





Trustworthy, Reliable and Engaging Scientific Communication Approaches

D3.3 Experimental survey analysis and synthesis report



This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 872855.



PROJECT DESCRIPTION

Acronym:	TRESCA
Title:	Trustworthy, Reliable and Engaging Scientific Communication Approaches
Coordinator:	Erasmus University Rotterdam
Reference:	872855
Туре:	Research and Innovation Action (HORIZON 2020)
Program:	Science with and for Society (SwafS)
Theme:	(Science) Communication
Start:	01.01.2020
Duration:	28 months
Website:	https://trescaproject.eu
E-Mail:	tresca@eur.nl

Consortium: Erasmus University Rotterdam, The Netherlands (EUR), Coordinator Consejo Superior de Investigaciones Científicas, Spain (CSIC) Observa Science in Society, Italy (OBS) Zentrum Für Soziale Innovation GMBH, Austria (ZSI) Science Business Publishing Limited, United Kingdom (SBP) In a Nutshell, Kurzgesagt GmbH, Germany (KURZ)





DELIVERABLE DESCRIPTION

Number:	D3.3
Title:	Experimental survey analysis and synthesis report
Lead beneficiary:	CSIC
Work package:	WP3
Dissemination level:	Public (PU)
Туре	Report (R)
Due date:	30/06/2021
Submission date:	30/06/2021
Authors:	Sara Degli Esposti, CSIC; Jay Lee, EUR; Francisco Herreros, CSIC; Marina Tulin, EUR.
Contributors:	Adolfo Antón Bravo, CSIC.
Reviewers:	Gabor Szudi, ZSI.

Acknowledgement: This project has received funding from the European Union's Horizon 2020 Research and Innovation Action under Grant Agreement No 826497. **Disclaimer:** The content of this publication is the sole responsibility of the authors and does not in any way represent the view of the European Commission or its services.



TABLE OF CONTENTS

Introduction	7
Methodology and research questions	7
Data collection procedure	9
Data analysis	12
4.1 Description of Appendix One data	12
Institutional Trust	14
Conspiracy thinking	17
4.2 Description of Appendix Two data	19
Experimental findings - Appendix One	22
Dependent variables	23
Emotional reactions	25
Experimental Findings	27
Veracity assessment	29
Description of dependent variables - Appendix Two	33
Limitations	36
Conclusion	38
Appendix	39



EXECUTIVE SUMMARY

This report presents descriptive statistics and preliminary results of the analysis of survey data collected during February 2021 by means of a factorial survey designed and administered in working package three (WP3). Results of survey data analyses are discussed here and in forthcoming scientific publications as part of WP6. The insights gained from the factorial survey are also included in TRESCA MOOC teaching material developed in WP5. 7120 responses were gathered in total from nationally representative samples of individuals (age 18-75) living in France, Germany, Hungary, Italy, the Netherlands, Poland, and Spain (Appendix One). In Germany, the Netherlands and Spain, about 500 responses were collected from young adults (age 16-25), reaching a total of 1533 additional records (Appendix Two).

As part of TRESCA thematic area 'Environ-Mental Health', Social Science research findings about the negative effects social media may have on young people's mental health were communicated to study participants. These findings were published in reputable scientific journals and also communicated through documentary films. Relying on this material, we created four versions of a post that included a few statements and a visual aid. Treatment A included a basic statement presenting the findings, the original graph and mentioning the original source (Nature 578: 226). Treatment B added a sentence describing a personal experience of the matter to the previous statement and image. Treatment C used the same statement of treatment A, but replaced the source and the image with a graph shown in the documentary film *The Social Dilemma* and a caption indicating this was the source of the information. Finally, treatment D used the same text of treatment B and the image used in treatment C. Treatments B and D tested the effects of adding a personal story to the scientific communication on people's veracity assessment and on behavioral intentions such as deleting social media accounts. Treatments C and D tested the effects of using dramatic colours (black and purple) in the image combined with a nonscientific source (Netflix). The third factor tested in the first experiment involved the use of a reliability score informing study participants of the accuracy of the post. Respondents of Appendix One survey were given the option of seeing or skipping this score: 77% of them decided to see the score. Rather than focusing on fact-checking, in Appendix Two, which was tailored at young adults, additional questions about sharing or not the post were included.

The preliminary analysis of the data shows in Spain, Hungary and Poland an effect of the presence of the personal story in scientific communication. Changing the presentation of the visual aid seems to have no effect. However, effects of the dramaticity of the image over emotional reactions of Spanish respondents indicate the need to better explore interaction effects at play in each national context. Similar emotional reactions are also present in France, Poland and Hungary in the case of the presence of a personal story.



1. INTRODUCTION

The report is organised in three main sections. In the first part, we present the research questions that informed the study and the methodology used to test the hypotheses in the online survey experiments. In the second section, we move on to describing the data collected by means of two versions of the questionnaire distributed to general and then more specific demographic segments, called respectively Appendix One and Appendix Two. While Appendix One was available to people of all ages in all seven countries, Appendix Two included some additional questions and was tailored only to young adults (age 16-25) in Germany, Netherlands and Spain. In this report, we mostly rely on descriptive statistics to present the data. In the final part, we present a summary of the experimental findings. More sophisticated testing techniques and models are included in related academic publications.

2. Methodology and research questions

To better understand how we can improve science communication (SciCom), in WP3 we used experimental survey design to answer the following research questions and corresponding hypotheses.

In raising people's awareness and understanding of SSH science through SciCom with visual aids for enjoyment and opinion-forming,					
Research questions	Hypotheses				
Question 1. Is it more effective to emphasise the scientific message associated with a data visualisation or to combine data visualisation with a <i>personal story</i> ?	Appendix One H1. Treatment B Preliminary findings presented here.				
Question 2. Is it more effective to present the scientific message using a neutral or a <i>dramatic style</i> of data visualisation?	Appendix One H2. Treatment D Preliminary findings presented here.				
Question 3. To what extent does the opportunity of using fact checking services showing the <i>reliability score</i> of the scientific message influence respondents' opinions?	Appendix One H3. Treatments A1, A2; B1, B2; C1, C2; D1, D2 Preliminary findings presented here.				
Question 4. How does the presentation of the message influence people's (age 16-25) <i>sharing</i> attitudes and what are the reasons for deciding to share or not to share the post?	Appendix Two Q11 & Q12 Preliminary findings presented in forthcoming academic publications.				

Tab. 1 Research questions and hypotheses investigated in the large-scale online survey



An experimental survey was designed as part of task 3.1, and a survey instrument was published in D3.1 "Set of illustrated vignettes with questions." Factors included in the experimental design are: (1) the visual presentation of the scientific message, (2) the presence of a personal story in the explanation of the message, and (3) the opportunity of knowing the degree of reliability of the message offered by a fact-checking service. Table 2 shows inputs used in the $2 \times 2 \times 2$ experimental design. Treatments B and D tested the effect of including a personal story in the scientific communication; treatments C and C tested the effect of manipulating the dramaticity of the colours of the graphs showing the findings; finally the last treatments tested the effect of showing a high (green shield, labeled '1') or medium (orange shield, '2') reliability score. Only respondents who checked the option "I would like to double-check the veracity of the story" saw these reliability scores.

The scientific communication used as input was true and included findings originated from the academic field investigating how digital technologies affect adolescents' psychological well-being. A meta-analysis of these findings finds a negative relationship between social media use and psychological well-being.¹ These findings clearly relate to the second TRESCA thematic area, which is called "Environ-Mental Health". The text used as input was extracted from the article written by Jonathan Haidt and Nick Allen in 2020 titled "Digital technology under scrutiny" and published in Nature (578, 7794: 226-227). The sentence signalling a personal story was also taken from this article. The sentence used in treatment A was also used in all other treatments. Treatment A, in fact, represents the control group who saw the basic scientific message with the neutral graph published in Nature.

	A - No	B - Yes
FIRST FACTOR: Personal story (no, yes)	"Several studies show that there is a positive correlation between time spent on screens and bad mental-health outcomes." (Nature 578: 226)	"If you focus on social media you'll find strong evidence of harm especially for its millions of under-age users. When most of the 11-year-olds in a class are on Instagram (as was the case in my son's school), there can be pervasive effects on everyone. Children who opt out can find themselves isolated." (Nature 578: 226)

Tab. 2 All treatments included in the 2 x 2 x 2 experimental design in Appendix One survey

¹ See Orben, Amy. 2020 ("Teenagers, screens and social media: a narrative review of reviews and key studies", *Social Psychiatry and Psychiatric Epidemiology*: 1-8) and Haidt, Jonathan, and Nick Allen 2020 ("Digital technology under scrutiny", *Nature* 578 (7794): 226-227).



SECOND FACTOR: Scientific message – Neutral	Figure 1. Dependence on the fiber of the fib	$F_{\text{test}} = 10^{-2} \frac{10^{-2}}{200} 10^{$
THIRD FACTOR: Reliability score (high; medium)	A1 A2	B1 B2
	C	D
	"Several studies show that there is a positive correlation between time spent on screens and bad mental-health outcomes." (The Social Dilemma, Netflix 2020)	"If you focus on social media you'll find strong evidence of harm especially for its millions of under-age users. When most of the 11-year-olds in a class are on Instagram (as was the case in my son's school), there can be pervasive effects on everyone. Children who opt out can find themselves isolated." (The Social Dilemma, Netflix 2020)
SECOND FACTOR: Scientific message – Dramatic	LUS SECIOL RECON Der Mark The LODOCODE BIN	EVER SUBJECT FAILURE Dearty Type (USCODD) OF 1
	C1 C2	D1 D2
	REAL NEWS Reliability: 85%	REAL NEWS Reliability: 85% MOSTLY REAL Reliability: 52%

3. DATA COLLECTION PROCEDURE

Participants' recruitment and data collection procedure. A subcontractor (Dynata Global Spain SL) was appointed by CSIC to carry out the data collection in all seven countries. During February 2021 until the 10th of March, the survey was available to Dynata's proprietary panel members based in France, Hungary, Germany, Italy,



Netherlands, Poland, and Spain. The CSIC team worked in cooperation with the Spanish subsidiary of Dynata to create the electronic versions of the questionnaire available in report D3.1. Two versions of the questionnaire were used: Appendix One and Appendix Two. The first version was administered in France, Germany, Hungary, Italy, the Netherlands, Poland, and Spain, while the second version was used in Germany, the Netherlands and Spain.

Selection of countries. Hofstede's six-values model of culture (Hofstede 1984) offers an overview of the cultural dimensions amongst the seven countries used in this study. These six dimensions are:

- 1. power distance the extent to which members of a society accept unequal power distributions;
- 2. individualism a preference for loose social structures in which individuals mostly take care of themselves and their relatives;
- 3. masculinity an emphasis for achievement and competitiveness over cooperation and consensus;
- 4. uncertainty avoidance the degree to which members of a society feel uncomfortable when facing uncertainty or ambiguity;
- 5. long-term orientation low LTO societies prefer to maintain traditions and view societal change with suspicion; high LTO societies are more socially liberal;
- 6. indulgence the degree to which societies allow free gratification of basic human drives linked with hedonism.



Fig. 1 Values of Hofstede's six-values model of culture in Hungary, Italy, Poland and Spain

Key: Hungary (light blue); Italy (purple); Poland (green); Spain (Orange).



Fig. 2 Values of Hofstede's six-values model of culture in France, Germany and the Netherlands



Key: France (light blue); Germany (purple); Netherlands (green). Source: Hofstede Insights

Sample design. 7120 responses were gathered in total from nationally representative samples of individuals (age 18-75) living in France, Germany, Hungary, Italy, the Netherlands, Poland, and Spain (Appendix One). In Germany, the Netherlands and Spain, about 500 responses were collected from young adults (age 16-25), reaching a total of 1533 additional records (Appendix Two). Each national sample reflected populational characteristics in terms of gender, age and regional distribution. Table 3 reports the gender composition of individuals who filled in Appendix Two version of the electronic survey. Expected and observed gender and age distribution of individuals who filled in Appendix One version of the electronic survey are included in tables A.1 in the Appendix. All samples were also representative of the national population in terms of geographical distribution of residence.

APPENDIX TWO										
	GERMANY			NETHERLANDS			SPAIN			Tabal
GENDER	%	N Exp.	N Obs.	%	N Exp.	N Obs.	%	N Exp.	N Obs.	Total
Male	52%	270	229	49%	255	192	52%	270	236	657
Female	48%	250	266	51%	265	314	48%	250	275	855
TOTAL	100%	520	495	100%	520	506	100%	520	511	1512

Tab. 3 Expected and observed gender distribution of Appendix Two survey participants

Electronic questionnaire. Reading the vignettes and answering the questions took less than 10 minutes. Logic checks were built into the script to ensure participants cannot



continue if they try to submit an illogical answer. The survey included at the end the opportunity to share opinions by writing a text answering the question "Is there any other comment, impression or suggestion you would like to share with the TRESCA team?".

Translation of the questionnaire. The subcontractor translated the English version of the questionnaire into French, German, Hungarian, Italian, Dutch, Polish and Spanish. TRESCA team members from all consortium partners revised the electronic questionnaire at the end of January 2021 to ensure that the programmed logic was operating correctly and translations were appropriate.

Datasets. Overuse of item non-response by respondents (missing data) were identified, and their records were removed from the final data during the quality checks and before the delivery of the complete datasets. Also, responses where the completion time is less than 25% of the median length of the survey were identified across the entire sample and, based on these checks, fast responders were classified as speeders and were removed from the final data. The two complete datasets (one for Appendix One and another for Appendix Two) were submitted as D3.2 at the beginning of March 2021.

Data protection. Certification of full compliance with data protection rules was provided by the subcontractor. The legal basis for lawful data processing was explicit consent. No marketing actions were directed at any panelist by the subcontractor as a result of their participation in TRESCA research activities. Any time the subcontractor holded the complete survey participation history of participants on the panel, we asked the subcontractor to carefully select panel members for each survey to ensure that they are not being over-contacted.

4. DATA ANALYSIS

In this section, we present frequency distributions and graphic representations of variables included in the two survey experiments.

4.1 Description of Appendix One data

About 1000 individuals participated in the study in France (n = 1030), Germany (n = 1012), Hungary (n = 1018), Italy (n = 1014), the Netherlands (n = 1018), Poland (n = 1019), and Spain (n = 1009). People of all age groups, born between the year 1946 and 2002 filled in the survey. 51% of respondents were female and 48% were male. 28% of respondents said to be always online and 67% are often connected. 42% of respondents have a paid job and 33% meet their economic commitments and have some savings. Most respondents completed their studies around when they were 18 years old, with some finishing during their 30s and a very few studying when they were older.







Tab. 5 Respondents' current activity and economic situation

Activity	%	Meeting bills and credit commitments	Freq.	%
1. Paid work	42%	It is a constant struggle	854	11,99%
2. Studying	14%	It is a struggle from time to time	1689	23,72%
 Looking for benefits or charity support 	4%	Without much difficulty but with no savings	1998	28,06%
4. Looking for a job	13%	Without any difficulty and with some savings	2356	33,09%
5. Looking after my family or other vulnerable people	15%	Don't know: somebody else deals with it in my household	223	3,13%





Institutional Trust

With respect to trust, respondents are divided between those who trust others and those who do tend to distrust others. Regarding institutions, the press is more trusted (mean 43.91) than social media (38.36), while the European Union (48.42) is more trusted than national governments (40.03) or political parties (34.51). All differences are significant, under paired Wilcoxon signed rank tests. These results are consistent with European Social Survey² statistics. The ordering of trust for all countries - in descending order of both mean and median - is the EU, press, government, social media, the church, and finally political parties.

Tab. 6 *Trust in other people and institutions*



² https://www.europeansocialsurvey.org/





Across individual countries, there appears to be some variation in the ordering of trust. While the EU maintains a high level of trust across all countries, the trust in social media is the lowest (NL) to being 2nd (PL). Despite that, social media's actual variance is the lowest (SD 25.5) among the 6 trust measures. The church also receives varying levels of trust, with its lowest in Germany (DE) to highest in Hungary (HU). Similarly, government's rank also has a wide span: #1 (DE) to #6 (HU). Church exhibits the highest numerical variance across (SD 29.8) and within each country.



Rank	France	Germany	Hungary	Italy	Netherla nds	Poland	Spain
1st	The Press	Governm ent	EU	EU	People	EU	People
2nd	EU	People	People	The Press	Governm ent	Social Media	EU
3rd	People	EU	The Church	Governm ent	EU	The Press	The Press
4th	The Church	The Press	Social Media	The Church	The Press	People	Social Media
5th	Governm ent	Political parties	The Press	People	Political parties	The Church	Governm ent
6th	Social Media	Social Media	Governm ent	Social Media	The Church	Governm ent	The Church
7th	Political parties	The Church	Political parties	Political parties	Social Media	Political parties	Political parties

Tab. 7 *Trust rankings for each country*

70% of respondents say they are interested in politics and consider it important. 23% position themselves in the Centre of the political ideological spectrum, while 43% are more inclined toward the Right and 33% toward Left-wing parties.

Tab. 8 Frequency distribution of Interest in Politics and Political Ideology





Conspiracy thinking

Items from a conspiracy thinking scale were included in the survey. 82% of respondents agree with the statement: "Politicians are frequently bribed by major companies or interest groups." 62% think that "Much of what happens in the world today is decided by a small and secretive group of individuals." 40% of respondents feel comfortable with the idea that "in general, when people in a group set priorities, superior groups should dominate inferior groups rather than pushing for group equality." Finally, 25% of respondents think that "People never really landed on the moon, everything was recorded in TV studios."

Previous research shows that conspiracy beliefs are associated with the need to reduce uncertainty, increase control in the face of external threats, avoiding isolation through group identity and belonging (Troian et al. 2020).³ Unfortunately the items selected cannot be summarised into a unique component because of low correlation, as shown in the table below. Nonparametric methods for statistical data analysis available in Stata,⁴ such as Kendall's Tau correlation, were used to analyse the data. The advantage of nonparametric statistics over parametric statistics is that the former does not require the make assumptions (such as behaving like a normal distribution) on the conjoint distribution of the probability density function of the variables used in the model.

	[1]	[2]	[3]	[4]
1. Politicians are frequently bribed by major companies or interest groups	0.9698			
2. Much of what happens is decided by a small and secretive group of individuals	0.3583*	0.9779		
3. People never really landed on the moon, everything was recorded in TV studios	0.0434*	0.2326*	0.9195	
4. Superior groups should dominate inferior groups rather than pushing for equality	0.0738*	0.2454*	0.3403*	0.9618

Tab. 9 Kendall's Tau correlation of the four dimensions of conspiracy thinking

³ Adam-Troian, J., Wagner-Egger, P., Motyl, M., Arciszewski, T., Imhoff, R., Zimmer, F., ... & van Prooijen, J. W. (2020). Investigating the links between cultural values and belief in conspiracy theories: The key roles of collectivism and masculinity. *Political Psychology*.

⁴ https://www.stata.com/features/nonparametric-methods/





Tab. 10 Frequency distribution of conspiracy thinking variables



4.2 Description of Appendix Two data

Responses to 'Appendix Two' survey were gathered in Germany (n = 502), Spain (n = 519) and in the Netherlands (n = 512) from young people (age 16-25). Respondents' ages range between 16 (year of birth: 2005) and 25 years old (year of birth: 1996). Respondents are satisfied with their life (mean 16,48). 30% of respondents declare to be always online and less than 1% are never connected.

64% consider politics somehow important and are interested in it. 45% position themselves on the Right-hand side political spectrum, only 12% sit in the Centre and 42% prefer Left-wing parties. 76% think that "politicians are frequently bribed by major companies or interest groups." 62% say that "Much of what happens in the world today is decided by a small and secretive group of individuals." 49% consider that "in general, when people in a group set priorities, superior groups should dominate inferior groups rather than pushing for group equality." 35% believe that "people never really landed on the moon, everything was recorded in TV studios."



Tab. 11 Number of respondents from each country and age

Tab. 12 Number of respondents from each region

GERMANY (ACNielsen region)	%	Expected	Observed
Hamburg, Bremen, Schleswig-Holstein, Niedersachsen	16,0%	83	88
Nordrhein-Westfalen	21,9%	114	109
Hessen, Rheinland-Pfalz, Saarland	13,6%	71	74



Baden-Württemberg	12,9%	67	68
Bayern	15,0%	78	77
Berlin	4,1%	21	20
Meclenburg-vorpommern, Brandenburg, Sachsen-Anhalt	8,3%	43	35
Thürlingen, Sachsen	8,2%	43	31
Total	100,0%	520	502
NETHERLANDS (ACNielsen region)	%	Expected	Observed
Amsterdam, Rotterdam, Den Haag	15%	77	97
Rest west (Utercht, Noord-Holland, Zuid- Holland excl. Amsterdam, Rotterdam en Den Haag)	30%	153	131
Noord (Groningen, Friesland, Drenthe)	11%	55	62
Oost (Overijssel, Gelderland, Flevoland)	21%	109	113
Zuid (Zeeland, Noord-Brabant, Limburg)	24%	126	109
TOTAL	100%	520	512
SPAIN (ACNielsen region)	%	Expected	Observed
Barcelona Metropolitan	9%	49	52
North East	12%	61	64
East	15%	76	74
South	21%	107	133
Madrid Metropolitan	13%	67	69
Centre	10%	51	41
North West	11%	57	36
North Centre	10%	53	31
North Centre Islands	10% 4%	53 21	31 19



Fig. 3 Frequency distribution of respondents' degree of satisfaction with their life (n = 1475)



Fig. 4 Frequency distribution of respondents' daily internet use from 'never' to 'always' (n = 1498)



smartphones)?



5. Experimental findings – Appendix One

In this section, we present an overview of the experimental findings, starting from a description of the number of survey respondents subject to each treatment.

In the overall sample, 7120 respondents were exposed to treatments (A, B, C or D). 77% of respondents wanted to check the veracity of the scientific communication when offered the opportunity to see a reliability score. One of two scores was shown next to the graph and the text respondents previously saw: one with a green shield and a score of 85% and another one with an orange shield and a score of 52%.



Tab. 13 Number of cases per treatment





Tab. 14 Number of cases per treatment by country

	Persona	al Story	Dramatic Scientific Message					
Country	NO YES		NO	YES	Total			
France	511 519		515	515	1030			
Italy	481 533		515	515 499				
Germany	508 504		506	506 506				
Spain	504	505	505	504	1009			
Hungary	509	509	508	510	1018			
Poland	509 510		512	507	1019			
Netherlands	508 510		509	509	1018			
Total	3530	3590	3570	3550	7120			



Fig. 6 Percentage of respondents who wanted to double-check the veracity of the scientific *communication presented*



Out of curiosity, would you like to spend a few seconds double-checking if what was said in the post is true or false?

"Reliability Flag" by country

	Reliab	ility Flag	
Country	Green	Orange	Total
France	387	392	779
Italy	455	415	870
Germany	323	327	650
Spain	417	417	834
Hungary	425	426	851
Poland	417	416	833
Netherlands	349	350	699
Total	2773	2743	5516





Dependent variables

We tested the effects of the treatments on specific dependent variables (DV) measuring the veracity of the story, respondents' willingness to know more about the topic presented in the communication, and respondents' propensity to take an action as a result of the content of the message. In this case, we used respondents' propensity to delete their social



media accounts as a behavioral reaction. The figures below show the frequency distribution of the DVs.







Emotional reactions

The treatment conditions also generated emotional reactions. Before presenting these reactions, it is important to notice that, overall, respondents said they were satisfied with their life. On a scale from -50 (not at all satisfied) to +50 (very satisfied), the average response was +14.6. 72.8% of respondents marked an answer greater than zero, communicating their level of satisfaction. Only 19,4% of respondents declared to be unsatisfied, despite filling in the survey during the fourth wave of the CIVID19 pandemic.

Fig. 8 Level of respondents' satisfaction with life



All things considered, how satisfied are you with your life as a whole nowadays?

As shown in the following graphs, emotional reactions were pretty stable and consistent after the treatments, mostly because the last treatment simply confirmed and reinforced respondents' previous beliefs. The graphs below, which are the boxplots for the three affective dimensions of Pleasure, Arousal and Dominance measured through the AffectButton⁵ v3.3, demonstrate this consistency.

The range of values of the AffectButton in terms of Pleasure, Arousal and Dominance values lies between -1 and 1. Respondents answer by clicking on a dynamically changing icon with a facial expression. Pleasure relates to the positiveness versus negativeness of affect. Arousal is associated with the level of activation. Dominance measures respondents' perceptions of whether they are in control of the events or the context is influencing their

⁵ Broekens, J., & Brinkman, W. P. (2013). "<u>AffectButton: a method for reliable and valid affective self-report</u>", International Journal of Human-Computer Studies, 71(6), 641-667.



actions or decisions. Users do not need an explanation of these three affective dimensions before using the AffectButton.

Fig. 9 Boxplot for the affective dimensions Pleasure, Arousal and Dominance after treatments A, B, C and D



Fig. 10 Boxplot for the affective dimensions Pleasure, Arousal and Dominance after observing one of the two reliability scores





Experimental Findings

Here we present results calculated with the Mann-Whitney test, also known as the Wilcoxon rank sum test or the Wilcoxon-Mann-Whitney test. We use this technique to test the null hypothesis that the treatments had no effect, that is, that the samples were drawn from the same distribution. The results of this analysis have been calculated for each of the seven countries.

As presented in the tables below, the treatments seem to function only in specific cultural contexts. For example, Spanish respondents are more willing to believe that the SSH SciCom is true and are more willing to delete their social media accounts, when we add a personal story. The presence of a personal story also increases Polish and Hungarian respondents' propensity of believing that the story is true. In France, Hungary and Poland we also find an emotional effect of the presence of the personal story in the scientific message. We will explore in forthcoming academic articles the role emotions can play in orienting people's opinions and behavior.

Overall, we find no effect of manipulating the dramaticity of the graph included in the scientific communication except in the case of Spain, where changing the presentation of the graph triggered strong emotional reactions.

After seeing the fact-checking reliability scores, people are more willing to delete their social media accounts in Spain. Reliability scores also produce emotional reactions in Italy and Poland.

DV	France	Hungary	Germany	Italy	Netherla nds	Poland	Spain
Asking for more information	•	•	•	•	•	-	•
Deleting social media	•	•	•	•	•	-	
Veracity of the message	•		•	•	•		
Affective reaction (Pleasure)			•	•	•	-	•
(Arousal)	-	-	-	-	-	-	-
(Dominance)	-		-	-	-		-

Tab. 17 Effects of adding a personal story to the scientific communication - Appendix One



Tab. 18 *Effects of playing with the dramaticity of the graph added to the scientific communication - Appendix One*

DV	France	Hungary	Germany	Italy	Netherla nds	Poland	Spain
Asking for more information	•	-	•	•	•	•	•
Deleting social media	-	-	•	•	-	-	-
Veracity of the message	-	-	•	•	-	-	-
Affective reaction (Pleasure)	•	-	•	•	•	•	
(Arousal)	-	-	•	-	•	-	
(Dominance)	-	•	•	•	•	-	



DV	France	Hungary	Germany	Italy	Netherla nds	Poland	Spain
Asking for more information	•	-	•	-	-	-	-
Deleting social media	•	-	-	-	-	-	
Affective reaction (Pleasure)	•	-	-	-	-	-	•
(Arousal)	-	-	-	-	•		•
(Dominance)	•		•				

Tab. 19 Effects of viewing the green and orange reliability scores - Appendix One

Veracity assessment

The large majority of respondents said that in their opinion the story communicated was true, no matter which treatment they were subject to. However, on average, we find a significant difference in the way people perceive the truthfulness of the story in the seven countries (Pearson $\chi^2(12) = 128.0697$, p < .001). Poland, Italy, and Spain are the countries where more people believe the story is true (88%, 87%, 86%). Netherlands, Germany, and Hungary are the countries where more people believe the story is false (15%, 10%, 9%). France is the country where more people are indecisive, and cannot say whether the story was true or false (14%).

Do	Do you believe that what said in the post you just saw was true or false?														
	France	France Italy		Spain	Hungary	Poland	Netherlan ds	Total							
	Freq. (%)	Freq. (%)	Freq. (%)	Freq. (%)	Freq. (%)	Freq. (%)	Freq. (%)								
False	52 (6)	50 (6)	81 (10)	61 (7)	78 (9)	56 (6)	128 (15)	506 (8)							
Neither true, nor false	119 (14)	67 (7)	79 (9)	62 (7)	74 (9)	54 (6)	79 (10)	534 (9)							
True	664 (80)	770 (87)	687 (81)	756 (86)	682 (82)	800 (88)	619 (75)	4987 (83)							

Tab. 20 Number of people per country that believe the story is true, false, or are indecisive



Fig. 11 Number of respondents who believed the story was true, false or were indecisive by country



Do you believe that what said in the post you just saw was true or false? BY Country

The figure below shows the effects of the treatments on believing that the story is true for all the countries of our sample, from a linear regression in which the perception of the truthfulness of the story is the dependent variable, with a range from -30 to +30, and controlling for socio-demographics (gender, age and education). The reference category is treatment A, the control. As we can see, the expected value of the dependent variable is significantly higher in Treatments B and D with respect to the control. So, people exposed to these two treatments are more willing to think that the story is true than those exposed to the control. These two treatments differ from the control in that they introduce a personal example in the story. Treatment D also shows a different graphic, a more dramatic one, from the Netflix documentary "The Social Dilemma", instead of the more neutral graphic from Nature shown in treatment B and the control. However, given that Treatments B and D have a similar effect, we can conclude that it is the introduction of a personal note in those two treatments as compared to the more neutral treatment the manipulation that has made the story more credible.





Fig. 12 *Treatment effect over the likelihood of believing the story is true*

It is possible, however, that behind these treatment effects for the whole sample of countries there are actually differences between countries. In Figure 13, we show the effects of the treatments for each of the seven countries. The result is not as clear-cut as that of Figure 12. The general pattern found when we estimate the model for the whole dataset only stands in the cases of Spain and Hungary. In these two countries, treatments B and D show a higher expected value of the dependent variable as compared with the control. A similar pattern is found in Poland, where the coefficient of Treatment B is significant, while Treatment D is positive but not significant. In France, Italy and the Netherlands, there are no treatment effects, while in Germany there is a negative effect of Treatment B, meaning that those exposed to that treatment are less willing to believe that the story is trustworthy than those in the control group.





Fig. 13 *Treatment effect over the likelihood of believing the story is true by country*

Besides this analysis, we also tested the probability of believing the story was false with respect to believing it was true. In doing so, we transformed the continuous variable veracity assessment into a discrete variable with only three categories: true, neither true nor false and false. The results of this analysis are presented in the table below. The analysis confirms previous findings showing that the personal story reduces the probability of believing that the scientific communication is false in Spain and in Hungary. The transformation of the dependent variable and the introduction of other covariates also emerge as an effect of the personal story in Germany, but with the opposite effect: Germans seem to be more skeptical of the veracity of the communication because of the presence of the personal story. No effect is found in the case of Poland, instead. Despite the fact that the transformation of the dependent variables into a discrete variable may be the cause behind these observed variations in the findings, it is clear that we need to dig more into the complexities of these data and explore these nuanced variations in forthcoming scientific publications.



	Multinomial Logistic Regression BY Country (base outcome: True)												
	Do you believ	/e that what sai	d in the post yo	ou just saw was	true or false?								
			FALSE										
FRANCE	ITALY	POLAND	NETHERLANDS										
Gender	Gender	Gender	Gender	Gender	Gender	Gender							
Age	Age	Age	Age	Age	Age	Age							
Satisfaction	Satisfaction	Satisfaction	Satisfaction	Satisfaction	Satisfaction	Satisfaction							
Internet use	Internet use	Internet use	Internet use	Internet use	Internet use	Internet use							
Pleasure	Pleasure	Pleasure	Pleasure	Pleasure	Pleasure	Pleasure							
Arousal	Arousal	Arousal	Arousal	Arousal	Arousal	Arousal							
Dominance	Dominance	Dominance	Dominance	Dominance	Dominance	Dominance							
Social trust	Social trust	Social trust	Social trust	Social trust	Social trust	Social trust							
DRAMATIC	DRAMATIC	DRAMATIC	DRAMATIC	DRAMATIC	DRAMATIC	DRAMATIC							
STORY	STORY	STORY	STORY	STORY	STORY	STORY							
Politics	Politics	Politics	Politics	Politics	Politics	Politics							
Ideology	Ideology	Ideology	Ideology	Ideology	Ideology	Ideology							
		NEITH	HER TRUE, NOR	FALSE									
Gender	Gender	Gender	Gender	Gender (D)	Gender	Gender							
Age	Age	Age	Age	- Age	Age	Age							
Satisfaction	Satisfaction	Satisfaction	Satisfaction	Satisfaction	Satisfaction	Satisfaction							
Internet use	Internet use	Internet use	Internet use	Internet use	Internet use	Internet use							
Pleasure	Pleasure	Pleasure	Pleasure	Pleasure	Pleasure	Pleasure							
Arousal	Arousal	Arousal	Arousal	Arousal	Arousal	Arousal							
Dominance	Dominance	Dominance	Dominance	Dominance	Dominance	Dominance							
Social trust	Social trust	Social trust	Social trust	Social trust	Social trust	Social trust							
DRAMATIC	DRAMATIC	DRAMATIC	DRAMATIC	DRAMATIC	DRAMATIC	DRAMATIC							
STORY	STORY	STORY	STORY	STORY	STORY	STORY							
Politics	Politics	Politics	Politics	Politics	Politics	Politics							
Ideology	Ideology	Ideology	Ideology	Ideology	Ideology	Ideology							
Key: Orange cells	s represent positi	ve regression coe	fficients; blue ce	lls represent neg	ative coefficients	i.							

Tab. 21 Results of the multinomial logistic regression of the discrete veracity assessment

6. Description of dependent variables – Appendix Two

89% of respondents think that the post was true and 88% would like to know more about the effects of social media on people's wellbeing after reading it. 60% of respondents would be willing to delete some of their social media accounts after reading the post. When asked which social media account respondents would like to delete, 13% mention Instagram, 10% Facebook, 3,1% TikTok, 2,5% WhatsApp, 1,9% Twitter, 1,6% Snapchat, 0,9% Telegram, 0,7% YouTube, 0,5% Pinterest or Tinder.

As displayed in the table below, the same number of respondents was assigned to treatments testing the first and second factors of the experimental design.



		-					-
Tah	22 Numb	or of ro	nondante	accigned to	oach treatments	first and second	factor)
Iau.	ZZ NUIIID		ponuents	ussigned to			JUCLOI
				· · · · · · · · · · · ·		0	J · · · · /

	Germany	Spain	Netherlands		
TREATMENT	Freq.	Freq.	Freq.		
TREATMENT A	124	129	130		
TREATMENT B	122	128	128		
TREATMENT C	128	129	126		
TREATMENT D	128	133	128		
Total	502	519	512		

Most respondents (66%) assess the veracity of the post based on facts. However, 26% of respondents also rely on personal experience. This result is based on the analysis of an additional question asking whether respondents relied on facts or on personal experience in assessing the truthfulness of the post (Q10_1 Why do you believe so? Is your assessment of the authenticity of the post based on facts or on personal experience?). In Appendix Two this additional question followed the Veracity Assessment item (Q9_1 Do you believe that what said in the post you just saw was true or false?).









When asked whether respondents would like to share the post with their friends, 63% of them are positive about it. Respondents say they would share the post because (a) they "want to know other people's views and opinions" (77%); (b) "the post seems accurate" (72%); (c) "sharing is a culture and [they] share like others do" (53); (d) it "comes from authoritative sources" (60%); (e) it "makes [them] look good to other people" (46%); (f) "what's written in the post sounds frightening" (69%); (g) it "helps [them] bookmark useful information" (69%); and (h) "it's consistent with [their] beliefs or assumptions" (74%).







Emotional reactions caused by the post are similar to those observed in the other experiment.





Tab. 25 Density distribution of emotional variables



7. LIMITATIONS

As any experimental design, we tried to reduce the effect of hidden variables or other accidental events that could inadvertently influence the results. For this reason, for example, we included at the end of the survey a question asking respondents whether they had watched the documentary film "The Social Dilemma". We asked for it because the graphs we used in treatments C and D came from that setting. As shown below, the



majority of survey participants did not know the documentary film or have only heard about it. Only a very small fraction had watched the movie.

Another limitation is that respondents' encountering the Reliability Score treatment is conditional on their own decision (i.e. they have to request to see it) rather than the treatment's appearance determined a priori. This will be handled appropriately in future analyses, e.g., nested models or separate analyses on subsets of respondents.

Fig. 16 Number of respondents who had watched the documentary film the Social Dilemma before taking the survey



Have you ever watched the 2020 documentary film The Social Dilemma, produced by Netflix?



8. CONCLUSION

This report presents analyses of data collected during the large scale experimental surveys carried out in seven EU countries in February 2021. A 2 x 2 x 2 experimental design was tested by means of two electronic questionnaires, called Appendix One and Appendix Two. Both data collections were carried out by the same subcontractor, Dynata Global Spain SL. Appendix One was administered to people of all ages, while Appendix Two was targeted to young adults, age 16-25.

The analysis of the data demonstrates the need to create forms of scientific communications taking into consideration cultural differences and biases. Here we tested the effect of adding a personal story to a scientific message. While this strategy seems to be effective in Spain, Poland or Hungary, it does not seem to produce effects in other European countries. As noticed in other reports produced by the TRESCA team, playing with emotions in SciCom can also produce unintended effects. While we find almost no effect of manipulating the dramaticity of the colours used in the visual aid present in the scientific message, we need to better model the effect of emotional reactions in each national context to be sure it had no effect. As already written in the executive summary, more detailed and complex analyses will be published in forthcoming academic publications.

Another important aspect is how scientific communication can influence not only people's beliefs, but also their actions and behavior. The analysis of the data shows an effect of adding a personal story to Spanish people's propensity to delete their social media accounts. On the other hand, if given the opportunity, two third of respondents were willing to use a fact-checking service to verify the veracity of the information received. The large majority of respondents were also able to understand by themselves that the story was true. Only dutch respondents were on average more inclined to believe that the story was false; an effect that could have been produced by Dutch people's high reliance and positive attitudes toward digital technologies.



Appendix

Tab. A.1 Expected and observed sample composition in terms of age and gender per country

	APPENDIX ONE																					
	FRANCE		E	GERMANY		HUNGARY			ITALY		NETHERLANDS		POLAND		D	SPAIN			Tot			
GEND ER	%	N Exp	N Obs	%	N Exp	N Obs	%	N Exp	N Obs	%	N Exp	N Obs	%	N Exp	N Obs	%	N Exp	N Obs	%	N Exp	N Obs	Tot.
Male	48%	490	486	52%	530	527	48%	490	489	48%	490	478	49%	500	497	48%	490	482	48%	490	489	3448
Fema le	52%	530	542	48%	490	480	52%	530	529	52%	530	534	51%	520	517	52%	530	535	52%	530	516	3653
AGE	%	N Exp	N Obs	%	N Exp	N Obs	%	N Exp	N Obs	%	N Exp	N Obs	%	N Exp	N Obs	%	N Exp	N Obs	%	N Exp	N Obs	Tot.
18-24	15%	153	109	13%	133	114	18%	184	108	14%	143	111	13%	133	127	18%	184	166	13%	133	112	847
25-34	23%	235	223	20%	204	206	22%	224	239	25%	255	243	24%	244	240	22%	224	242	25%	255	251	1644
35-44	24%	245	261	27%	275	281	21%	214	231	23%	235	260	25%	255	256	21%	214	215	25%	255	260	1764
45-54	23%	235	251	22%	224	222	24%	245	266	20%	204	213	23%	231	232	24%	245	240	20%	204	210	1634
55-75	15%	153	186	18%	184	189	15%	153	174	18%	184	187	16%	162	163	15%	153	156	17%	173	176	1231
Tot	100	1020	1028	100	1020	1007	100	1020	1018	100	1020	1012	100	1020	1014	100	1020	1017	100	1020	1005	7101