RESEARCH ARTICLE

Framing superbugs - testing whether advocacy frames change attitudes, intention or behaviour using an online randomised control experiment [version 1; peer review: 1 approved, 1 not approved]

Daniel Metcalfe, Esraa Aldalooj, Matthew Thakur , Lara Clements

Wellcome Trust, UK, London, UK

V1 First published: 26 May 2021, 6:131
<https://doi.org/10.12688/wellcomeopenres.16723.1>
 Latest published: 26 May 2021, 6:131
<https://doi.org/10.12688/wellcomeopenres.16723.1>

Abstract

Background: Antimicrobial resistance (AMR) presents a significant threat to global health, requiring multifaceted action by individuals and policymakers. Advocates must persuade others to act. Making communication about AMR more effective could plausibly increase support for action. The Wellcome Trust-funded 'Reframing Resistance' project used communications research to develop framing recommendations for the language practitioners use to describe AMR. The aim of this study was to explore how this language influenced attitudes and behaviours towards AMR.

Methods: This study was a randomised trial to evaluate the effects of different styles of AMR framing language upon attitudinal and behavioural measures. Participants (n=1,934) were recruited in October 2019 using an on-line tool called "Prolific" and randomly assigned to review five variations of AMR narratives: four experimental frames which incorporated different combinations of language recommended by the framing guidelines, or a control frame without these features, taken from UN AMR communications. Participants were then asked a series of attitudinal and behavioural questions in relation to the AMR narrative they reviewed. Attitudes were measured using five-point Likert-type scales and behaviours were measured using binary variables. Descriptive analysis was used to explore respondents' characteristics and multivariable logistic regression models were used to establish independent associations between AMR frames and respondents' attitudes and behaviours.

Results: Participants who reviewed narratives that followed framing language guidelines were more likely to donate money or sign a petition, and rated narratives as more usable and important than participants who reviewed the control framing.

Conclusions: While larger trials with more diverse participants are needed to confirm generalisability, these results suggest that applying

Open Peer Review

Approval Status

	1	2
version 1
26 May 2021	view	view
1. Caline Mattar , Washington University in St Louis, St. Louis, USA		
2. Simon J G Otto , University of Alberta, Edmonton, Canada		

Any reports and responses or comments on the article can be found at the end of the article.

framing to communications could help advocates of AMR to drive action. Furthermore, the study reinforces the value of randomised studies to empirically test the impact of frames upon behaviour and indicates the potential for a more extensive programme of research.

Keywords

Framing, AMR, Antibiotic Resistance, Drug Resistance, Behaviour change

Corresponding author: Matthew Thakur (m.thakur@wellcome.org)

Author roles: **Metcalfe D:** Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Writing – Original Draft Preparation; **Aldalooj E:** Formal Analysis; **Thakur M:** Writing – Review & Editing; **Clements L:** Funding Acquisition, Supervision

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2021 Metcalfe D *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Metcalfe D, Aldalooj E, Thakur M and Clements L. **Framing superbugs - testing whether advocacy frames change attitudes, intention or behaviour using an online randomised control experiment [version 1; peer review: 1 approved, 1 not approved]** Wellcome Open Research 2021, 6:131 <https://doi.org/10.12688/wellcomeopenres.16723.1>

First published: 26 May 2021, 6:131 <https://doi.org/10.12688/wellcomeopenres.16723.1>

Introduction

Antimicrobial resistance (AMR) occurs when infections and illnesses treated by antimicrobial drugs evolve and become resistant to the medicines designed to kill them. Key drivers of AMR are a vast increase in the selective pressure upon microbial infections caused by the rapidly growing use of antibiotics in healthcare and agriculture, and market incentive issues which mean that no new classes of antibiotics have been developed in thirty years (World Health Organisation, 2019). The O'Neill review, led by a former Chief Economist of Goldman Sachs, estimated around 700,000 AMR-related deaths per annum and, if no action is taken, AMR could cause the deaths of 10 million people per annum by 2050 (O'Neill, 2016). A range of policy actions have been proposed to mitigate this problem and, whilst there has been some new investment and action since the publication of the O'Neill review, the actions undertaken to date are widely viewed to be insufficient (Clift, 2019).

Framing and communicating the risks that are associated with AMR are well-established in a range of policy and academic fields. Although there have been national level AMR awareness campaigns, the issue remains relatively poorly known (Huttner *et al.*, 2019). Encouragingly, this may mean that framing has more potential to be efficacious with respect to AMR than for more contentious issues.

However, there is little published evidence about the most appropriate means for presenting and communicating AMR. In this context, the Wellcome Trust commissioned communications research in 2018 on the optimal 'framing' of AMR. The research project was conducted by a commercial market research company in seven countries and utilised a methodology that included quantitative online message testing with over 12,000 people and qualitative research using 18 focus groups. This research recommendations were not validated using real-world data. Therefore, this paper provides evidence for the impact of different narratives of framing on AMR issues. We tested the hypothesis that the description of AMR developed using communications research or theoretical underpinnings will increase attitudinal response toward the perceived usability of the frame and the perceived importance of AMR.

The research question for this study was: do frames developed using *de novo* communications research or the behavioural evidence base change attitudinal and behavioural response, or increase understanding?

Methods

Study design and settings

Five concise AMR treatment frames were created – structured narratives that aim to communicate AMR risks to promote certain reader attitudinal shifts and/or actions. All were between 175 and 250 words and are reproduced in full in Figure 1 and Figure 2. The Control - labelled United Nations (UN) - was adapted from the AMR section of the website of the UN body,

the Food and Agriculture Organisation. It was selected as a real-world example of current communication practice (Box 1).

Box 1. Control frame

Control. United Nations Frame (UN)

"Antimicrobial resistance (AMR) is a major global threat of increasing concern to human and animal health. It also has implications for both food safety and food security and the economic wellbeing of millions of farming households.

AMR refers to when microorganisms – bacteria, fungi, viruses, and parasites – evolve resistance to antimicrobial substances, like antibiotics. This can occur naturally through adaption to the environment, the pace of AMR's spread is now on the uptick due to inappropriate and excessive use of antimicrobials

As a result of AMR, medicines that were once effective treatments for disease become less so – or even useless, leading to a reduced ability to successfully treat infections; increased mortality; more severe or prolonged illnesses; production losses in agriculture; and reduced livelihoods and food security.

The health consequences and economic costs of AMR are respectively estimated at 10 million human fatalities a year and a 2 to 3.5 percent decrease in global Gross Domestic Product (GDP), amounting to US\$ 100 trillion by 2050. However, the full impact remains hard to estimate."

The Control - labelled United Nations (UN) - was adapted from the AMR section of the website of the UN body, the Food and Agriculture Organisation. It was selected as a real-world example of current communication practice

Treatment 1 (RRes, Box 2) was taken directly from the recommendation in the 'Reframing Resistance' report (Wellcome, 2019, p. 30).

Box 2. Treatment frames

Treatment 1. Reframing Resistance (RRes).

"Common infections and injuries that were once easily treatable are becoming more dangerous and killing once again. This is because of drug-resistant infections which are undermining modern medicine.

Infections become drug-resistant when the bacteria that cause them adapt and change over time, developing the ability to resist the drugs designed to kill them.

The result is that many drugs – like antibiotics – are becoming less effective at treating illnesses. Our overuse of antibiotics in both humans and animals is speeding up this process.

Without working antibiotics, routine surgery like hip replacements, common illnesses like diarrhoea, and minor injuries from accidents, even including cuts, can become life-threatening.

People are already dying from drug-resistant infections, and as more drugs stop working, more lives will be put in danger. Drug-resistant infections can affect anyone; we are all at risk of infections from drug-resistant bacteria.

We can solve this problem. By taking action now to develop new drugs, and to make sure the drugs we already have stay effective, we can protect ourselves, our families and our communities.

This variant is directly taken from the 'Reframing Resistance' report, which recommends the use of this narrative.

Figure 1. Methodology flow diagram. The survey excluded those who failed an attention test, reported residence in the wrong country, or never consumed any of the listed media channels. 1,934 participants met all eligibility criteria.

. regress sumean genderd countryd age educationd mediad employmentd, beta						
Source	SS	df	MS	Number of obs	=	1,934
Model	30.4206474	6	5.07010791	F(6, 1927)	=	13.00
Residual	751.606369	1,927	.390039631	Prob > F	=	0.0000
				R-squared	=	0.0389
				Adj R-squared	=	0.0359
Total	782.027017	1,933	.404566486	Root MSE	=	.62453
sumean	Coef.	Std. Err.	t	P> t	Beta	
genderd	.095777	.0291174	3.29	0.001	.0742317	
countryd	.0209704	.0293316	0.71	0.475	.0164882	
age	.0615641	.0116765	5.27	0.000	.1241203	
educationd	-.0230039	.029459	-0.78	0.435	-.0178355	
mediad	.1281131	.0322444	3.97	0.000	.0947255	
employmentd	.0131313	.0357328	0.37	0.713	.00841	
_cons	4.029186	.0546604	73.71	0.000	.	

Figure 2. Linear regression Stata output for dependent variable 'importance' (sumean).

Treatment 2. RRes+Messenger and Episodic (RRes_ME)

I'd like to tell you about an issue many people have never heard of. Common infections and injuries that were once easily treatable are becoming more dangerous and killing once again. This is because of drug-resistant infections which are undermining modern medicine.

Infections become drug-resistant when the bacteria that cause them adapt and change over time, developing the ability to resist the drugs designed to kill them.

The result is that many drugs – like antibiotics – are becoming less effective at treating illnesses. Our overuse of antibiotics in both humans and animals is speeding up this process.

Without working antibiotics, routine surgery like hip replacements, common illnesses like diarrhoea, and minor injuries from accidents, even including cuts, can become life-threatening.

One person who has been affected by resistant infections this year is Anna. After a caesarian section, she acquired a multidrug-resistant wound infection that required antibiotics for 6 straight weeks, with several readmissions to the hospital. Then she needed daily home healthcare and treatment from an advanced wound center for 15 weeks. The costs to her family were substantial, and took a toll on both financially and emotionally. Her entire maternity leave was taken up with this infection.

People are already suffering and dying because of drug-resistant infections and, as more drugs stop working, more lives will be put in danger. Drug-resistant infections can affect anyone; we are all at risk of infections from drug-resistant bacteria.

We can solve this problem. By taking action now to develop new drugs, and to make sure the drugs we already have stay effective, we can protect ourselves, our families and our communities.

Doctor Karen Yamaguchi, Chief Medical Officer, World Health Organisation

This variant adds messenger and episodic factors to the RRes frame. The messenger has the formal authority of a fictional senior Doctor from the World Health Organisation. The episodic focus was adapted from a case study on the 'Resistance Fighters' campaign website (<https://antimicrobialresistancefighters.org/>) and was selected as a caesarean section is a procedure with high recognition levels.

Treatment 3. UN+Messenger and Episodic (UN_ME)

"I'd like to tell you about an issue many people have never heard of. Antimicrobial resistance (AMR) is a major global threat of increasing concern to human and animal health. It also has implications for both food safety and food security and the economic wellbeing of millions of farming households.

AMR refers to when microorganisms – bacteria, fungi, viruses, and parasites – evolve resistance to antimicrobial substances, like antibiotics. This can occur naturally through adaption to the environment, the pace of AMR's spread is now on the uptick due to inappropriate and excessive use of antimicrobials

As a result of AMR, medicines that were once effective treatments for disease become less so – or even useless, leading to a reduced ability to successfully treat infections, increased mortality; more severe or prolonged illnesses; production losses in agriculture; and reduced livelihoods and food security.

One person who has been affected by AMR this year is Anna. After a caesarian section, she acquired a multidrug-resistant wound infection that required antibiotics for 6 straight weeks, with several readmissions to the hospital. Then she needed daily home healthcare and treatment from an advanced wound center for 15 weeks. The costs to her family were substantial, and took a toll on both financially and emotionally. Her entire maternity leave was taken up with this infection.

The health consequences and economic costs of AMR are respectively estimated at 10 million human fatalities a year and a 2 to 3.5 percent decrease in global Gross Domestic Product (GDP), amounting to US\$ 100 trillion by 2050. However, the full impact remains hard to estimate.

Doctor Karen Yamaguchi, Chief Medical Officer, World Health Organisation."

This variant adds messenger and episodic factors to the Control UN frame. As per Treatment 2, the messenger has the formal authority of a fictional senior Doctor from the World Health Organisation. The episodic focus was adapted from a case study on the 'Resistance Fighters' campaign website (<https://antimicrobialresistancefighters.org/>) and was selected as a caesarean section is a procedure with high recognition levels.

Treatment 4. UN+Gain (UN_G).

"Antimicrobial resistance (AMR) is a major global threat of increasing concern to human and animal health. It also has implications for both food safety and food security and the economic wellbeing of millions of farming households.

AMR refers to when microorganisms – bacteria, fungi, viruses, and parasites – evolve resistance to antimicrobial substances, like antibiotics. This can occur naturally through adaption to the environment, the pace of AMR's spread is now on the uptick due to inappropriate and excessive use of antimicrobials

As a result of AMR, medicines that were once effective treatments for disease become less so – or even useless, leading to a reduced ability to successfully treat infections. If we can mitigate AMR we can save millions of lives and billions of dollars, prevent severe or prolonged illnesses, prevent production losses in agriculture, and protect livelihoods and food security.

There is so much that can be done to mitigate AMR and prevent further loss of life. By developing new medicines we can develop new options for patients. Reducing unneeded antibiotic use in animals helps to slow down the development of resistance. Improving sanitation and hygiene can reduce the number of people getting infections in the first place."

Treatment 1 utilises a loss frame in relation to AMR. Following the adaptation of Prospect Theory to health information, a gain frame may be more appropriate for a prevention issue such as AMR – tested in this treatment.

The messenger and episodic treatments. As credible messengers and an episodic focus are well evidenced strength factors (Busby *et al.*, 2018) two frames were developed to test their impact in this context. Neither the Control (UN) frame nor Treatment 1 (RRes) frames use a specific messenger or episodic focus, so it was straightforward to create new treatment frames incorporating these factors - Treatment 2 (RRes_ME) and Treatment 3 (UN_ME) – reproduced in Box 2. The messenger has the formal authority of a fictional senior Doctor from the World Health Organisation. The episodic focus was adapted from a case study about a woman who had suffered protracted complications due to AMR following a caesarean section, taken from the 'Resistance Fighters' campaign website¹.

By adding two strength factors we do create a risk of conflating the impact of both. However, following precedent from the Behavioural Insight Team (Bholat *et al.*, 2018) it seemed

¹ <https://antimicrobialresistancefighters.org/>

plausible that a greater effect size would be created by the additive combination of factors. Given limited budgets and a focus on practitioner applicability, studying the combination was judged to be of greater research relevance than testing solely one of the two factors.

The gain frame treatment. Given the strong theoretical basis and wide, if not definitive, evidence base for the impact of gain and loss frames (Rothman & Salovey, 1997), we decided to test this in the context of AMR. Following the adaptation of Prospect Theory to advocacy communication in Spence & Pidgeon (2010) above, a gain frame may be more appropriate for a prevention issue such as AMR. The control (UN) version uses a loss frame (this is assumed to be an intuitive expert decision). The fourth treatment, 'UN_G', adds a gain frame, which is predicted by Rothman & Salovey (1997) to increase behavioural responses.

As can be seen in Box 2, the 'Reframing Resistance' commercial research utilised a hybrid gain-loss frame - adapting Treatment 1 (RRes) would have been less straightforward. A further test could have been to create two new treatment versions of RRes, one purely loss and one purely gain. However, moving from a hybrid frame to a loss or gain has less precedent in the literature, so it is harder to be confident that there could be a significant effect. Within the limited number of frames this study could test findings that were likely to increase response were prioritised as this would be of more practical use to AMR advocates.

Variables of interest

Framing studies use problematically inconsistent categorisations of persuasiveness - frequently conflating attitudes and intentions with behaviour (Gallagher & Updegraff, 2012). This study assesses attitudinal and behavioural measures to gain a more nuanced understanding of frame persuasiveness. The primary dependent variable was behavioural response. Given the constraints of an online study, respondents were asked to take two behavioural actions. The first asked if they would sign a petition for or against greater government investment in AMR and the second whether they would donate a percentage of their fee for completing the survey to an AMR charity. This was recoded into a combined dummy variable where 0 = no behavioural action and 1 = one or more behavioural actions.

In addition, this study assesses three attitudinal response variables - Usability (HCMean), Importance (SUMean) and Understanding. Usability was assessed by asking two questions about the perceived clarity and helpfulness of the frame - for example, "*How clear/helpful, or otherwise, do you find this as an explanation of what is happening with AMR? Please rate on a scale from 1 to 5, where 1 is extremely unclear/unhelpful, and 5 is extremely clear/helpful.*" Importance was assessed by asking about how urgent and significant respondents felt AMR to be - for example, "*Based on this statement, how urgent a priority do you now believe AMR is? Select scale where 1 is not at all urgent and 5 is highly urgent.*" These were combined into two dummy variables, using the mean of each respondent's

combined response. A 1–5 Likert was chosen as the literature suggests that it is less confusing, increasing both response rate and response quality whilst reducing respondents' "frustration level" (Babakus & Mangold 1992; Buttle, 1996).

One question was asked to gauge if the frames influenced respondents' understanding of AMR. The qualitative stage of the Reframing Resistance (Wellcome, 2019) project had identified 'widespread' misattribution of AMR - with some respondents believing that humans, not infections, became resistant to medicines. Respondents were asked to choose one of three options, for example, "*Drug Resistant Infections occur when your body becomes resistant to medicines, meaning that those medicines no longer work as well on you*", two of which were incorrect. This was recoded into a dummy variable where 0 = incorrect and 1 = correct. Independent variables of gender, age, education level, media consumption and country of residence were also recorded. To test for variance in different English-speaking cultural contexts the survey was run in the US and the UK.

Sample size and participant recruitment

Given that such a large number of psychological studies are significantly underpowered (Maxwell *et al.*, 2015) calculating sample size *a priori* is crucial. Based on comparable studies (Opiyo *et al.*, 2013; Rosenbaum *et al.*, 2010 and Santesso *et al.*, 2015) that indicate small to medium effect sizes we estimated a mean difference of 10% between answers within the treatment and control groups. At an alpha of 0.05 and power of 0.9, with an estimated standard deviation of 0.3 in both samples, Stata calculates a sample size of 190 per group. Allowing for attrition/ineligibility at 5%, we aimed to recruit 200 people to each arm of the trial, so 1,000 per country.

The survey was developed in the Qualtrics online survey platform and carried out using the Prolific online respondent platform. UK and US based respondents participated in a between-subjects survey design on the 23rd October 2019 and were paid approximately US\$1 for completion of the survey, which took an average of 6 minutes to complete.

Participants were randomly assigned to one of five conditions using the Qualtrics randomisation algorithm. The survey asked subjects to provide key demographic information, read one of five framed statements about AMR, and answer seven follow up questions about Usability, Importance, Understanding and Behaviour. The sample was split evenly between the UK and the US to enable analysis by country. To reflect different media consumption patterns in the US and the UK, two versions of the survey were created, each containing different lists of popular publications based on country of residence. The UK version of the survey is provided as extended data (Metcalfe *et al.*, 2021b).

The survey excluded those who failed an attention test, reported residence in the wrong country, or never consumed any of the listed media channels. The rationale for the latter exclusion was twofold - it mirrored the selection criteria that Wellcome

had used for the quantitative online message testing in 'Reframing Resistance', and the theory of change underpinning the work is media advocacy - the primary mechanism for distribution of framed messages would be through media channels, so focusing on users of those channels makes the findings more pertinent.

Ethical approval, consent and data collection

All participants provided full and informed consent at the time the study is carried out via the start of the survey. Each participant was assigned a unique numeric identifier so no personal information could be accessed by researchers. Prior to recruitment, the London School of Economics (LSE) ethics checklist was completed. No significant issues arose and no further approval was required. The completed checklist is provided as extended data (Metcalf *et al.*, 2021b).

Statistical methods

Data was analysed in Stata 15 (the full Stata code used is included as extended data, (Metcalf *et al.*, 2021b)).

Descriptive statistics were collected and ranked by mean response. The differences between the control and four treatments were analysed via the creation of aggregated categories for the four dependent variables. Where dependent variables were assessed via a continuous variable the significance of differences between responses to the treatment frames were assessed using the Wilcoxon rank-sum method due to the nonparametric data skew. The remaining (dichotomous) dependent variables were assessed via a two-way proportions test. The alpha level used for the tests was 0.05 (Hair *et al.*, 2006). Linear and logistic regression analysis, analysis of variance and post-hoc pairwise comparisons were also conducted to further interrogate results.

Critical values were adjusted to allow for family-wise error rates. Because several dependent variables were tested at the same time, the Holm-Bonferroni (H-B) adjustment was used - this method performs better than the Bonferroni correction procedure in terms of maintaining statistical power (Holm, 1979). This approach first ranks the p-values from smallest to largest. The alpha value 0.05 of the lower p-value is divided by the number of dependent variables. For example, the revised H-B significance value for the first ranked p-values follows:

$$HB = 0.05/4-1+1 = .0125$$

The second smallest p-value then has the alpha value divided by the number of tests minus 1. This process continued one-by-one with the progressively greater p-values until statistical significance was no longer met.

Results

Of 2,101 respondents, 1,934 participants met all eligibility criteria and were included in the analysis (Figure 1) (Metcalf *et al.*, 2021a). Dummy variables were created for gender, media consumption, country of residence and employment status. Sample characteristics are summarised below (see the linked Underlying Data for full details):

- 58% of the sample were female
- 33% consumed media from the provided UK or US lists every day
- 58% had a degree or postgraduate qualification
- 79% were working

Collectively, the subjects:

- Assessed the five frames as usable - rating clarity at a mean of 4.41 (SD 0.75) and helpfulness at a mean of 4.31 (SD 0.77)
- Felt AMR was both a significant threat (mean 4.40, SD 0.67) and an urgent concern (mean attitude rating of 4.33, SD 0.72)
- Understood the nature of the problem (91% gave a correct answer)
- Almost two thirds (65%) undertook at least one of the prompted behaviours.

A preliminary step was to check if the frames had influenced the dependent variables in the expected direction. The results showed that, in the control group, the dependent variables Usability and Importance had means of 4.23 and 4.36 on a 1–5 scale. The dichotomous Understanding and Behavioural dependent variables had means of 0.91 and 0.59, respectively. The Understanding mean showed no clear difference between control and treatment frames. Among participants reading the treatment frames, Usability, Importance and Behaviour means were all higher than the control means. A Spearman's correlation between the dependent variables found a moderate correlation of 0.361 between Usability and Importance, and weak correlations between Behaviour and Usability (0.194), and Behaviour and Importance (0.261), indicating they were linked but distinctive variables (Akoglu, 2018).

Treatment 1 vs control

Initial analysis investigated the relative performance of the Control frame (UN) and Treatment Frame 1 (RRes), representing 775 observations (Table 1). In terms of the continuous dependent variables, a skewness/kurtosis test for both Usability and Importance finds that the three p-values are all less than 0.001 – so we can conclude that the distribution of attendance is statistically different from a normal distribution. Both exhibited a negative skew (Usability -1.26, Importance -0.92) and were strongly platykurtic (Usability 4.99, Importance 4.44). Because the distribution is not normal, and the variable is continuous a Mann-Whitney test was used. For the dichotomous variables Understanding and Behaviour, a two-sample test of proportions was used.

The Usability and Behaviour variables show a clear difference in mean between Control (UN) and Treatment 1 (RRes), with significant z scores and p values, even when adjusted with the Holm-Bonferroni correction.

Table 1. Analysis of dependent variables for control (UN) vs treatment 1 (RRes).

Dependent Variable	Results	H-B adjusted sig. level	p	z	
Usability	Control (UN) Mean 4.23 (SD 0.71) Treatment 1 (RRes) Mean 4.43 (SD 0.66)	0.0125	<0.001**	4.23	Mann-Whitney
Importance	Condition (UN) Mean 4.28 (SD 0.66) Treatment 1 (RRes) Mean 4.41 (SD 0.56)	0.025	0.021	2.30	
Behaviour	Control (UN) Mean 0.59 (SE 0.025) Treatment 1 (RRes) Mean 0.69 (SE 0.023)	0.017	0.003**	-2.94	Two sample test of proportions
Understanding	Control (UN) Mean 0.91 (SE 0.014) Treatment 1 (RRes) Mean 0.89 (SE 0.016)	0.054	0.340	0.95	

* $p < 0.05$; ** $p < 0.01$

Initial analysis investigated the relative performance of the Control frame (UN) and Treatment Frame 1 (RRes), representing 775 observations. The frame created a statistically significant increase in perceived Usability and Behaviour. There was no significant difference in Importance or Understanding between the control (UN) and treatment (RRes) frames.

The influence of the frame upon respondent behaviour was assessed using the dependent variable Behaviour, where 0 = no action taken and 1 = at least one action taken. A two-sample test of proportions shows a significant z score of -2.94 and a p-value of 0.003, we can reject the null hypothesis of no difference between groups for behavioural response.

For the continuous dependent variable Usability, the non-adjusted p value is <0.001, indicating significance. A simple analysis of variance (ANOVA) also finds an R squared of 0.02 - the proportion of variance explained by the treatment dummy variable is very small. We can reject the null hypothesis for this attitudinal response variable. The variable Importance also increases, but not to an extent where we can reject the null hypothesis.

The survey also contained one question to gauge Understanding. The mean for control (UN) was 0.91 and for the treatment (RRes) was 0.89 - there is no significant difference as reflected in a p-value of 0.34 and we cannot reject the null hypothesis that there is no difference in this attitudinal response variable between groups.

The 'Reframing Resistance' frame created a statistically significant increase in perceived Usability and Behaviour. There was no significant difference in Importance or Understanding between the control (UN) and treatment (RRes) frames.

Impact of gain frame: UN_G vs. control (UN)

Following the approach used for Treatment 1, the distribution is not normal for dependent variables Usability and Importance, so a Mann-Whitney test was used, and a two sample test of proportions was used for the variables Behaviour and Understanding. The results are described in Table 2.

Contrary to hypothesis, the addition of a Gain frame to control did not create a statistically significant increase in the dependent variable of Behaviour, or the secondary dependent variables of Importance, Usability and Understanding. Overall, we cannot reject the null hypothesis.

Impact of the messenger-episodic frames

Two treatment frames (UN_ME and RRes_ME) incorporating messenger and episodic variants were tested, one adapting the Control (UN) frame, and the second adapting the (RRes) frame. Augmenting the UN frame with messenger and episodic factors makes a statistically significant difference to behavioural measures versus the UN control. Secondary dependent variables of Usability and Importance also increased significantly. As in all treatments, Understanding does not alter significantly versus control. The RRes_ME results see significant increases in Behaviour and Usability, but not Importance or Understanding. The results are described in Table 3. Table 4 summarises the effects observed across all 4 experimental frames versus control.

Regression of dependent variables: Understanding and Importance

As the dependent variables Importance and Understanding are continuous it is possible to run linear regressions to model the relationships between variables. The model output for Importance is described in Figure 2.

F-statistics of 5.19 (Usability) and 13.00 (Importance), both with corresponding p-values of <0.001, indicate the models fit better than a model with no predictors. In terms of Usability, the independent variables age and media consumption are positively associated, showing high t and low p-values. This suggests that older respondents and those who read from the target

Table 2. Impact of gain frame across all dependent variables.

Dependent Variable		H-B adjusted sig. level	p	z	Test used
Usability	Control (UN) Mean 4.23 (SD 0.71) Treatment 4 (UN_G) Mean 4.35 (SD 0.63)	0.0125	0.012	2.44	Mann-Whitney
Importance	Condition (UN) Mean 4.28 (SD 0.66) Treatment 4 (UN_G) Mean 4.28 (SD 0.67)	0.025	0.97	-0.04	
Behaviour	Control (UN) Mean 0.59 (SE 0.025) Treatment 4 UN_G) Mean 0.63 (SE 0.024)	0.017	0.22	-1.22	Two sample test of proportions
Understanding	Control (UN) Mean 0.91 (SE 0.014) Treatment 4 (UN_G) Mean (0.93) SE (0.013)	0.05	0.43	-0.78	

*p<0.05; **p<0.01

Contrary to hypothesis, the addition of a Gain frame to control did not create a statistically significant increase in the dependent variable of Behaviour, or the secondary dependent variables of Importance, Usability and Understanding.

Table 3. Impact of messenger and episodic treatments on dependent variables.

Dependent Variable		H-B adjusted sig. level	p	z	Test used
Usability	Control (UN) Mean 4.23 (SD 0.71)				Mann Whitney
	Treatment 2 (RRes_ME) Mean 4.43 (SD 0.67)	0.0125	0.000**	-4.64	
	Treatment 3 (UN_ME) Mean 4.47 (SD 0.66)	0.025	0.004**	-2.87	
Importance	Condition (UN) Mean 4.28 (SD 0.66)				
	Treatment 2 (RRes_ME) Mean 4.40 (SD 0.63)	0.017	0.008**	-2.67	
	Treatment 3 (UN_ME) Mean 4.43 (SD 0.63)	0.0125	0.000**	-3.42	
Behaviour	Control (UN) Mean 0.59 (SE 0.025)				Two sample test of proportions
	Treatment 2 (RRes_ME) Mean 0.66 (SE 0.024)	0.025	0.030*	-2.17	
	Treatment 3 UN_ME) Mean 0.70 (SE 0.023)	0.017	0.001**	-3.33	
Understanding	Control (UN) Mean 0.91 (SE 0.014)				
	Treatment 2 (RRes_ME) Mean 0.89 (SE 0.015)	0.05	0.32	0.99	
	Treatment 3 UN_ME) Mean 0.91 (SE 0.014)	0.05	0.78	-0.26	

*p<0.05; **p<0.01

Two treatment frames (UN_ME and RRes_ME) incorporating Messenger and Episodic variants were tested, one adapting the Control (UN) frame, and the second adapting the (RRes) frame. Augmenting the UN frame with Messenger and Episodic factors makes a statistically significant difference to behavioural measures versus the UN control. Secondary dependent variables of Usability and Importance also increased significantly. As in all treatments, Understanding does not alter significantly versus control. The RRes_ME results see significant increases in Behaviour and Usability, but not Importance or Understanding.

media list daily were more likely to find the frames helpful and clear. Level of education, country of residence, gender and employment status are not significantly associated. For the dependent variable Importance, age and media were again significantly associated. In addition, gender exhibited a t statistic of 3.29 and a p value of 0.001, indicating that women were

significantly more likely to think of AMR as a significant and urgent problem. Standardised regression coefficients suggest that the strongest predictor variable for Usability is media, where a one standard deviation increase leads to a 0.08 standard deviation increase in Usability. For Importance the strongest predictor variable is age, where a one standard deviation increase

would lead to a predicted 0.12 standard deviation increase in Importance. However, adjusted R^2 values of 0.012 (Usability) and 0.03 (Importance) indicate that only around 1 and 3 percent of variability is explained by the factors included in this model.

Additional analysis of dependent variable: Behaviour

As the main dependent variable of interest, Behaviour, is dichotomous it is recommended to run a logistic regression rather than a typical linear regression model. The results are described in [Figure 3](#).

Table 4. Summary of effects of each condition on dependent variables.

Candidate predictors Odds Ratio of reporting at least one behaviour (likelihood)	Behaviour	Usability	Understanding	Importance
Condition (vs UN framing)				
RRes	1.6	1.9	0.8	1.2
RRes_ME	1.4	1.9	0.8	1.3
UN_ME	1.6	1.3	1.1	1.5
UN_G	1.2	1.3	1.2	0.9

. logistic behaviour i.condition genderd age educationd countryd mediad employmentd						
Logistic regression			Number of obs		=	1,934
			LR chi2(10)		=	30.64
			Prob > chi2		=	0.0007
Log likelihood = -1232.5178			Pseudo R2		=	0.0123
behaviour	Odds Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
condition						
1	1.546221	.2338979	2.88	0.004	1.149502	2.079857
2	1.38771	.2081102	2.18	0.029	1.034301	1.861875
3	1.643367	.2502206	3.26	0.001	1.219354	2.214824
4	1.190008	.1766681	1.17	0.241	.8895718	1.591912
genderd	.9316333	.0922082	-0.72	0.474	.7673572	1.131078
age	.8920269	.0350369	-2.91	0.004	.8259326	.9634104
educationd	.9466716	.0948638	-0.55	0.584	.7778616	1.152116
countryd	1.297706	.1293157	2.62	0.009	1.067467	1.577605
mediad	1.09679	.1202817	0.84	0.400	.8846561	1.359791
employmentd	.9232346	.1130589	-0.65	0.514	.7262302	1.17368
_cons	2.126436	.4484904	3.58	0.000	1.406449	3.214998
Note: _cons estimates baseline odds.						

Figure 3. Logistic Regression with dependent variable 'behaviour'.

This model confirms high z scores and low (non-adjusted) p values for treatments 1 (RRes), and 3 (UN_ME), and non-significant effects for treatments 2 (RRes_ME) and 4 (UN_G). The likelihood ratio chi-square of 30.64 with a p -value of <0.001 indicates the model as a whole fits significantly better than an empty model (i.e. a model with no predictors).

In terms of independent variables, country and age are significantly correlated with the dependent variable, showing high z scores and low p -values. In the case of age, the z score is negative - this suggests UK respondents and younger respondents were more likely to take action than US and older participants. The odds ratios show that the odds of taking one behavioural action are increased by a factor of 1.30 if you are from the UK. However, a pseudo R^2 of 0.012 indicates that this model is around 1 percent better than a null model.

The difference in behavioural response between the US and the UK warranted further investigation. Behavioural response was based on two questions - whether respondents were willing to sign a petition and/or make a small donation. A two sample test of proportion outputs found no significant difference between countries in terms of respondent likelihood to sign a petition ($\Pr(|Z| > |z|) = 0.269$). The same test of likelihood to make a donation found UK respondents were significantly more likely to take action ($\Pr(|T| > |t|) = 0.000$). The overall behavioural difference appears to have been driven by different levels of willingness to make a donation.

To create a pairwise analysis of all conditions for the dependent variable Behaviour the mean square error was calculated via an analysis of variance, and then a pairwise comparison was run. To reduce the probability of a Type 1 error a Tukey honestly significant difference post hoc test was used. The outputs are described in [Figure 4](#).

This more conservative analysis also finds that condition 3 (UN_ME) and condition 1 (RRes) were significantly different from the control (UN) in terms of behavioural response. Treatments 2 (RRes_ME) and 4 (UN_G) do not demonstrate a significant difference from control. Comparisons between treatments one to four do not show significant differences, even though some t scores are relatively high.

Discussion

The study is the first randomised experiment of advocacy communication frames on AMR. The results suggest that the way in which this important issue is communicated can change the way in which the issue is both perceived and acted upon. Despite the significant commentary on AMR in the academic and policy literature, there has been surprisingly little systematic research about optimal ways to advocate for change, or optimal processes by which to derive strong frames. The findings are compatible with the existing framing literature in the finding that strong frames make a small but significant difference to attitudinal measures. The study also indicates that applying findings from the behavioural science evidence base and the communications research process can develop strong

frames. The results of the study have implications for AMR advocates and indicate that a more ambitious research programme in this area would be beneficial.

This section considers the implications and potential drivers of the findings from the tests of the research hypotheses, and options for future research.

Treatment 1

The Reframing Resistance frame (RRes) had a significant impact versus control on behaviour. In terms of underlying mechanisms, processing fluency may have played a part. The UN control contains more complex language than the RRes treatment and more formal language such as 'food security' and 'Gross Domestic Product'. Post hoc testing of the UN and RRes frames found a Flesch reading ease score of 21.8 and 41.8, respectively. This places the UN frame score in the most difficult band of comprehension, and the RRes frame in the second most difficult. It is possible that the RRes treatments' primary advantage is in using more straightforward language. According to [Oppenheimer \(2006\)](#), "*Simpler writing is easier to process, and studies have demonstrated that processing fluency is associated with a variety of positive dimensions.*" (p.140). If the consultative aspects of the market research process leads to frames that are easier to read, this could drive positive behavioural outcomes. This raises the testable possibility that even more impactful frames could be created when processing fluency is an explicit consideration in frame development.

Treatments 2-4

We predicted that building frames based on transferable findings from behavioural science (such as a gain frame) would increase behavioural response versus the UN control frame. The results here were mixed - this approach is clearly not a panacea, but has potential nonetheless.

The addition of a gain frame (UN_G) was not more effective than a loss frame (UN Control) across most measures. Given the extensive literature concerning gain and loss frames it's a disappointing finding, that emphasises the importance of testing applications of theory in context rather than assuming transferability. One explanation may be that AMR is a bridge too far from the theoretical underpinnings of prospect theory, and the findings do not apply to more abstract issues.

Intriguingly, the gain frame did drive a robust difference on one attitudinal dependent variable, Usability, but not on the other, Importance. Why did respondents regard a gain frame to be more helpful and clearer than a loss frame? One difference is that Usability questions are focused less on the issue of AMR and more towards the words themselves. Perhaps gain frames are more likeable than loss frames, even if they are not necessarily more motivating.

The addition of the messenger and episodic devices identified as significant factors in a meta analytic review focusing on behavioural impact ([Gallagher & Updegraff, 2012](#)) did

. pwcompare condition, mcompare(tukey) sort effects						
Pairwise comparisons of marginal linear predictions						
Margins : asbalanced						
	Number of Comparisons					
condition	10					
	Contrast	Std. Err.	Tukey		Tukey	
			t	P> t	[95% Conf. Interval]	
condition						
4 vs 3	-.0714972	.0341209	-2.10	0.222	-.1646596	.0216651
4 vs 1	-.0586107	.0341209	-1.72	0.423	-.151773	.0345517
4 vs 2	-.032804	.0341872	-0.96	0.873	-.1261472	.0605392
2 vs 1	-.0258067	.0341431	-0.76	0.943	-.1190297	.0674163
3 vs 1	.0128866	.0340768	0.38	0.996	-.0801553	.1059285
3 vs 2	.0386933	.0341431	1.13	0.789	-.0545297	.1319163
4 vs 0	.0429704	.0341429	1.26	0.717	-.050252	.1361927
2 vs 0	.0757744	.0341651	2.22	0.173	-.0175086	.1690573
1 vs 0	.101581	.0340988	2.98	0.024	.008479	.194683
3 vs 0	.1144676	.0340988	3.36	0.007	.0213656	.2075696

Figure 4. Pairwise comparison of behaviour, by condition.

drive a significant increase in response for the UN_ME frame on the dependent variables versus control. However, the RRes_ME frame behavioural results were not statistically significant versus control. Results for RRes_ME were overall very similar to the RRes results. Given that the _ME additions to the control and RRes frames were identical, this difference is intriguing. It is possible that the RRes frame represents the upper bounds of the potential difference a frame can make, and significant improvements upon it are difficult. Alternatively, there is some evidence that a credible messenger is used as a heuristic to 'offload' the cognitive effort required to interpret information (Martin & Marks, 2019). Given that we know the UN frame is more difficult to process perhaps a credible source operated in a comparable way, whereas the more straightforward language of the RRes frame reduced the need for 'offloading'. Promisingly, this latter possibility could be tested experimentally utilising psycho-physiological measures (such as

measuring microsaccades of pupil dilation) to assess cognitive load during response to frames (Clark *et al.*, 2018).

Understanding the Understanding variable

This dependent variable was clearly non-discriminatory, and highlights a potentially problematic outcome of the market research process. Qualitative research led to a belief that understanding was a widespread problem, which is referenced extensively in the 'Reframing Resistance' report (Wellcome, 2019). Given that, on average, only 10% of respondents gave an incorrect answer, it is possible that the questions we used to test this understanding were not specific or challenging enough, or that the proximity of the question to a statement about AMR simply meant the correct answer was more salient. Alternatively, it is possible that public understanding of AMR is higher than the market research had indicated. This is potentially a very useful finding for advocates - indicating there may

be less of a need to educate public audiences about the nature of AMR than the conclusions of the O'Neill (2016) review suggest. It would be straightforward to more extensively test understanding levels, for example building on the methodology of Bholat *et al.* (2018) in their study of Central Bank communication.

The influence of independent variables

The difference in overall behavioural response between US and UK respondents was interesting. Given that there was no significant difference between the countries in willingness to sign a petition, (or indeed, the other dependent variables), the higher UK willingness to donate drove the overall finding. There may be different cultural norms over donating to charity which drove this difference. The fact that older respondents were more likely to think the issue of AMR was Important, but younger respondents were more likely to take a behavioural action is similarly counterintuitive - potentially reflecting different behavioural motivation or different levels of comfort with online donation. Both findings highlight the challenges in comparing the results of using a donation question to gauge behavioural response.

These identifiable differences in the study, driven by independent variables, highlight the importance of developing a more nuanced understanding of the impact of frames and other nudges upon segments of target populations. Costa & Kahn (2013) show, for example, that republicans in the US do not respond to social norm messaging about electricity consumption in the same way as democrats. Here, the limited demographic and supplementary information of this study hinders detailed analysis. It is possible that factors such as a value orientations, personality characteristics and recent experience of being prescribed or taking antibiotics make a difference to response. Certainly, the low R^2 scores for all the dependent variable regression models indicate a significant amount of variability is not accounted for by the independent variables captured in this study.

Limitations

The study approach described above has several limitations. Not least, it has the inherent artificiality of an online experiment completed by a survey panel. Paid online survey respondents are an imperfect representation of target audiences, and whilst the selection criteria strove to maximise the relevance of the respondents it is by no means a perfect sample. Larger, more representative samples, additional countries and mixed methods including field experimentation would increase the robustness of the findings.

Accuracy of behavioural proxies

Whilst the measures of a signed petition and percentage donation were a sincere attempt to move beyond the more usual attitudinal and stated preference questions that dominate the literature and to better understand behavioural impact, it is by no means certain they meaningfully correspond with real-world behaviour. Future research would ideally be conducted in more authentic settings - for example, if different treatments were

placed as 'advertorials' in online newspapers with an option to sign up to a real petition would a framing effect also be apparent? Moreover, this research design creates no understanding of the longitudinal impact, such as how quickly framing effects dissipate, or what level of information 'dosage'/frame repetition is optimal.

Underlying mechanisms and treatment comparability

Ideally, this study would have also explored underlying mechanisms behind frame strength. As Aarøe (2011) points out, "*an impressive body of research has investigated framing effects, we still only have very limited knowledge of the factors that shape frame strength*" (p. 207). The transferability of findings is enhanced by a greater understanding of psychological mechanisms, and further research around framing AMR would also ideally explore considerations such as the role of self-efficacy or processing fluency in creating frame strength.

It could be argued that the UN version is at an inherent disadvantage compared to the new treatments which have new forms of information such as additional narrative. The UN version may also have been designed for use by a variety of audiences - the public, industry and policymakers - rather than being optimised for any particular audience or function. However, the control is a real-world example that was authored with at least implicit assumptions about utility. Therefore, following the logic of Thaler & Sunstein (2003), it would be perverse to argue that explicitly considering choice architecture invalidates any improvement in the treatment conditions.

Generalisability beyond study cohort

Participants were recruited to broadly reflect the audience that the original Reframing Resistance study envisaged as likely recipients of messaging for media advocacy. This, together with the biases resulting from the use of an online platform to perform the testing, mean that it cannot be assumed that the findings would automatically generalise to other populations. The theories that inform the work would suggest broad applicability, especially in Western settings - but the extent to which the same principles of framing can be applied across languages and cultures remains to be seen.

Future studies

This study raises a number of further potential research questions. For AMR advocates, the real-world relevance of the findings needs to be tested. Given that some AMR frames had an effect in a relatively large online panel survey, can this finding be replicated in a real-life setting? Kahan (2013), argues eloquently for the need for climate communications to be evidence based 'all the way down' and for communicators and scholars to work more effectively together. We believe this argument holds with respect to the lower profile issue of AMR. Whilst the work on using social norms to reduce unneeded frontline prescribing in the UK (Hallsworth *et al.*, 2016) is useful, there is a need for more, and more ambitious, work considering AMR. In particular, increasing our understanding of how to optimise the strength of effect and persistence of framing,

alongside the influence of cultural heterogeneity, values, experience and other predictors, would enable progressively more targeted and effective communication.

A range of broader questions for anyone releasing communications are also raised. Should organisations and agencies evolve their processes? Are there ways to optimise and predict the impacts of different frames more quickly? The potential for better empirical results is exciting. For scholars, potential theoretical questions include building our understanding of the underpinning mechanisms behind strong and weak frames, exploring the constraints of prospect theory and utilising the opportunities that new technology presents to test at speed and at scale. And, given the well-documented issue of positive result publication bias (Olson *et al.*, 2002), caution and studying the replicability of findings is advisable

On a cautionary note, the absolute difference in the means of response between the treatments and control across Usability, Importance and Behaviour were all relatively small. For example, the absolute difference in behavioural response between control and the strongest treatment (UN_ME) was only 19%. This change is welcome, and could plausibly make a difference to advocacy outcomes, but it is not in and of itself sufficient. Addressing systemic issues like AMR requires a multifaceted response, and we hope this study plays a small part in contributing to that.

Conclusion

AMR is a significant global health challenge for policy-makers. There is a need for approaches that effectively, and cost-effectively, support advocates. This experiment shows that strong frames can make a small but significant difference to changing people's attitudes about AMR and, potentially, their behaviour. Whilst extensively used in political communication and health promotion, perhaps the power of framing could be applied more systematically outside of these fields.

The finding that the Reframing Resistance project created a frame that changed behaviour as well as attitudes is noteworthy. A very different, less labour intensive, process of developing frames by transferring learnings from the evidence base also created one comparably strong frame. This indicates an alternative to common practice may be of use, opening up a new option for practitioners seeking to create strong frames. The fact that two other frames developed in this way were not significant improvements highlights that evidence is only sometimes transferable and reinforces the value of systematic,

randomised testing in framing research. As Sally Davies, the UK Special Envoy on AMR, says in the preface of the Reframing Resistance report:

“By basing our communications in evidence – as we would be required to do for any other intervention we develop – we can better unlock the huge potential we have as advocates to more effectively galvanise support for the antimicrobial resistance cause and stimulate action.” (Wellcome, 2019, p. 4)

There is a clear opportunity to expand research in this area, for example by better understanding the mechanisms underpinning framing effects in this context, exploring the replicability of these findings in different countries, assessing longitudinal impact or assessing the impact of different AMR frames upon real world behaviour.

Data availability

Underlying data

Figshare: Framing Survey Raw Analysis and Summary Data. <https://doi.org/10.6084/m9.figshare.14186441.v2> (Metcalf *et al.*, 2021a).

This project contains the following underlying data:

- Framing Survey Summary Data.xlsx
- Framing Survey Raw Analysis Data.xlsx

Extended data

Figshare: Framing for AMR - extended data inc appendixes. <https://doi.org/10.6084/m9.figshare.14186513> (Metcalf *et al.*, 2021b).

This project contains the following extended data:

- Appendix 1: UK Survey questions
- Appendix 2: Copy of stata commands
- Appendix 3: Ethics assessment

Data are available under the terms of the [Creative Commons Attribution 4.0 International license](#) (CC-BY 4.0).

Acknowledgements

The authors would like to thank Tim Jinks, Paul Dolan and Louisa Tribe for their input delivering the framing project, and Sian Williams for internal review of the manuscript

References

- Aarøe L: **Investigating Frame Strength: The Case of Episodic and Thematic Frames.** *Political Communication*. 2011; **28**(2): 207- 226.
[Publisher Full Text](#)
- Akoglu H: **User's guide to correlation coefficients.** *Turk J Emerg Med*. 2018; **18**(3): 91-93.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

- Babakus E, Mangold WG: **Adapting the SERVQUAL Scale to Hospital Services: An Empirical Investigation.** *Health Serv Res*. 1992; **26**(6): 767.
[PubMed Abstract](#) | [Free Full Text](#)
- Bholat D, Broughton N, Parker A, *et al.*: **Staff Working Paper No . 750 Enhancing central bank communications with behavioural insights.** 2018.
[Reference Source](#)

Busby E, Flynn DJ, Druckman JN: **Studying Framing Effects: Existing Research And Lingering Questions**. 2018.
[Publisher Full Text](#)

Buttle F: **SERVQUAL: Review, critique, research agenda**. *Eur J Mark*. 1996; 30(1): 8–32.
[Publisher Full Text](#)

Costa DL, Kahn ME: **Energy conservation “nudge” and environmentalist ideology: evidence from a randomized residential field experiment**. *J Europ Econ Assoc*. 2013; 11(3): 680–702.
[Publisher Full Text](#)

Clark KR, Leslie KR, Garcia-Garcia M, *et al.*: **How advertisers can keep mobile users engaged and reduce video-ad blocking: Best practices for video-ad placement and delivery based on consumer neuroscience measures**. *J Advert Res*. 2018; 58(3): 311–325.
[Publisher Full Text](#)

Clift C: **Review Of Progress On Antimicrobial Resistance : Background And Analysis**. 2019.
[Reference Source](#)

Gallagher KM, Updegraff JA: **Health Message Framing Effects On Attitudes, Intentions, And Behavior: A Meta-analytic Review**. *Ann Behav Med*. 2012; 43(1): 101–116.
[PubMed Abstract](#) | [Publisher Full Text](#)

Hair JF, Black WC, Babin BJ, *et al.*: **Multivariate data analysis (6th ed.)**. NJ: Prentice Hall. 2006.
[Reference Source](#)

Hallsworth M, Chadborn T, Sallis A, *et al.*: **Provision of social norm feedback to high prescribers of antibiotics in general practice: a pragmatic national randomised controlled trial**. *Lancet*. 2016; 387(10029): 1743–52.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Holm S: **A Simple Sequentially Rejective Multiple Test Procedure**. *Scand J Statist*. 1979; 6(2): 65–70.
[Publisher Full Text](#)

Huttner B, Saam M, Moja L, *et al.*: **How To Improve Antibiotic Awareness Campaigns : Findings Of A Who Global Survey**. *BMJ Glob Health*. 2019; 4(3): e001239.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Kahan DM: **Making climate-science communication evidence-based: All the way down**. *Culture, Politics and Climate Change: How Information Shapes Our Common Future*. 2013; 203–220.
[Publisher Full Text](#)

Martin S, Marks J: **Messengers Who We Listen To Who We Don't, And Why**. Random House Publishing, 2019.
[Reference Source](#)

Maxwell SE, Lau MY, Howard GS: **Is psychology suffering from a replication crisis? What does “failure to replicate” really mean?** *Am Psychol*. 2015; 70(6): 487–98.
[PubMed Abstract](#) | [Publisher Full Text](#)

Metcalfe D, Aldalooj E, Thakur M, *et al.*: **Framing Survey Raw Analysis and Summary Data**. *figshare*. Dataset. 2021a.
<http://www.doi.org/10.6084/m9.figshare.14186441.v2>

Metcalfe D, Aldalooj E, Thakur M, *et al.*: **Framing for AMR - extended data inc appendixes**. *figshare*. Online resource. 2021b.
<http://www.doi.org/10.6084/m9.figshare.14186513>

Olson CM, Rennie D, Cook D, *et al.*: **Publication Bias in Editorial Decision Making**. *JAMA*. 2002; 287(21): 2825–2828.
[PubMed Abstract](#) | [Publisher Full Text](#)

O'Neill J: **Tackling Drug-resistant Infections Globally: Final Report And Recommendations**. 2016.
[Reference Source](#)

Opiyo N, Shepperd S, Musila N, *et al.*: **Comparison of Alternative Evidence Summary and Presentation Formats in Clinical Guideline Development: A Mixed-Method Study**. *PLoS One*. 2013; 8(1): e55067.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Oppenheimer DM: **Consequences of Erudite Vernacular Utilized Irrespective of Necessity : Problems with Using Long Words Needlessly**. 2006; 20(2): 139–156.
[Publisher Full Text](#)

Rosenbaum SE, Glenton C, Oxman AD: **Summary-of-findings Tables In Cochrane Reviews Improved Understanding And Rapid Retrieval Of Key Information**. *J Clin Epidemiol*. 2010; 63(6): 620–626.
[PubMed Abstract](#) | [Publisher Full Text](#)

Rothman AJ, Salovey P: **Shaping perceptions to motivate healthy behavior: the role of message framing**. *Psychol Bull*. 1997; 121(1): 3–19.
[PubMed Abstract](#) | [Publisher Full Text](#)

Santesso N, Rader T, Nilsen ES, *et al.*: **A summary to communicate evidence from systematic reviews to the public improved understanding and accessibility of information: A randomized controlled trial**. *J Clin Epidemiol*. 2015; 68(2): 182–190.
[PubMed Abstract](#) | [Publisher Full Text](#)

Spence A, Pidgeon N: **Framing and communicating climate change: The effects of distance and outcome frame manipulations**. *Glob Environ Change*. 2010; 20(4): 656–667.
[Publisher Full Text](#)

Thaler RH, Sunstein CR: **Libertarian Paternalism**. *American Economic Review*. 2003; 93(2): 175–179.
[Publisher Full Text](#)

Wellcome: **Reframing resistance: How to communicate about antimicrobial resistance effectively**. 2019.
[Reference Source](#)

World Health Organisation: **2019 antibacterial agents in clinical development: an analysis of the antibacterial clinical development pipeline**. 2019.
[Reference Source](#)

Open Peer Review

Current Peer Review Status:

Version 1

Reviewer Report 17 April 2023

<https://doi.org/10.21956/wellcomeopenres.18443.r55622>

© 2023 Otto S. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Simon J G Otto

University of Alberta, Edmonton, Alberta, Canada

This is an interesting study that sought to assess the impacts of different framings of antimicrobial resistance narratives on attitudinal and behavioural responses related to AMR. Given the challenges we have experienced globally in terms of making headway on the AMR crisis, it is important that we better understand how to frame scientific information to general public and to health practitioners in order to influence real change with measurable improvements.

However, I have major concerns with how the statistical analysis was conducted for the majority of this study. The analysis is not described in enough detail to make it clear to the reader what was done where. More importantly, many of the major outcomes, specifically the behavioural responses that were measured using a 5-point Likert scale, were analyzed using the outcome as a continuous variable. They frequently report the mean of these outcomes per group.

I realize that there is precedent for reporting ordinal Likert data such as this using parametric methods for continuous data, but this typically requires that the data are normally distributed, and the approach is still not universally agreed upon. The authors present no descriptive or other results to support this. Further, they often report means of these outcomes where medians are likely more appropriate. See: Sullivan, G. & Artino Jr., A. R. (2013)¹.

They used linear regression for these outcomes where ordinal logistic regression is more appropriate if the normality assumption cannot be confirmed. See: Liang J, Bi G, Zhan C (2020)².

In my opinion, these data should be reported as proportions of respondents that selected each of the five categories and then analyzed using ordinal logistic regression. I have reviewed the Introduction to the Results, including Tables/Figures, but I did not review the Discussion in detail because I am not comfortable assessing the conclusions of the analytical approach. In particular, I take this quote from the Discussion, "On a cautionary note, the absolute difference in the means of the response between the treatments and control across Usability, Importance and Behaviour were all relatively small." I do not believe this is how these results should be compared. There may be stark differences in the proportions of respondents that selected any of 1-5 for each outcome

in each of the treatment and control groups that result in similar mean responses.

Specific comments are included herein.

Methods/Results:

- Page 6: "Variables of interest" paragraph. The authors state the concerns with measuring the impact of a treatment on a behavioural response and then provide the behavioural responses that they decided to use, but did not justify why they think these responses will not be impacted by the concerns they mention. It would be nice if they could provide some justification for this (and perhaps cite another example where similar responses were used and shown to be a valid measurement).
- Page 6: the last paragraph at the bottom left + Tables 1-3: the description of the attitudinal responses. The authors describe how they asked participants to provide Likert rankings for the three variables (Usability, Importance, and Understanding). They took a mean of each respondent's rankings for Usability and Importance to provide a value for a combined response (though I cannot see where they used this). The mean Likert responses for these three variables are provided in Table 1. This is not a good approach. Results of Likert rankings should be presented and analyzed as the proportion of responses in each Likert category (1-5). Summarizing a mean response of a numerical Likert scale applies meaning to the numbers where one may/does not exist.
- Further, Table 1 also reports a mean Behaviour score (presumably the mean between 0 and 1 for a yes/no answer as to whether they would sign the petition or donate to an AMR charity). Again, it is not appropriate to report these as means. They should be analyzed as proportions of treatment vs control groups that answered yes or no.
- Page 6 Sample Size and Participant Recruitment: linked to my concerns above. When the authors state that they wanted to estimate a mean difference in effect size of 10%, was it based on a 10% difference in mean scores for these variables where they took the mean of a categorical ordinal response? If so, I question the validity of the sample size calculation.
- Page 7 Statistical methods: the authors provide a short description of their analysis was pair-wise comparison. Their Stata output also shows multivariable models. The authors do not state that they accounted for clustering in their data by including a fixed effect for country, while it appears that they did to this, which is good. Explicitly stating this would be good. They also provide no details as to how they built the multivariable regression models (see more specific comments below).
- Further to my comments above regarding treatment ordinal variables as numerical outcomes, it is not easily discerned from the description of the statistical methods or presentation of results how the outcomes were analyzed. Page 7 does state that the differences between the control and four treatment groups were analyzed via the creation of aggregated categories. It would be helpful if they could describe how they did this. Did they aggregate the Likert categories for the variables measured on the 5-point scale? Presumably, the yes/no variables are simply treated as such in a logistic model. They go on to state that they did assess dependent variables as continuous variables. Based on my understanding of how they described their outcomes, I do not think this is a valid way to

analyzed these ordinal outcomes of the Likert or yes/no scale for the reasons stated. For example, Figure 2 shows a linear regression model on the mean Likert score for Importance (sumean).

- One can deduce from the Stata output which multivariable models were run (one linear and one logistic model). However, the shorthand used for the outcome and explanatory variable names should be provided in the figure caption to make them easier to interpret. Some of these explanations are provided in the text, but including them in the figure caption would help.
- Page 8 (right column) + Tables 1-3: the reporting of means for their outcome measures is not valid for reasons stated above. I want to see the proportions of respondents who answered 1-5 or yes/no for the described outcomes.
- Results section (pages 8-9): all of this text discusses the statistical results of the analyses that treated these ordinal outcomes as continuous variables. I have not reviewed it in detail because I do not feel this analysis is valid.
- Page 10 + Figure 3: I was glad to see that they treated Behaviour as a dichotomous outcome and used logistic regression to analyze these data. However, I do not completely understand their statement "The likelihood ratio chi-square... p-value of <0.001 indicates the model as a whole fits significantly better than an empty model (with no predictors)." Interpreted literally, this means they tested the difference between a model with all of their predictors and one with none and got this result. That is not very helpful. I want to know how they built the multivariable model and determined what variables to include and exclude and why. What was the LR test p-value for inclusion of the treatments? These are presumably the variable listed as "Condition" labeled as 1-4, but the authors do not tell us this in the Figure caption, which should be added.
- Figure 3. The authors state that there is a significant difference between countries related to the outcome of Behaviour. Did they consider any possibility of interaction between variables like Country (or others) and the Treatment ("Condition") variable? If they considered interactions, they should tell us how they did this and the results. If interactions were not evaluated, then this should be reconsidered. How did the authors consider potential confounding variables? This is also not mentioned and ties back to the lack of any discussion of how they built their multivariable regression models.
- Page 6 (final paragraph bottom right): the authors state that the survey excluded those who failed an attention test or never consumed any of the listed media channels. Did the authors consider whether this created any opportunity for selection bias, whether differential or non-differential. And the potential impact this could have on their results? They provided a justification for why it was done, but I suggest that this consideration could be added to the Discussion.

More General Comments:

- Page 3: Research question "The research question for this study was: do frames developed using de novo communications research or the behavioural evidence base change attitudinal and behavioural response, or increase understanding?" I suggest that this should be more specific. The hypothesis stated previously is clear, but if read alone, this

research question does not include any language about AMR. Suggestion: "The research question for this study was: do AMR frames developed using de novo communications research or the behavioural evidence base change attitudinal and behavioural response towards, or increase understanding of AMR?"

- Methods (Page 3): "Five concise AMR treatment frames were created – structured narratives that aim to communicate AMR risks to promote certain reader attitudinal shifts and/or actions. All were between 175 and 250 words and are reproduced in full in Figure 1 and Figure 2." I do not believe that Figures 1 and 2 represent the AMR control and treatment frames. From what I can see, Figure 1 is the methodology flow diagram and Figure 2 is a regression result. I believe that the AMR control and treatment frames are represented in Box 1 (Control) and Box 2 (Treatment Frames 1-4). Box 2 is split across multiple locations, so it may be more appropriate to label each as Boxes 2-5, each one containing one of Treatment Frames 1-4. This will make things clearer and can be referred to in the text of this quoted paragraph to be more clear. The following text in the methods seems to be correct in this regard.
- Page 5: text between the Boxes (before the heading "The messenger and episodic treatments". Because these small blocks of text are between the boxes, my first thought was that they pertained to the Treatments listed in the nearest boxes. Upon reading further, this appeared to not be the case (they all pertain to Treatment 1 on Page 3). I found this confusing and had to spend some time to understand this. Considering a reorganization of the text vs boxes or clearly making sure each block of text mentions Treatment 1 to be clear would improve the clarity and ease of understanding for the reader.
- Page 5: Likewise, for the paragraph with the heading "The messenger and episodic treatments" – add "Treatments 2 and 3" to the heading to be clear. One can deduce that this is the case from the text, but this could be made clearer to the reader. It could also be explicitly stated in this paragraph that Treatments 2 and 3 specifically combining the episodic focus with either the "Reframing Resistance" or "UN" messengers.
- There is also variable text font size in the main text of pages 3 and 5.
- Page 6: "Gain frame treatment" heading – similar comment – add "Treatment 4" to this heading +/- text in the paragraph to be crystal clear that it is referring to Treatment 4.

References

1. Sullivan GM, Artino AR: Analyzing and interpreting data from likert-type scales. *J Grad Med Educ.* 2013; **5** (4): 541-2 [PubMed Abstract](#) | [Publisher Full Text](#)
2. Liang J, Bi G, Zhan C: Multinomial and ordinal Logistic regression analyses with multi-categorical variables using R. *Annals of Translational Medicine.* 2020; **8** (16). [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

No

Is the study design appropriate and is the work technically sound?

No

Are sufficient details of methods and analysis provided to allow replication by others?

No

If applicable, is the statistical analysis and its interpretation appropriate?

No

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

No

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Antimicrobial resistance, antimicrobial stewardship, integrated surveillance of antimicrobial resistance and antimicrobial use, One Health, veterinary medicine, epidemiology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to state that I do not consider it to be of an acceptable scientific standard, for reasons outlined above.

Reviewer Report 15 June 2021

<https://doi.org/10.21956/wellcomeopenres.18443.r44106>

© 2021 Mattar C. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Caline Mattar

Division of Infectious Diseases, Department of Internal Medicine, Washington University in St Louis, St. Louis, MO, USA

This is a very important article and trial addressing framing of AMR/Drug-resistant infections. The first Wellcome Report on reframing resistance was an essential piece to stimulate thinking about communication around AMR, but not much change was seen by ways of AMR communications especially at the UN and WHO level. This article is very thoughtfully designed to attempt to provide an evidence base around the best approach to framing. Overall, the article brings good knowledge and actionable points that can be used for advocacy messages on AMR. Overall, the structure of the article and layout is unusual for a scientific article. I would remove the figure 2 stata output and either include it at the end in a appendix or supplemental materials. This figure needs text to explain the content, variables and utility.

With regards to the proxies for engagement, while signing a petition may be more generalizable,

charitable donations for specific causes such as AMR, climate change medical research tends to be rooted in local culture with significant variability in uptake within the United States. For future studies, in addition to attitudes towards AMR, consider assessing consumer choices when it comes to poultry, meat etc and antibiotics, in addition to questions around attitudes towards being the receiver of antibiotic prescriptions for minor ailments for ex, procurement etc which would be more relevant for a broader context. Excellent that the full stata code is shared. For the results, I recommend having a summary table especially for the demographics and later for the results. the current methods of using separate tables to report on the individual comparisons makes it difficult for the reader to have the full picture. I would recommend all limitations to go under the limitations section. Some are already mentioned in the methods section.

Overall great article, some minor changes to the writing can make the article stronger and more accessible.

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

I cannot comment. A qualified statistician is required.

Are all the source data underlying the results available to ensure full reproducibility?

Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Antimicrobial resistance, dissemination and implementation of AMR interventions in low and middle income countries, clinical trials

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.