shows the mean (standard deviation) for each measure, where Attract = Attractiveness, Comp = Competence, Mor = Morality, Soc = Sociability, Age = Apparent age; Gen = Gender, Ethn = Ethnicity, Vid Len = Video length (seconds), Vid Age = Video age (normalised), Res Qual = Research Quality, Ent Value = Entertainment Value.

To test whether people's evaluations of science TED talks are predicted by trait judgments made on the basis of thin-slice videos, we conducted a multivariate regression with research quality and entertainment value as the outcomes, and the speaker's perceived attractiveness, competence, sociability, morality, age, gender and ethnicity as the predictors (standardised). A Type III Manova (conducted using the "car" package for R; Fox & Weisberg, 2011) revealed a significant overall model: Pillai's $\text{Trace}_{\text{Intercept}}$ = 0.99, $F(2, 91) = 4291.1$, $p < .001$. However, no individual effect of the predictors on the two outcomes was significant (Table 5).

Predictors	Pillai's Trace	F-value (2, 91)	p-value
Attractiveness	0.026	1.198	.307
Competence	0.018	0.834	.437
Morality	0.001	0.041	.960
Sociability	0.011	0.523	.594
Apparent age	0.020	0.951	.390
Gender	0.003	0.138	.871
Ethnicity	0.013	0.580	.562

Table 5. Pillai's Trace, F- and p-values for a type III multivariate regression, regressing research quality and entertainment value onto the scientist's attractiveness, competence, morality, sociability, age, gender and ethnicity.

Our results suggest that impressions formed from thin-slices of the videos do not predict the perceived quality of the research and the entertainment value of the full length talk. However, participants in Part 1 had slightly different contextual information from those in Part 2, because the former were not told that the speakers were giving a science talk. We therefore ran a supplementary study that replicated the competence rating task with a fresh sample of participants who were aware that the speakers were giving a science TED talk. The results were very similar to those of the main study, including no meaningful relationship between apparent competence and the outcome variables (Supplementary Materials).

Exploratory Analyses

In an exploratory analysis, we computed Bayesian credible intervals and Bayes factors for the correlations between predictors and outcome (Table 4) using JASP's default prior (JASP Team, 2018). All of the credible intervals span zero (Table 6) and the Bayes factors are all between 1 and 10, indicating that the data favour the null hypothesis in each case -- by factors ranging from 1.76:1 to 7.99:1.

Predictors	Research Quality				Entertainment Value			
	r	95% CIs	p-value	BF_{01}	r	95% CIs	p-value	BF_{01}
Attractiveness	.15	-.05, .33	.141	2.75	.01	-.19, .20	.960	7.99
Competence	-.05	-.24, .15	.659	7.27	.02	-.18, .21	.867	7.89
Morality	.05	-.14, .25	.591	6.94	.03	-.17, .22	.774	7.68
Sociability	-.06	-.25, .14	.583	6.89	-.04	-.23, .16	.719	7.50
Apparent age	-.18	-.36, .02	.079	1.76	-.11	-.29, .09	.301	4.72
Gender	.11	-.09, .29	.299	4.70	.02	-.17, .22	.825	7.81
Ethnicity	.06	-.13, .25	.539	6.64	.14	-.06, .32	.179	3.28
Video length	-.05	-.24, .15	.658	7.26	-.08	-.27, .11	.408	5.71
Video age	-.17	-.35, .03	.086	1.88	-.05	-.24, .15	.641	7.18

Table 6. Pearson's r, lower and upper credible intervals, p-values and Bayes' factors for the correlations between thin-slice predictors and full talk outcomes.

We also explored whether our ratings correlated with an ecological index of talk popularity. For each talk we recorded the number of views listed on the TED web page as at April 11th 2018. This is an imperfect index: the talks can be viewed on other platforms, and many extraneous factors drive the viewing figures. Nonetheless, Table 7 reports the correlation coefficients for the relationship between views (after $\log_{10}$ transformation to correct for skew) and the other variables. None of the Bayes factors indicate convincing evidence against the null, and in many cases there is moderate support for the hypothesis of no association. Thus, a crude ecological measure mirrors our laboratory findings.

	r	95% CIs	p-value	BF_{01}
Attractiveness	-0.12	-0.30, 0.08	0.244	4.10
Competence	-0.03	-0.22, 0.17	0.796	7.74
Morality	-0.08	-0.27, 0.12	0.449	6.03
Sociability	-0.05	-0.24, 0.15	0.640	7.18
Apparent age	-0.01	-0.20, 0.19	0.928	7.97
Gender	-0.20	-0.37, 0.00	0.049	1.20
Ethnicity	-0.11	-0.30, 0.09	0.274	4.43
Video length	0.07	-0.12, 0.26	0.466	6.16
Video age	-0.22	-0.40, -0.03	0.027	0.71
Entertainment Value	0.16	-0.04, 0.34	0.106	2.20
Research Quality	0.11	-0.09, 0.29	0.291	4.62

Table 7. Pearson's r, lower and upper credible intervals, p-values and Bayes' factors for the correlations between number of talk views and the other variables.

Overall, our results suggest that first impressions based on thin-slices of TED talks do not predict the perceived research quality or entertainment value of the full length TED talks.

Discussion

Our experiment produced 3 findings. First, we found further evidence that competence, sociability, and morality comprise distinct traits. In contrast to studies that have grouped the latter two dimensions into a single "warmth" construct, we found similar small-to-medium correlations between all three traits. This adds to theoretical and empirical work suggesting that morality is a distinct social attribution (Brambilla & Leach, 2014; Brambilla et al., 2011; Gheorghiu et al., 2017; Goodwin, 2015; Goodwin, Piazza, & Rozin, 2014).

Second, our data indicate that people differentiate between the scientific quality of a researcher's work and the entertainment value of their communications. The dimensions on which members of the public evaluate science communications have received little attention from psychologists, and it will be important to conduct future work with a wider variety of evaluative judgment questions and behavioural outcome measures (e.g., whether the viewer actually forwards a link to the video to a friend), because the validity of our outcome variables has not yet been established. Nonetheless, the dissociation between entertainment value and research quality accords with recent evidence for the same distinction reported by Gheorghiu et al. (2017), and implies that the research which garners the public's interest or approval may not be that which they judge to be of the highest quality, with corresponding implications for the use of social media sharing ("altmetrics") or public engagement as metrics against with which scientists are evaluated (Barnes, 2015).

Our third and most important finding was of very little evidence that evaluations of TED talks are predicted by superficial characteristics of the speakers. Neither our multivariate analysis nor simple correlations found a meaningful association between any of the predictor variables and either of the two outcome measures. As is always the case with null hypothesis significance testing, a larger sample might have shrunk the confidence intervals on the measured effects sufficiently for them to exclude zero. However, our sample of video clips gave reasonably high power, and our post hoc Bayesian analyses indicate that the data provide non-trivial evidence in favour of the null for most of the predictor-outcome combinations under consideration, and never provide support for the alternative hypothesis. Likewise, although the null results involving the morality predictor might reflect the relatively low reliability of this variable, all of the other predictor and outcome variables showed very good consistency and reliability. In short, we can be reasonably confident that relationships between the predictor variables and the outcome measures are, at best, small.

What are the implications of this finding? Considering first the demographic variables, the fact that both the perceived scientific quality and the entertainment value of the communications were independent of the gender, ethnicity, and apparent age of the communicators is surprising, because there is high-profile evidence for gender discrimination in science (e.g., Moss-Racusin et al., 2012). However, our results accord with recent research suggesting that the evidence for bias may be weaker than is often assumed (Williams & Ceci, 2015), and that the considered opinions of a reasonably engaged viewing public may be relatively free from the influence of gender, age, or racial stereotypes (Gheorghiu et al., 2017). Notably, there was a pronounced gender imbalance among the speakers themselves (with males outnumbering females by approximately two to one), and the analysis of viewing figures indicate slightly lower success for females, implying that there may be bias or self-selection elsewhere in the communication process. Likewise, our video sample size necessitated a crude treatment of ethnicity (as white vs non-white) which may have obscured inter-racial differences.

Turning to the predictive power of socio-cognitive trait attributions made on the basis of thin slices of behaviour: the uniformly null results were unexpected. We were open-minded about which traits would predict which outcomes, and in what direction, but the wide literature on impression formation and on biases in science and communication led us to expect non-trivial consequences of superficial impressions on people's evaluations of these talks. One possible explanation is that the predictor constructs are simply irrelevant to the measured outcomes. That is, when participants viewed TED talks in their entirety, their evaluations of the scientific content and entertainment value of the videos may have been independent of their assessments of the speakers' competence, morality, sociability, and attractiveness. This seems unlikely because these traits are widely regarded as foundational attributions in social judgments (Feingold, 1992; Fiske, 2018; Goodwin, 2015; Wojciszke, 2005), and because some of the predictors conceptually overlap with some of the outcomes. For example, competence encompasses intelligence, skill, and organization (Fiske, Cuddy, & Glick, 2007), which are central to scientific ability. Indeed, Gheorghiu et al. (2017) found strong correlations between face-based assessments of competence and judgments of whether the person was a "good scientist". Our preferred explanation is therefore that the predictor traits *are* relevant to the outcome variables but the assessment of those traits from thin-slices is largely independent of the assessment based on viewing the full video. In other words, people's superficial impressions of a TED speaker's competence, sociability, morality, and attractiveness may be overwhelmed by other sources of information about these traits.

Why would this happen for the science communication outcomes studied here, when judgments of competence, warmth, and attractiveness made on the basis of facial appearance have been found to predict electoral, legal, and financial outcomes? One possibility is that publication bias has obscured null results in those domains. Another is that our motivated participant sample and controlled testing conditions meant that participants attended closely to the content of the talks – perhaps more than some voters in a congressional election, for example, who may not engage much with campaign materials beyond the photograph of the candidate; likewise, students providing teacher evaluations (which are often influenced by attractiveness and can be predicted from thin-slice evaluations; Ambady & Rosenthal, 1993; Talamas, Mavor, & Perrett, 2016) may lack the motivation to consider all relevant factors, or be responding in noisy environments where only the most salient social cues are considered. A related possibility is that the *scientific* nature of the communication encourages a more deliberative approach to evaluating the talk; society tolerates and sometimes encourages going with one's "gut reaction" about a politician or a financial decision (e.g., Kandasamy et al., 2016), but a central feature of scientific work is that it should be considered dispassionately (Shapin, 1996). Notably, although Gheorghiu et al. (2017) found that people rated science news stories more favourably when they were purportedly written by a more competent-looking scientist, the effect was quite small, despite this study contrasting some of the lowest- and highest-scoring faces on the competence dimension. A 10-minute TED talk provides even more information than does a written article, and this information, when presented to participants who are motivated and able to process it fully, may drown out the effects of superficial visual cues – although our exploratory analysis of viewing figures suggest that the thin-slice cues may not be predictive even under noisier viewing conditions.

In short, while we cannot disentangle the various possible explanations for our null results, we tentatively attribute our null results to wealth of information available when people view full talks but which was not accessible to people who viewed short, silent clips. However, this conclusion is speculative, and our results are tempered by several caveats and methodological limitations. Although we recruited participants from both inside and outside our university, we recruited more females than males, and our participants typically had above-average education and may well have had higher than usual interest in scientific research. In mitigation, the latter is also likely to be true of "real world" viewers of TED talks, but nonetheless we cannot tell how far our results generalize to that population or to other groups.

Similarly, our procedures do not perfectly capture ecological viewing conditions and behavioural outcomes: we had participants watch 5 videos in an individual testing cubicle and evaluate them with rating scales, rather than (for example) viewing a single talk on a home computer, perhaps with friends, and deciding how much to watch, whether to share it, or what comments to type on a social media page. Our controlled, lab-based approach was driven by the aim of exploring the factor structure of talk-evaluations (which requires judgments on multiple dimensions) and by our interest in the processes by which people evaluate scientific talks when they are engaged with them and exposed to the entirety of the presentation. Our exploratory analysis of real-world viewing data produced the same null results as our lab studies, but we offer this finding with caution: TED talks are hosted on many different platforms, so it is impossible to obtain total viewing figures, and viewing numbers are shaped by extraneous variables (e.g., positive feedback loops; Salganik, Dodds, & Watts, 2015) and may not reflect the perceived quality of the talk.

A final limitation is that our approach was correlational; had we found a relationship between thin-slice judgments and communication outcomes, it might have been because (for example), more competent-looking people really do have better scientific ideas. (For this to explain our results, one would have to assume that thin-slice judgments are negatively correlated with the quality of a talk's content and structure, and that people who view complete talks integrate both sources of information when making their judgments.) In future it would be good to have a new group of raters evaluate the talk transcripts in order to examine whether thin-slice evaluations predict the quality of a talk's content and structure when superficial cues have been stripped away completely. Likewise, an audio-only rating task would allow us to assess the contribution of vocal cues to impressions and evaluations.

The present work provides an important first step towards understanding the social cognition of video-based science communication. Our results suggest that people's evaluations of such communications are relatively independent of their physical characteristics and of attributions made on the basis of thin-slices of their behaviour. This is potentially heartening news given the importance of effective science communication to the development of informed and engaged citizenry, and to the careers of scientists, but it will be important to explore the generality of our findings to other stimuli, populations, and viewing conditions.

References

Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream. *Advances in Experimental Social Psychology*, *32*, 201–271. https://doi.org/10.1016/S0065-2601(00)80006-4

Ambady, N., & Gray, H. M. (2002). On being sad and mistaken: Mood effects on the accuracy of thin-slice judgments. *Journal of Personality and Social Psychology*, *83*(4), 947–961. https://doi.org/10.1037//0022-3514.83.4.947

Ambady, N., Hallahan, M., & Conner, B. (1999). Accuracy of judgments of sexual orientation from thin slice behavior. *Journal of Personality and Social Psychology*, *77*(5), 538–547.

Ambady, N., & Rosenthal, R. (1993). Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, *64*(3), 431–441. https://doi.org/10.1037/0022-3514.64.3.431

Barnes, C. (2015). The use of altmetrics as a tool for measuring research impact. *Australian Academic and Research Libraries*, *46*(2), 121–134. https://doi.org/10.1080/00048623.2014.1003174

Boring, A. (2017). Gender biases in student evaluations of teaching. *Journal of Public Economics*, *145*, 27–41. https://doi.org/10.1016/j.jpubeco.2016.11.006

Brambilla, M., & Leach, C. W. (2014). On the importance of being moral: The distinctive role of morality in social judgment. *Social Cognition*, *32*(4), 397–408. https://doi.org/10.1521/soco.2014.32.4.397

Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for honesty: The primary role of morality (vs. sociability and competence) in information gathering. *European Journal of Social Psychology*, *41*(2), 135–143. https://doi.org/10.1002/ejsp.744

Briton, N. J., & Hall, J. A. (1995). Beliefs about female and male nonverbal communication. *Sex Roles*, *32*(1–2), 79–90. https://doi.org/10.1007/BF01544758

Carney, D. R., Colvin, C. R., & Hall, J. A. (2007). A thin slice perspective on the accuracy of first impressions. *Journal of Research in Personality*, *41*(5), 1054–1072. https://doi.org/10.1016/j.jrp.2007.01.004

Eriksson, K. (2012). The nonsense math effect. *Judgment and Decision Making*, *7*(6), 746–749. https://doi.org/10.1073/pnas.1216778110.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin*, *111*(2), 304–341. https://doi.org/10.1037/0033-2909.111.2.304

Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science*, *27*(2), 67–73. https://doi.org/10.1177/0963721417738825

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Ford, H., Brick, C., Blaufuss, K., & Dekens, P. . (2018). Gender inequity in speaking opportunities at the American Geophysical Union Fall Meeting. *Nature Human Behavior*.

Gheorghiu, A. I., Callan, M. J., & Skylark, W. J. (2017). Facial appearance affects science communication. *Proceedings of the National Academy of Sciences*, *114*(23), 5970–5975. https://doi.org/10.1073/pnas.1620542114

Goodwin, G. P. (2015). Moral character in person perception. *Current Directions in Psychological Science*, *24*(1), 38–44. https://doi.org/10.1177/0963721414550709

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, *106*(1), 148–168. https://doi.org/10.1037/a0034726

Harold, J., Lorenzoni, I., Shipley, T. F., & Coventry, K. R. (2016). Cognitive and psychological science insights to improve climate change data visualization. *Nature Climate Change*, *6*(12), 1080–1089. https://doi.org/10.1038/nclimate3162

JASP_Team. (2018). JASP version 0.8.6. Retrieved from https://jasp-stats.org

Johnson, B. D., & King, R. D. (2017). Facial profiling: Race, physical appearance, and punishment*. *Criminology*, *55*(3), 520–547. https://doi.org/10.1111/1745-9125.12143

Kandasamy, N., Garfinkel, S. N., Page, L., Hardy, B., Critchley, H. D., Gurnell, M., & Coates, J. M. (2016). Interoceptive ability predicts survival on a london trading floor. *Scientific Reports*, *6*, Article 32986 1-7. https://doi.org/10.1038/srep32986

Lok, C. (2010). Science funding: Science for the masses. *Nature*, *465*(7297), 416–418.

https://doi.org/10.1038/465416a

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474–16479. https://doi.org/10.1073/pnas.1211286109

Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*, *34*(2), 83–110. https://doi.org/10.1007/s10919-009-0082-1

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Rule, N. O., Ambady, N., Adams, R. B., Ozono, H., Nakashima, S., Yoshikawa, S., & Watabe, M. (2010). Polling the face: Prediction and consensus across cultures. *Journal of Personality and Social Psychology*, *98*(1), 1–15. https://doi.org/10.1037/a0017673

Salganik, M. J., Dodds, P. S., & Watts, D. (2015). Experimental study of inequality and unpredicatibility in an artificial culture market. *Science*, *299*(5614), 1874–1877. https://doi.org/10.1126/science.1121066

Scheufele, D. A. (2014). Science communication as political communication. *Proceedings of the National Academy of Sciences*, *111*, 13585–13592. https://doi.org/10.1073/pnas.1317516111

Schinske, J., Cardenas, M., & Kaliangara, J. (2015). Uncovering scientist stereotypes and their relationships with student race and student success in a diverse, community college setting. *CBE Life Sciences Education*, *14*(3), 1–16. https://doi.org/10.1187/cbe.14-12-0231

Shapin, S. (1996). *The scientific revolution*. Chicago: University of Chicago Press.

Talamas, S. N., Mavor, K. I., & Perrett, D. I. (2016). Blinded by beauty: Attractiveness bias and accurate perceptions of academic performance. *PLoS ONE*, *11*(2), e0148284 1-18. https://doi.org/10.1371/journal.pone.0148284

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Psychology: Inferences of competence from faces predict election outcomes. *Science*, *308*, 1623–1626. https://doi.org/10.1126/science.1110589

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of*

Psychology, *66*(1), 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

Todorov, A., & Porter, J. M. (2014). Misleading first impressions: Different for different facial images of the same person. *Psychological Science*, *25*(7), 1404–1417. https://doi.org/10.1177/0956797614532474

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, *12*(12), 455–460. https://doi.org/10.1016/j.tics.2008.10.001

Uleman, J., & Kressel, L. M. (2013). A brief history of theory and research on impression formation. In D. E. Carlston (Ed.), *Oxford handbook of social cognition* (pp. 53–73). New York: Oxford University Press.

Weisberg, D. S., Taylor, J. C. V., & Hopkins, E. J. (2015). Deconstructing the seductive allure of neuroscience explanations. *Judgment and Decision Making*, *10*(5), 429–441.

Williams, W. M., & Ceci, S. J. (2015). National hiring experiments reveal 2:1 faculty preference for women on STEM tenure track. *Proceedings of the National Academy of Sciences*, *112*(17), 5360–5365. https://doi.org/10.1073/pnas.1418878112

Wojciszke, B. (2005). Affective concomitants of information on morality and competence. *European Psychologist*, *10*(1), 60–70. https://doi.org/10.1027/1016-9040.10.1.60

Zebrowitz, L. A., & Mcdonald, S. M. (1991). The impact of litigants' babyfacedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior*, *15*(6), 603–623.

SUPPLEMENTARY MATERIALS

Contents

Factor Analysis

We ran a Principal Axis Factoring extraction with Direct Oblimin rotation on the 6 judgements from part 2 (Table 2). The correlation matrix between the 6 questions indicated significant, medium-to-large relationships between Q1 and Q2 (tapping into research quality), and between Q3 - Q6 (presentation quality and likelihood of dissemination), as shown in Table S1.

	Q2	Q3	Q4	Q5	Q6
Q1	.84*	.39*	.37*	.54*	.44*
Q2		.50*	.46*	.66*	.57*
Q3			.83*	.81*	.66*
Q4				.85*	.74*
Q5					.84*

Table S1. Correlations between the items forming each factor (* indicates p <.05).

One factor (Research Quality, formed of Q1 and Q2, initial Eigenvalue of 1.06) accounted for 14.87% of the shared variance, while the other factor (Entertainment Value, formed of Q3-Q6, initial Eigenvalue of 4.20) accounted for 67.05% of the shared variance, cumulating in 81.92% of the variance explained. The rotated solution had a reasonably simple structure, with Q1 and Q2 loading more strongly on Research Quality, and Q3-Q6 loading more strongly on Entertainment Value (highlighted in Table S2).

	Research Quality	**Entertainment Value**
Q1	**.863**	-.015
Q2	**.953**	.054
Q3	-.025	**.871**
Q4	-.126	**1.002**
Q5	.162	**.863**
Q6	.121	**.749**

Table S2. Factor loadings following the PFA extraction with Direct Oblimin rotation from the pattern matrix.

The two factors were significantly, but moderately correlated (Pearson's $r = .55$, $p < .001$).

Supplementary Study

Method

The study followed the methodology of the Main Study, part 1: participants viewed thin-slices of TED talks and rated the speaker on their competence. The main difference is the participants were informed they are watching a science TED talk, as per the suggestions we received from one of the reviewers.

Open practices

The study was pre-registered on the Open Science Framework (osf.io/cy7xz). The data are available from the Open Science Framework (osf.io/a4jzr).

Sample size and exclusion rules

The same sample size and exclusion rules as for the Main Study Part 1 were followed: we aimed for a minimum of 16 participants, since we collected ratings only for competence. If the reliability (Cronbach's alpha) of the competence ratings was low (below 0.7), we increased the sample size to 25. We replaced participants who were not native English speakers (n = 1). No participants had to be replaced for having no variance in their ratings, or for recognising more than 40% of the videos.

Participants

We recruited 16 participants (12 women) from the University of Bath's psychology volunteers pool and staff; each was paid £15. Ages ranged from 20 to 54 (M_{age} = 25.4, SD_{age} = 8.1). All participants reported being British and native speakers of English.

Stimuli and Procedure

We used the same stimuli and procedure as in the Main Study, part 1.

Results

Trials on which the participant recognised the video were excluded on a case-by-case basis (14/1600 judgements excluded).

Mean competence was computed for each thin-slice video; competence had a reliability of Cronbach's Alpha = 0.79, so no additional data were collected.

We computed zero-order correlations, Bayesian credible intervals, and Bayes factors (using JASP's default prior; JASP Team, 2018) for the correlations between the new competence ratings and the previous predictors and outcome variables (Table S3). The credible intervals for the correlation between new competence ratings and both outcome variables span zero, and the Bayes factors (7.97 and 7.88) indicate that the data favour the null hypothesis.

Predictors/Outcomes	***r***	**95% CIs**	***p*-value**	**BF_{01}**
Attractiveness	.15	-.05, .33	.132	2.62
Competence	.75	.64, .82	<.001	<.001
Morality	.25	.06, .42	.011	0.33
Sociability	.55	.39, .67	<.001	<.001
Apparent age	.22	.03, .40	.026	0.70
Gender	-.08	-.27, .11	.407	5.70
Ethnicity	-.12	-.30, .08	.250	4.16
Video Length	.04	-.16, .23	.701	7.44
Video Age	-.01	-.20, .19	.956	7.99
Research Quality	-.01	-.20, .19	.934	7.97
Entertainment value	-.02	-.21, .18	.859	7.88

Table S3. Pearson's *r*, lower and upper credible intervals, *p*-values and Bayes factors for the correlations between new competence ratings ($M_{NewComp}$ = 6.26, $SD_{NewComp}$ = 0.70), and the predictors (including the previous competence judgments) and outcomes from the main study.

To test whether people's evaluations of science TED talks are predicted by trait judgments made on the basis of thin-slice videos when people know they are watching a science TED talk, we conducted a multivariate regression with research quality and entertainment value as the outcomes, and the speaker's perceived competence (new ratings), apparent age, gender and ethnicity as the predictors (standardised). Perceived attractiveness, morality, and sociability were not included in the analysis because the judges who made these ratings (in the main study) were not aware they were watching a science TED talk. A Type III MANOVA (conducted using the "car" package for R; Fox & Weisberg, 2011) revealed a significant overall model: Pillai's $Trace_{Intercept}$ = 0.99, $F(2, 94) = 4385$, p <.001. However, no individual effect of the predictors on the two outcomes was statistically significant (Table S4).

Predictors	**Pillai's Trace**	***F* (2, 91)**	***p***
New Competence	0.001	0.065	.938
Apparent age	0.024	1.161	.318
Gender	0.008	0.353	.704
Ethnicity	0.016	0.743	.479

Table S4. Pillai's Trace, *F*- and *p*-values for a Type III multivariate regression, regressing research quality and entertainment value onto the scientist's competence, apparent age, gender, and ethnicity.

Exploratory analysis

In line with our analysis of the main study, we explored whether the new competence ratings correlated the number of views listed on the TED web page as of April 11th 2018. The correlation between the new competence ratings and the number of views (after a $\log_{10}$ transformation to correct for skew) was small and not significant, and the Bayes factor indicated some evidence for the null: $r = -0.082$, 95% credible interval = [-0.27, 0.12], $p = .418$, $BF_{01} = 5.79$.

Full participant instructions

Main Study Part 1

Prior to sitting down to read the consent form, the participants were briefly told what the task involved and what to expect. If they were happy to continue, they were presented with the consent form below. The questions after the consent form comprise the science engagement questionnaire mentioned in the main text (Gheorghiu et al., 2017).

Consent form

Consent form

This study is being conducted on behalf of Ana Gheorghiu and Dr. Mitch Callan at the University of Essex., and Dr. William Skylark at the University of Cambridge.

You will be asked to view short, silent videos of people giving a TED talk, and to indicate your assessment of the person in the video on a social dimension.

The study takes approximately 60-75 minutes, and can be divided into 2 separate sessions. You may withdraw at any time without giving a reason and without penalty. You will receive **£8** for your time – if the study has been divided into 2 sessions, payment will be made **after the final session**.

As well as your responses to the task, we will ask you to provide demographic information and your name in case we need to contact you and to make sure that the same person does not take part more than once. We will be able to link this information to your responses on the task, and may use your responses on this task to decide your suitability for future studies.

Any published report of the experiment will present the data anonymously.

The data you provide will normally only be accessed by the investigators, William Skylark, Mitch Callan and Ana Gheorghiu. We will not share your personal information with anyone else. Your data will be stored on PCs and in lockable university offices, although we cannot completely guarantee their security.

Please sign below to indicate that you have read and understood the information above.

Name: _____________________________ Signature: ___________________________

Date: __________________

If you are happy to take part, please complete the questions overleaf.

TO BE COMPLETED BY EXPERIMENTER

Participant ID: _____________ *Experimental booth:* ______________ *Task:*____________

Please indicate the extent to which the following statements apply to you. Use a scale from 1 to 7, where larger numbers indicate greater agreement.

1) I am knowledgeable about science

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

2) I find scientific ideas fascinating

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

3) I do not understand most scientific research

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

4) I like to read about scientific discoveries

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

5) I enjoy watching and listening to people describe scientific ideas

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

6) I have little interest in science

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

7) I am well-equipped to evaluate scientific arguments

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

8) I fully understand the scientific method

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

Your age: _________ Your gender: _________ Your nationality: ___________________________

What is your first language ("mother tongue")? ___________________________

Instructions

After signing the consent form, participants were taken to individual testing booths, where the study was run on computers. Participants were free to start, and continue through the task, at their own pace.

The main instructions at the beginning of the study are shown below. “Competent” was replaced with “Physically attractive”, “Moral” or “Sociable”, as appropriate.

> In this task, we would like you to watch short videos, and then answer some questions about the PERSON YOU SAW IN THE VIDEO.
>
> We would like you to judge how **COMPETENT** each person is.
>
> Please use a scale from 1 to 9, where **1 = not at all** and **9 = extremely**.
> If you **recognise** the person, please press “**r**” instead of providing a rating.
>
> Take your time over each judgment, and enter your response by pressing one of the **number keys at the top of the keyboard only (or the “r” key to indicate recognition)**.
>
> After your button press, the next video will load automatically.
>
> Press SPACEBAR to begin.

Before each video, participants were instructed:

“Please prepare to watch the next video. Please press SPACEBAR when you are ready”.

After each video, participants were cued to make a reponse with the following prompt. As above, “Competent” was replaced with “Physically attractive”, etc. as appropriate.

“How COMPETENT is this person?

1=Not at all 9=Extremely

Press "r" if you recognise the person”

At the end of the experiment, participants were thanked, paid, and verbally debriefed.

Main Study Part 2

Prior to sitting down to read the consent form, the participants were briefly told what the task involved and what to expect. If they were happy to continue, they were presented with the consent form below. The questions after the consent form comprise the science engagement questionnaire mentioned in the main text (Gheorghiu et al., 2017).

Consent form

Consent form

This study is being conducted on behalf of Ana Gheorghiu and Dr. Mitch Callan at the University of Essex., and Dr. William Skylark at the University of Cambridge.

You will be asked to view videos of people giving a talk, and to indicate your assessment of the video on several dimensions.

The study takes approximately 60 minutes, and you will watch 5 videos. You may withdraw at any time without giving a reason and without penalty. You will receive **£8** for your time.

As well as your responses to the task, we will ask you to provide demographic information and your name in case we need to contact you and to make sure that the same person does not take part more than once. We will be able to link this information to your responses on the task, and may use your responses on this task to decide your suitability for future studies.

Any published report of the experiment will present the data anonymously.

The data you provide will normally only be accessed by the investigators, William Skylark, Mitch Callan and Ana Gheorghiu. We will not share your personal information with anyone else. Your data will be stored on PCs and in lockable university offices, although we cannot completely guarantee their security.

Please sign below to indicate that you have read and understood the information above.

Name: _____________________________ Signature: ___________________________

Date: __________________

If you are happy to take part, please complete the questions overleaf.

TO BE COMPLETED BY EXPERIMENTER

Participant ID: _____________ *Experimental booth:* ______________ *Task:*____________

Please indicate the extent to which the following statements apply to you. Use a scale from 1 to 7, where larger numbers indicate greater agreement.

1) I am knowledgeable about science

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

2) I find scientific ideas fascinating

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

3) I do not understand most scientific research

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

4) I like to read about scientific discoveries

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

5) I enjoy watching and listening to people describe scientific ideas

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

6) I have little interest in science

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

7) I am well-equipped to evaluate scientific arguments

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

8) I fully understand the scientific method

Strongly disagree 1	2	3	4	5	6	Strongly agree 7

Your age: _________ Your gender: _________ Your nationality: ___________________________

What is your first language (“mother tongue”)? ___________________________

Instructions

After signing the consent form, participants were taken to individual testing booths, where the study was run on computers. Participants were free to start, and continue through the task, at their own pace.

The main instructions at the beginning of the study were as per below.

> In this task, we would like you to watch videos of people giving a talk. After each video, you will be asked to rate various aspects of what you have seen.
>
> At that point, you should use a scale from 1 to 9, where **1 = not at all** and **9 = extremely**.
> For any Yes/No questions, please use the keys **“y”** or **“n”**.
>
> Take your time over each judgment, and enter your response by pressing one of the **number keys at the top of the keyboard only, or the “y”/”n” keys if the question requires a Yes/No answer.**
>
> Please pay attention to each talk, as you will be asked questions about it afterwards.
>
> Please put your headphones on NOW.
>
> Press SPACEBAR to begin.

Before each video, participants were instructed:

“Please prepare to watch the next video. Please press SPACEBAR when you are ready”.

After each video, participants were cued to make their reponses with the following prompts.

“How would you rate the overall quality of this scientist's research?

1=Very poor   9=Very good

How good were this scientist's ideas?

1=Not at all   9=Extremely

How easy to follow/comprehend was this scientist's presentation?

1=Not at all   9=Extremely

How engaging/entertaining was this scientist's presentation?

1=Not at all   9=Extremely

How much would you recommend that we show this video to the next group of participants?

1=Not at all   9=Extremely

How likely would you be to share this video on any social media platform?

1=Not at all   9=Extremely

Have you heard/read about this research before (not just this general topic, but this specific piece of research) or seen this scientist before?

y=Yes   n=No"

At the end of the experiment, participants were thanked, paid, and verbally debriefed.

Supplementary Study

Prior to sitting down to read the consent form, the participants were briefly told what the task involved and what to expect. If they were happy to continue, they were presented with the following consent form and information sheet.

Consent form

Department of Psychology
Tel: 01225 38 3251
Mitchell Callan
m.j.callan@bath.ac.uk

UNIVERSITY OF BATH

CONSENT FORM: Impressions of People Delivering Science TED Talks

Please answer the following questions to the best of your knowledge

	YES	NO
DO YOU CONFIRM THAT YOU:		
• are a native English speaker	☐	☐
• are at least 18 years old	☐	☐
HAVE YOU:		
• been given information explaining about the study?	☐	☐
• had an opportunity to ask questions and discuss this study?	☐	☐
• received satisfactory answers to all questions you asked?	☐	☐
• received enough information about the study for you to make a decision about your participation?	☐	☐
DO YOU UNDERSTAND:		
that you are free to withdraw from the study and free to withdraw your data prior to anonymisation		
• at any time?	☐	☐
• without having to give a reason for withdrawing?	☐	☐

I hereby fully and freely consent to my participation in this study

I understand the nature and purpose of the procedures involved in this study. These have been communicated to me on the information sheet accompanying this form.

I understand and acknowledge that the investigation is designed to promote scientific knowledge and that the University of Bath will use the data I provide for no purpose other than research.

I understand that the data I provide will be kept **confidential**, and that on completion of the study my data will be **anonymised** by removing all links between my name or other identifying information and my study data. This will be done by 15 September 2018, and before any presentation or publication of my data.

I understand that after the study will be made "open data". I understand that this means the anonymised data will be publicly available and may be used for purposes not related to this study, and it will not be possible to identify me from these data.

Participant's signature: ______________________ Date: __________

Name in BLOCK Letters: ______________________

Final consent

Having participated in this study

I agree to the University of Bath keeping and processing the data I have provided during the course of this study in accordance with the information I received at the outset and the Data Protection Regulation.

Participant's signature: ______________________ Date: __________

Name in BLOCK Letters: ______________________

If you have any concerns related to your participation in this study please direct them to the Department of Psychology Research Ethics Committee, via email: psychology-ethics@bath.ac.uk.

Information Sheet

Impressions of People Delivering Science TED Talks

Dear Participant,

In this study, you will be asked to view several short, silent videos of people delivering science TED talks. You will also indicate your assessments of the videos and the people in the videos in terms of basic social traits (e.g., competence).

This study takes approximately 75 minutes to complete. You will receive £15 for your time.

As well as your responses to the tasks, we will ask you to provide basic background information about yourself (e.g., age, gender).

This study does not have any associated risks beyond what you would normally experience in day-to-day life. All the data will be anonymous; your name or any other personally identifying information will not be associated with the data you provide. You may withdraw at any time from the study without giving a reason and without penalty.

If you would like any more information regarding this study and/or have any questions please feel free to ask the researcher.

……………………………………………………………………………

Contact details:

Mitch Callan (m.j.callan@bath.ac.uk)

Department of Psychology

University of Bath

BA2 7AY

Bath, England

Instructions

After reading the information sheet and signing the consent form, participants were taken to individual testing booths where the study was run on computers. Participants were free to start, and continue through the task, at their own pace.

The main instructions at the beginning of the study were as per below.

> In this task, we would like you to watch short, silent videos of a science TED talk, and then answer a question about the PERSON GIVING THE SCIENCE TALK.
>
> We would like you to judge how **COMPETENT** each person is.
>
> Please use a scale from 1 to 9, where **1 = not at all** and **9 = extremely**.
> If you **recognise** the person, please press "**r**" instead of providing a rating.
>
> Take your time over each judgment, and enter your response by pressing one of the **number keys at the top of the keyboard only (or the "r" key to indicate recognition)**.
>
> After your button press, the next video will load automatically.
>
> Press SPACEBAR to begin.

The interim instructions and prompts were identical to those for the competence judgments in the Main Study, Part 1.

At the end of the experiment, participants were thanked, paid, and presented with the following debrief sheet.

Debrief sheet

Thank you for taking part in the study.

We are interested in the first impressions people form of scientists, and in what social dimensions are important in science communication.

Please feel free to ask the researcher if you have any further questions now. Alternatively, please contact Mitch Callan (M.J.Callan@bath.ac.uk) at a later date.