Appendix

Appendix A: Deviation from Pre-registration 2
Appendix B: Stimulus Transcripts 13
Appendix C: Ethical Considerations 14
Appendix D: Power Analysis 15
Appendix E: Manipulation Check 16
Appendix F: Measurements 18
Appendix G: Regression Tables 21
Appendix H: Detailed Analysis Report and Robustness Checks 24

Appendix A: Deviation from Pre-registration

We deviate from our pre-registration as we (1) changed the phrasing three of our hypothesis and one research question, (2) formulated a hypothesis instead of a research question, (3) included an exploratory part, (4) excluded a pre-registered part from the manuscript and (5) conducted additional exploratory analyses. For transparency, these deviations are reported here in detail.

1. Formulation of hypothesis and research questions

Three of our hypotheses (H1a, H1b, H2) are formulated differently in the manuscript than in our pre-registration to enhance clarity. **Importantly, the expectations remain the same**.

H1a was pre-registered as:
Participants exposed to real visual media find these more credible compared to those exposed to manipulated visual media (decontextualization, cheapfake, deepfake).
In the manuscript, the hypothesis is now formulated as follows:
An authentic video is evaluated as more credible compared to a manipulated video (decontextualization, cheapfake, deepfake).

H1b was pre-registered as:
A manipulated video using decontextualization is assessed as more credible compared to a cheapfake or deepfake.
In the manuscript, the hypothesis is formulated as follows:
A decontextualized video is evaluated as more credible compared to a cheapfake or deepfake.

H2 was pre-registered as:
Credibility evaluations of real and manipulated visual media (decontextualization, cheapfake, deepfake) are moderated by issue attitudes in such a way that stronger anti-immigration attitudes lead to higher credibility evaluations for all videos compared to weaker anti-immigration attitudes.
In the manuscript, the hypothesis is formulated as follows:
Credibility evaluations of manipulated videos (decontextualization, cheapfake, deepfake) are moderated by issue agreement in such a way that stronger issue agreement leads to higher credibility evaluations for all videos compared to weaker issue agreement.

RQ1 in the manuscript was pre-registered as RQ3. The pre-registration reads as:
Is there a difference in the degree to which misperceptions are caused by different forms of manipulated visual media?
In the manuscript, the research question reads as:
Is there a difference in the degree to which different forms of manipulated videos (decontextualization, cheapfake, deepfake) cause misperceptions?

2. Hypothesis instead of research question

We decided to formulate a hypothesis (H1c) instead of a research question (RQ1 in the pre-registration) due to an unforeseen event, which, according to Lakens (2024), is a valid reason to deviate from a pre-registration. While working on this manuscript, a study was published that tested the difference between a cheapfake and a deepfake in terms of perceived credibility (Hameleers, 2024). To ensure our paper is aligned with the current state-of-the-art in visual disinformation research, we revised RQ1: *"Is there a difference in the credibility evaluation of a cheapfake versus a deepfake?"* into the following hypothesis: *"A cheapfake is evaluated as more credible compared to a deepfake (H1c)."*

This adjustment does not alter our analysis, as we would have conducted the same test (one-way ANOVA) to answer the research question. However, this modification positions our findings within the context of the latest research and enhances the contribution of our paper. By updating our approach to reflect recent advancements, we ensure our study remains relevant and provides valuable insights into the credibility evaluation of visual disinformation.

Lakens, D. (2024). When and How to Deviate from a Preregistration. *Collabra: Psychology*, *10*(1). https://doi.org/10.1525/collabra.117094

3. Inclusion of exploratory part

In our pre-registration, we had an exploratory part addressing the effects of visual disinformation on attitudes towards the portrayed politician. Because this part was pre-registered as an exploratory analysis, at the time, we only posed a research question, RQ4: *"Is there a difference in attitude towards the politician after exposure to real visual media vs.*

manipulated visual media (decontextualization, cheapfake, deepfake)?" However, as we decided to integrate this analysis into our main findings, we also reviewed relevant literature and formulated a hypothesis in line with state-of-the-art research (Barari et al., 2021; Dobber et al., 2021). Thus, the manuscript now presents the following research question (RQ2) and hypothesis (H3):

RQ2: Is there a difference in the degree to which different forms of manipulated videos (decontextualization, cheapfake, deepfake) affect perceptions of the portrayed politician?
H3: A deepfake leads to more negative perceptions of the portrayed politician compared to an authentic video.

4. Exclusion of pre-registered part

In the pre-registration, we outlined several expectations concerning the impact of visual disinformation on social media engagement intentions, as well as the moderating role of anti-immigration attitudes. These expectations include:

H3: Exposure to real visual media leads to higher engagement intentions than exposure to manipulated visual media (decontextualization, cheapfake, deepfake).
RQ2: Is there a difference in elicited engagement intentions between the different forms of manipulated visual media (decontextualization, cheap fake, or deepfake)?
H4: The intention to engage with real and manipulated visual media (decontextualization, cheapfake, deepfake) on social media is moderated by issue attitudes in such a way that stronger anti-immigration attitudes lead to higher intent to engage compared to weaker anti-immigration attitudes.

We decided not to include these in our manuscript. However, for transparency, the findings are reported in detail below.

Measurement engagement intention: (1 = Stimme überhaupt nicht zu; 7 = Stimme voll und ganz zu)
German: *Inwieweit stimmen Sie den folgenden Aussagen zu? Diesen Post...*

... sollte man mit anderen teilen.
... würde ich teilen.
... würde ich anderen empfehlen.

... würde ich kommentieren.

... würde ich liken oder anderweitig darauf reagieren.

English: *To what extent do you agree with the following statements?*

This post should be shared with others.

I would like to share this post.

I would recommend this post to others.

I would comment on this post.

*I would like this post or react in some other way***.**

Analysis:

A one-way ANOVA revealed a non-significant main effect of the condition on engagement intention F (1, 800) = 0.783, p = 0.376. Authentic (M = 2.45, SD = 1.63), Decontextualized (M = 2.22, SD = 1.48), Cheapfake (M = 2.33, SD = 1.47), Deepfake (M = 2.26, SD = 1.61). This was the same for Kruskal-Wallis-Test H (3) = 3.34, p = 0.34. We also performed OLS-based linear regressions to test whether anti-immigration attitudes would moderate this effect R^2 = -0.00, F (7, 794) = 0.59, p = 0.77. However, we did not find any significant two-way interaction effects between engagement intentions and anti-immigration attitudes for our conditions.

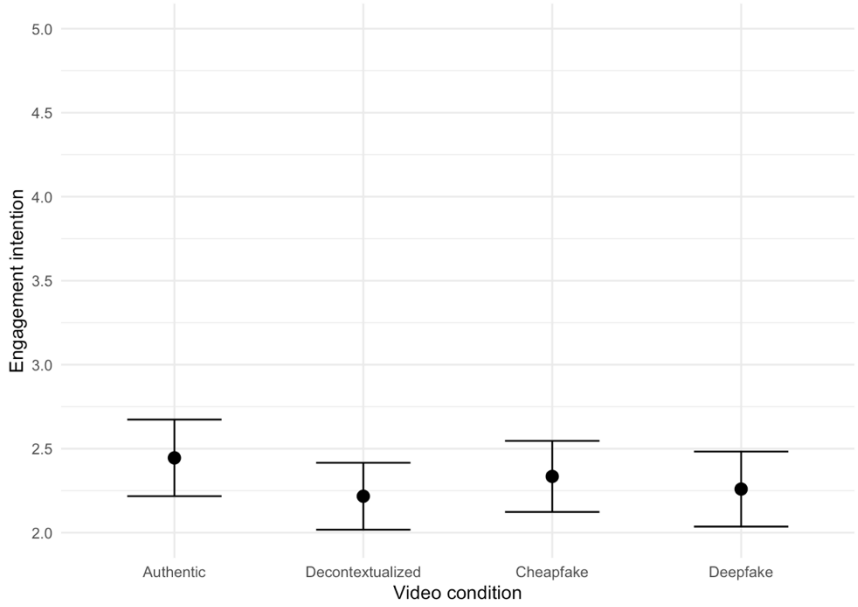


Figure 1. Interval plot for the variable engagement intention.

Table 1. Direct effect on engagement intentions per condition, including anti-immigration attitudes.

	Dependent variable: **Engagement intentions**			
	(1)	(2)	(3)	(4)
decontextualized	-0.129			-0.229
	(0.124)			(0.153)
cheapfake		0.030		-0.110
		(0.129)		(0.158)
deepfake			-0.070	-0.188
			(0.126)	(0.155)
anti-imm. attitudes				0.022
				(0.038)
Constant	2.346***	2.305***	2.329***	2.348***
	(0.064)	(0.062)	(0.063)	(0.198)
Observations	802	802	802	802
R^2	0.001	0.0001	0.0004	0.004
Adjusted R^2	0.0001	-0.001	-0.001	-0.001
Residual Std. Error	1.548 (df = 800)	1.549 (df = 800)	1.549 (df = 800)	1.550 (df = 797)
F Statistic	1.096 (df = 1; 800)	0.054 (df = 1; 800)	0.308 (df = 1; 800)	0.726 (df = 4; 797)

Note: OLS-based linear regressions with k-1 binary dummy variables. The authentic condition serves as the constant/intercept. * *p<0.1;*** ***p<0.05;*** *** p<0.01

Table 2. Individual interactions of anti-immigration attitudes with engagement intentions per condition.

	Dependent variable: **Engagement intentions**		
	(1)	(2)	(3)
decon. x anti-imm. attitudes	-0.061		
	(0.084)		
cheapfake x anti-imm. attitudes		0.094	
		(0.093)	
deepfake x anti-imm. attitudes			-0.020
			(0.088)
Constant	2.178***	2.296***	2.212***
	(0.205)	(0.197)	(0.200)
Observations	802	802	802
R^2	0.002	0.002	0.001
Adjusted R^2	-0.001	-0.002	-0.003
Residual Std. Error (df = 798)	1.549	1.550	1.551
F Statistic (df = 3; 798)	0.644	0.466	0.233

Note: OLS-based linear regressions with k-1 binary dummy variables. The authentic condition serves as the constant/intercept. * *p<0.1;*** ***p<0.05;*** **** p<0.01

Table 3. Complete model of all interactions with anti-immigration attitudes on engagement intentions.

	Dependent variable: **Engagement intentions**
decontextualized	-0.019 (0.475)
cheapfake	-0.412 (0.507)
deepfake	-0.106 (0.492)
decon. x anti-imm. attitudes	-0.048 (0.104)
cheapfake x anti-imm. attitudes	0.070 (0.111)
deepfake x anti-imm. attitudes	-0.019 (0.107)
Constant	2.333*** (0.342)
Observations	802
R^2	0.005
Adjusted R^2	-0.004
Residual Std. Error	1.551 (df = 794)
F Statistic	0.588 (df = 7; 794)
Note:	*p<0.1; **p<0.05; ***p<0.01

5. Additional exploratory analysis

Moderation:

To further examine potential interaction effects with our pre-treatment variable *anti-immigration attitudes* and our dependent variable *misperception* and *attitude towards the politician*, we conducted OLS-based linear regressions using dummy variables with the authentic/control condition serving as the reference group. The complete models are reported in the tables below.

Table 4. Complete model of all interactions with anti-immigration attitudes on misperception regarding von der Leyen.

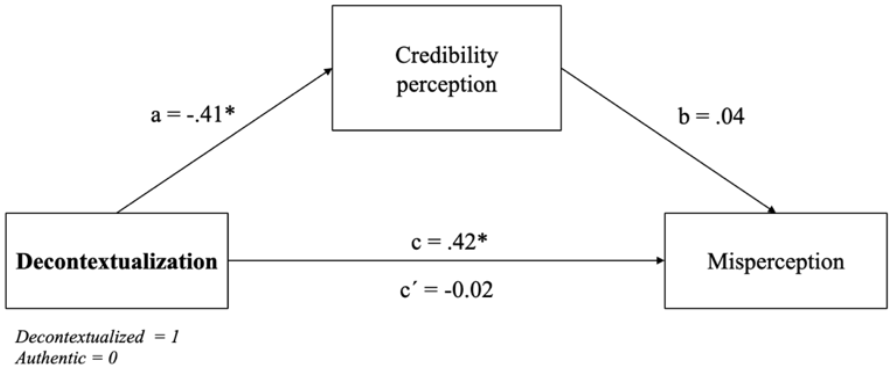
	Dependent variable:
	Misperception (politician)
decontextualized	-0.011
	(0.574)
cheapfake	0.537
	(0.613)
deepfake	-0.242
	(0.595)
decon. x anti-imm. attitudes	0.095
	(0.125)
cheapfake x anti-imm. attitudes	0.057
	(0.135)
deepfake x anti-imm. attitudes	0.181
	(0.129)
Constant	3.951***
	(0.413)
Observations	802
R^2	0.025
Adjusted R^2	0.016
Residual Std. Error	1.874 (df = 794)
F Statistic	2.919*** (df = 7; 794)
Note:	*p<0.1;* **p<0.05;** p<0.01

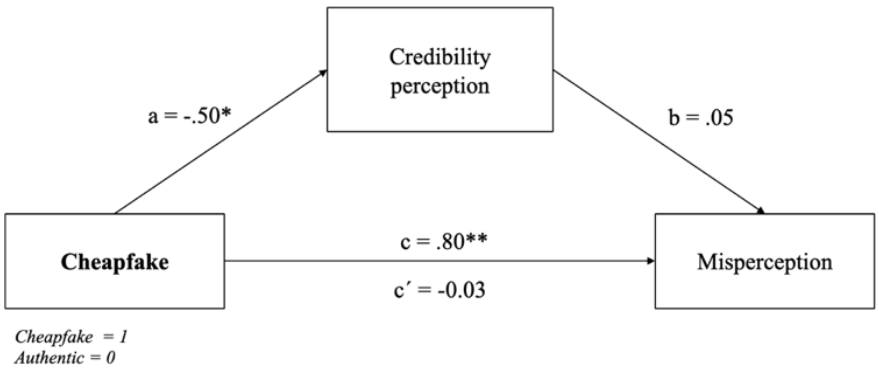
Table 5. Complete model of all interactions with anti-immigration attitudes on attitudes towards von der Leyen.

	Dependent variable:
	Attitude towards politician
decontextualized	0.268
	(0.496)
cheapfake	0.192
	(0.530)
deepfake	0.562
	(0.514)
decon. x anti-imm. attitudes	-0.109
	(0.108)
cheapfake x anti-imm. attitudes	-0.118
	(0.116)
deepfake x anti-imm. attitudes	-0.237**
	(0.111)
Constant	5.456***
	(0.357)
Observations	802
R^2	0.146
Adjusted R^2	0.139
Residual Std. Error	1.620 (df = 794)
F Statistic	19.457*** (df = 7; 794)
Note:	*p<0.1;* ***p<0.05;*** p<0.01

Mediation:

To lend further support to our claim that credibility perceptions and misperceptions can arise independently, we conducted an additional mediation analysis using Hayes' PROCESS Model 4, with 5,000 bootstrap resamples and a 95% confidence interval (CI). In this analysis, we tested credibility as a mediator for the misperception item, comparing each disinformation condition to the control group. We argue that mediation is conceptually plausible, as credibility is traditionally considered a mediating variable—although it has rarely been examined as such in deepfake research. Additionally, since we measured credibility before the misperception item in our questionnaire, the temporal sequence aligns with the mediation assumptions. Our mediation analysis results show that while the disinformation conditions (compared to the control group) significantly decrease credibility and increase misperceptions, there is no significant indirect effect. In other words, no mediation occurs, suggesting that credibility perceptions and misperceptions develop independently. To illustrate these findings, we have visualized the mediation models below.





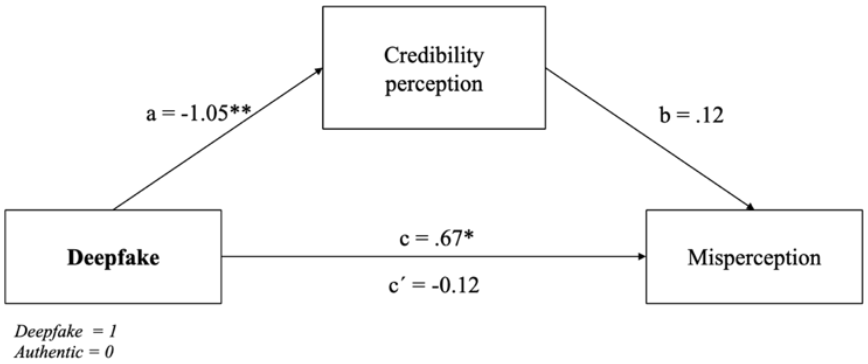


Figure 2. Mediation models

The following table gives an overview of all direct and indirect effects per condition group.

Table 6. Mediation analysis

Predictor (0=Control, 1=Condition)	Outcome	R^2	F (df1, df2)	p (F-test)	Effect Type	B	SE	t	p	95% CI (LL, UL)
Decon-textualized	Credibility	.0169	7.08 (1, 411)	.0081	Direct	-0.4142	0.1557	-2.661	.0081	[-0.7202, -0.1082]
	Mis-perception	.0138	2.86 (2, 410)	.0583	Direct	0.4185	0.1777	2.355	.0190	[0.0692, 0.7678]
					Indirect	-0.0167	0.0277			[-0.0760, 0.0355]
Cheapfake	Credibility	.0231	9.06 (1, 384)	.0028	Direct	-0.5010	0.1664	-3.010	.0028	[-0.8282, -0.1738]
	Mis-perception	.0476	9.58 (2, 383)	.0001	Direct	0.8091	0.1851	4.371	<.001	[0.4451, 1.1731]
					Indirect	-0.0251	0.0354			[-0.1044, 0.0380]
Deepfake	Credibility	.0917	40.26 (1, 399)	<.001	Direct	-1.0462	0.1649	-6.345	<.001	[-1.3703, -0.7220]
	Mis-perception	.0325	6.68 (2, 398)	.0014	Direct	0.6717	0.1931	3.479	.0006	[0.2921, 1.0513]
					Indirect	-0.1242	0.0709			[-0.2747, 0.0048]

Note: B = unstandardized regression coefficient; SE = standard error; t = t-statistic; p = p-value; CI = confidence interval (LL = lower limit, UL = upper limit). The direct effect is the effect of X on Y without considering M. The indirect effect represents the effect of X on Y through M (credibility).

Appendix B: Stimulus Transcripts

1. Facebook-Post caption:

Authentic

German: Ursula von der Leyen über den Stellenwert des Christentums für den Frieden in Europa.

English: Ursula von der Leyen on the importance of Christianity for peace in Europe.

Disinformation

German: Ursula von der Leyen gibt offen zu, dass nur christliche Geflüchtete aus der Ukraine willkommen sind!

English: Ursula von der Leyen openly admits that only Christian refugees from Ukraine are welcome!

2. Transcript:

Authentic/Decontextualized

German: Das Christentum ist tief eingewoben in Europa. In der Tat – Sie haben die christlichen Wurzeln beschrieben – und man spürt jeden Tag in unserem täglichen Handeln den Ursprung der christlichen Werte, wenn es um das große Thema des Friedens geht.

English: Christianity is deeply woven into Europe. Indeed - you have described the Christian roots - and you can feel the origin of Christian values every day in our daily actions when it comes to the great issue of peace.

Cheapfake

German: Das Christentum ist tief eingewoben in Europa. Und man spürt jeden Tag in unserem täglichen Handeln den Ursprung der christlichen Werte. Und insofern sind die christlichen Geflüchteten willkommen, wir tun alles, um sie zu schützen.

English: Christianity is deeply woven into Europe. And you can feel the origin of Christian values every day in our daily actions. And in this respect, Christian refugees are welcome and we do everything we can to protect them.

Deepfake

German: Das Christentum hat einen sehr hohen Stellenwert in Europa. Und dementsprechend sind Geflüchtete aus der Ukraine, die dieser Religion angehören, ganz besonders willkommen und wir tun alles, um sie zu schützen.

English: Christianity has a very high status in Europe. And accordingly, refugees from Ukraine who belong to this religion are particularly welcome and we do everything we can to protect them.

Appendix C: Ethical Considerations

This study obtained ethical approval from the university's Institutional Review Board (IRB). When designing the study, we reflected extensively on the potentially harmful nature of the stimuli, which involve disinformation about a real and active politician and could potentially evoke negative sentiments towards non-Christian refugees. The discrimination of refugees on the Ukrainian border is a sensitive topic that must be handled with care. We chose this topic to resemble the real-life phenomenon we set out to investigate. As such, our stimuli resemble real cases of decontextualized videos, cheap- or deepfakes citizens may come across in their everyday news consumption. For example, recently surfaced deceptive images include similarly harmful content, such as a cheapfake of Joe Biden saying the U.S. is "doomed" because of African Americans (Settles, 2021) or a photograph of Volodymyr Zelensky wearing a t-shirt with a Ukrainian military symbol, combined with the false claim it constitutes the Nazi's iron cross (AFP, 2022). In line with these occurrences, previous studies have created stimulus material casting politicians in a negative light (Dobber et al., 2021). For example, Hameleers et al. (2022) created a deepfake of a Dutch politician stressing the importance of Dutch traditions alongside the expression of anti-immigration attitudes ("immigrants and refugees are increasingly able to influence our culture and society with outdated ideas."). In this case, such a statement was necessary to test the moderating effect of issue-congruent attitudes. Similarly, our experiment aims to understand the moderating effect of anti-immigration attitudes on credibility perceptions. We consulted some of the authors of the cited studies to get first-hand advice on how to handle the use of real political actors and sensitive topics and followed their recommendations closely.

To ensure the stimulus is not disseminated outside of the experiment, several precautions were implemented. The survey was programmed to limit participants from re-watching the video multiple times using a timer. Options to right-click or download the video were disabled, and participants were reminded to complete the study until the end so they would not miss the debrief. Participants were required to spend a specific amount of time on the debriefing page to ensure they read the entire content. The comprehensive debriefing included showing participants the authentic video on which the stimulus was based, providing guidance on identifying manipulated visual media, and requiring participants to acknowledge

that they understood Ursula von der Leyen never made the statement attributed to her in the manipulated conditions.

References:

AFP. (2022, March 23). *Zelensky hit by false claim about "Nazi T-shirt" after US Congress meeting*. AFP Fact Check. https://factcheck.afp.com/doc.afp.com.326M9FR

Settles, G. (2021, April 21). *No, Biden didn't say this country is doomed because of African Americans*. Politifact. https://www.politifact.com/factchecks/2021/apr/23/instagram-posts/no-biden-didnt-say-country-doomed-because-african-/

Appendix D: Power Analysis

To estimate how many participants are needed to test the hypotheses for the main effect, a power analysis was conducted through simulation in R – an approach inspired by Johannes (2022). Expected group means and standard deviations are based on research by Hameleers et al. (2022) and Dobber et al. (2021), who found similar differences between groups. However, not enough prior research exists to make secure estimates, thus we decided not to work with planned contrasts, which would allow us to power more accurately. Overall, we expect the real video ($M = 4.5$, $SD = 1$) to be substantially more credible than the deepfake ($M = 3.5$, $SD = 1$). Moreover, Hameleers et al. (2022) contrasted a video constituting *malinformation* – closely resembling the *decontextualized* video in this study – with a deepfake and a real video. Based on these findings as well as our theoretical framework, we expect the decontextualized video to be more credible than the deepfake, but less than the real video ($M = 3.9$, $SD = 1$). Lastly, we expect the cheap fake to be less credible than the decontextualized video, but more so than the deepfake ($M = 3.7$, $SD = 1$). We simulated a linear regression including post-hoc comparisons with 200 runs to model the effect of the four categorical predictors on the expected mean scores. The p-value for the post-hoc contrast was set to 0.05. The power plot shows that $N = 600$ participants are needed for 80% power and $N = 800$ for 90% power.

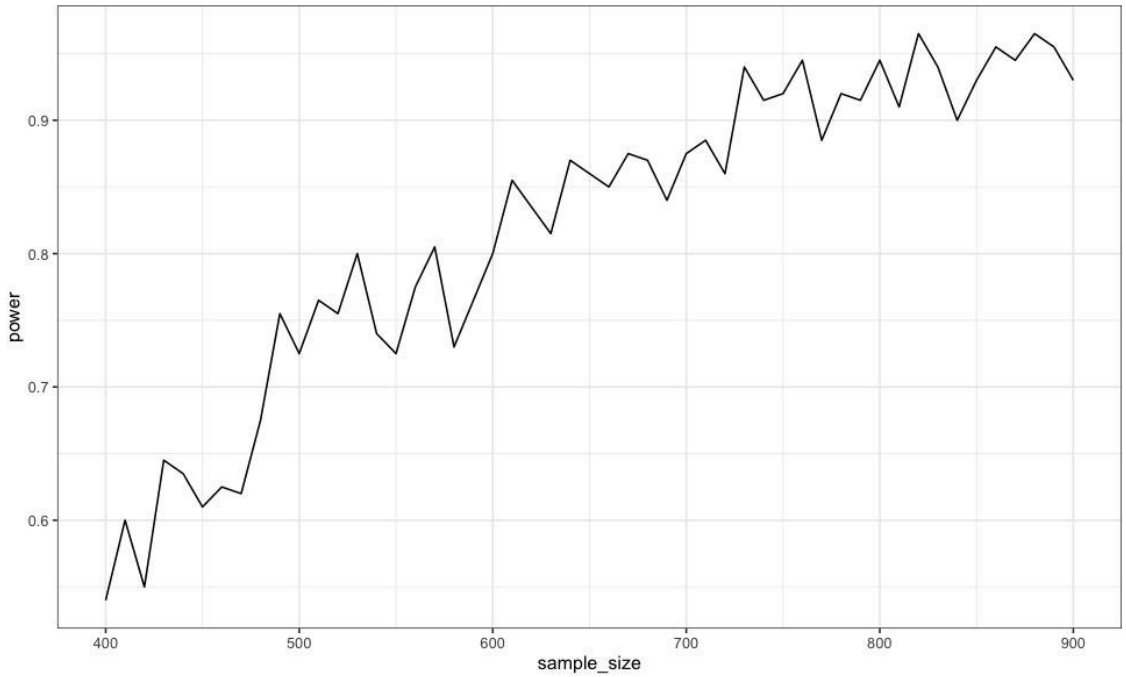


Figure 3. Plotted power for a sample size of 400-900 in steps of 100.

Appendix E: Manipulation Check

1. Manipulation checks

To investigate whether participants successfully recalled the central themes of the stimulus, we asked them about the extent to which they thought the post included the statements (1) "Ursula von der Leyen sees Christianity as a central aspect of Europe" and (2) "Ursula von der Leyen only welcomes Christian refugees from Ukraine" (1 = not at all; 7 = absolutely).

For the first statement, linear regression analysis with dummy coded variables (0 = "Authentic video"; 1 = "Fake conditions combined") to check whether there was a difference between the authentic video and the manipulated video conditions. The analysis revealed a statistically significant negative association between the stimulus type and the manipulation check score β = -0.4420, t (800) = -3.071, p = 0.0022, suggesting there are significant differences between conditions. A follow-up ANOVA with Bonferroni-correction shows that participants in the deepfake condition (M = 4.80, SD = 1.65) were less convinced that this statement was included in the post than in the authentic video (M = 5.59, SD = 1.71) and the decontextualized video (M = 5.46, SD = 1.70), but not less than in the cheapfake condition (M = 5.16, SD = 1.92), F (1, 800) = 23.38, p < 0.001. The authentic video, the decontextualized video, and the cheapfake do not differ significantly in terms of the perceived central message

of the post. The fact that each of them had the exact same wording at the beginning of the video ("Christianity is deeply woven into Europe") whereas the deepfake's message was slightly different ("Christianity has a very high status in Europe") could be an explanation for this. Importantly, the mean for the perceived presence of the *statement* is $M \geq 4.80$ in all conditions, suggesting that the statement was registered in all conditions.

Regarding the second statement, that von der Leyen only welcomes Christian refugees from Ukraine, linear regression analysis showed that participants who saw one of the manipulated videos [decontextualized ($M = 3.60$, $SD = 2.14$), cheapfake ($M = 4.03$, $SD = 2.09$) or deepfake ($M= 4.22$, $SD = 2.05$)] perceived the message that only Christians are welcome as significantly more prominent than those who saw the authentic video (M = 2.31, $SD = 1.65$), $\beta = 1.6287$, $t(800) = 9.945$, $p < 0.001$. Again, we conducted a follow-up ANOVA with Bonferroni-correction $F(1, 800) = 92.04$, $p < 0.001$. It shows that the deepfake differed significantly from the decontextualized video regarding the second manipulation check. The difference between cheapfake and decontextualized video or deepfake is not significant. Given that the statement was most strongly included in the deepfake condition (through audio, video and text), it appears plausible that the statement was also more strongly registered in this condition in comparison to the decontextualized condition. Importantly, the statement was perceived as more present in the decontextualized video in comparison to the authentic video ($M = 2.31$, $SD = 1.65$), showing that the manipulation worked through the accompanying caption alone.

2. Fabrication checks

Participants were asked at the end of the questionnaire to what extent they recognized the different elements of the post that were fabricated (1 = Does not apply at all; 7 = Fully applicable). First, they were asked whether they thought exclusively the text above the video was manipulated, which was only the case in the decontextualized video condition. However, participants found this statement to be significantly more applicable in the deepfake ($M = 4.07$, $SD = 1.85$) compared to the authentic video ($M = 3.34$, $SD = 1.68$), whereas the text was evaluated as equally manipulated in the authentic condition, the decontextualized condition ($M = 3.74$, $SD = 1.85$) and cheapfake condition ($M = 3.76$, $SD = 1.60$), $F(1, 800) = 16.03$, $p < 0.001$. We suspect that participants assumed that *spoken* text was meant instead of the *written* text in the Facebook post, as the spoken text was only manipulated in the deepfake.

Linear regression analysis with dummy-coded variables was performed to investigate whether exposure to the authentic and decontextualized video *combined* versus the cheapfake

and deepfake *combined* predicted whether participants thought the audio was manipulated. Indeed, this was the case – being exposed to either cheapfake or deepfake led participants to conclude that the audio must have been manipulated more than being exposed to an authentic or decontextualized video $\beta = 1.2658$, $t\ (800) = 10.39$, $p < 0.001$. In addition, being exposed to a cheapfake as opposed to any of the other conditions led participants to conclude the video was cut $\beta = 0.7901$, $t\ (800) = 5.24$, $p < 0.001$, and exposure to a deepfake was a significant predictor for whether people found the video to be digitally manipulated $\beta = 0.9446$, $t\ (800) = 6.70$, $p < 0.001$.

In general, these fabrication checks reveal that participants not only detected manipulated aspects in the videos they viewed but also identified the specific alterations, indicating the successful manipulation of various fabrication levels.

Appendix F: Measurements

1. Perceived credibility: (1 = Überhaupt nicht; 7 = Voll und ganz)

German: *Was halten Sie von dem Post? Finden Sie ihn...*

... faktisch richtig

... unglaubwürdig (reverse-coded)

... authentisch

English: *What do you think of the post? Do you find it...*

... factually correct

... implausible (reverse-coded)

... authentic

For perceived credibility, Cronbach's alpha level was questionable at .66. Accordingly, stepwise item deletion was performed, which revealed that the alpha would improve to .74 if the reverse-coded item "not believable" was removed from the analysis. The remaining two items show good inter-item correlation .59. The items are used for the main analysis ("factually correct", "authentic"), the reverse-coded item ("unbelievable") is included for robustness checks (appendix F).

2. Anti-immigration attitudes: (1 = Stimme überhaupt nicht zu; 7 = Stimme voll und ganz zu)

German: *Inwiefern stimmen Sie mit den folgenden Aussagen zur Europäischen Union (EU) überein?*

- *Durch Menschen, die aus anderen Ländern in die EU kommen, um hier zu leben, wird das kulturelle Leben untergraben.*
- *In die EU sollten viele Menschen aufgenommen werden, die aus ärmeren Ländern außerhalb Europas kommen.*
- *Menschen, die aus anderen Ländern in die EU kommen, um hier zu leben, stärken die Wirtschaft.*
- *Die EU sollte nur wenige Menschen aufnehmen, die nicht derselben ethnischen Gruppe oder Glaubensgemeinschaft angehören, wie die meisten Menschen in Europa.*

English: *To what extent do you agree with the following statements about the European Union (EU)?*

- *Cultural life is undermined by people coming from other countries to live in the EU.*
- *The EU should accept many people who come from poorer countries outside Europe.*
- *People who come from other countries to live in the EU strengthen the economy.*
- *The EU should only accept a few people who do not belong to the same ethnic group or religious community as most people in Europe.*

3. Perception of the politician:

German: *Was denken Sie über die Politikerin Ursula von der Leyen?*
Finde Sie diese...

unehrlich – ehrlich
unangenehm – freundlich
unfair – fair
korrupt – aufrichtig

English: *What do you think about the politician Ursula von der Leyen?*
Do you find her...

dishonest - honest
unpleasant - friendly
unfair - fair
corrupt - sincere

4. Misperceptions: (1 = Überhaupt nicht zutreffend; 7 = Voll und ganz zutreffend)

German: *Wie zutreffend finden Sie die folgenden Aussagen zur Europäischen Union (EU)?*

- *In der EU sind Kirche und Politik klar voneinander getrennt.*
- *Die meisten EU- Politiker*Innen lehnen die Aufnahme von muslimischen Geflüchteten ab.*
- *Ursula von der Leyen ist für die Bevorzugung von christlichen Geflüchteten in der EU.*
- *Geflüchtete aus christlichen Ländern genießen laut den EU- Richtlinien besondere Vorteile.*

English: *How accurate do you find the following statements about the European Union (EU)?*

- *In the EU, church and politics are clearly separated from each other.*
- *Most EU politicians reject the admission of Muslim refugees.*
- *Ursula von der Leyen is in favour of giving preference to Christian refugees in the EU.*
- *Refugees from Christian countries enjoy special advantages according to EU directives.*

Appendix G: Regression Tables

Table 7. Direct effect on perceived credibility per condition, including anti-immigration attitudes.

	Dependent variable: **Perceived credibility**			
	(1)	(2)	(3)	(4)
decontextualized	0.104 (0.133)			-0.413*** (0.160)
cheapfake		-0.013 (0.139)		-0.502*** (0.165)
deepfake			-0.742*** (0.133)	-1.039*** (0.162)
anti-imm. attitudes				-0.068* (0.040)
Constant	3.258*** (0.069)	3.289*** (0.067)	3.473*** (0.067)	4.069*** (0.207)
Observations	802	802	802	802
R^2	0.001	0.00001	0.038	0.054
Adjusted R^2	-0.0005	-0.001	0.036	0.049
Residual Std. Error	1.665 (df = 800)	1.666 (df = 800)	1.634 (df = 800)	1.623 (df = 797)
F Statistic	0.618 (df = 1; 800)	0.009 (df = 1; 800)	31.182*** (df = 1; 800)	11.301*** (df = 4; 797)

Note: OLS-based linear regressions with k-1 binary dummy variables. The authentic condition serves as the constant/intercept. * *p<0.1;* ** ***p<0.05;*** *** p<0.01

Table 8. Individual interactions of anti-immigration attitudes with perceived credibility per condition.

	Dependent variable: **Perceived credibility**		
	(1)	(2)	(3)
decon. x anti-imm. attitudes	-0.046 (0.091)		
cheapfake x anti-imm. attitudes		0.067 (0.100)	
deepfake x anti-imm. attitudes			-0.039 (0.092)
Constant	3.525*** (0.220)	3.676*** (0.211)	3.723*** (0.211)
Observations	802	802	802
R^2	0.005	0.005	0.041
Adjusted R^2	0.001	0.001	0.038
Residual Std. Error (df = 798)	1.663	1.664	1.633
F Statistic (df = 3; 798)	1.399	1.266	11.419***

Note: OLS-based linear regressions with k-1 binary dummy variables. The authentic condition serves as the constant/intercept. * *p<0.1;* ** ***p<0.05;*** *** p<0.01

Table 9. Complete model of all interactions with anti-immigration attitudes on perceived credibility.

	Dependent variable: **Perceived credibility**
decontextualized	-0.093 (0.498)
cheapfake	-0.555 (0.532)
deepfake	-0.762 (0.516)
decon. x anti-imm. attitudes	-0.074 (0.108)
cheapfake x anti-imm. attitudes	0.013 (0.117)
deepfake x anti-imm. attitudes	-0.063 (0.112)
Constant	3.923*** (0.358)
Observations	802
R^2	0.055
Adjusted R^2	0.046
Residual Std. Error	1.626 (df = 794)
F Statistic	6.568*** (df = 7; 794)

Note: OLS-based linear regressions with k-1 binary dummy variables. The authentic condition serves as the constant/intercept. * *p<0.1;* ** ***p<0.05;*** *** p<0.01

Appendix H: Detailed Analysis Report and Robustness Checks

The variables measured on 7-point Likert scales (anti-immigration attitudes, perceived credibility, misperceptions) and semantic-differential scales (perception of the politician) do not follow a normal distribution. Accordingly, non-parametric Kruskal-Wallis-Tests (and Dunn-Tests for post-hoc comparisons) were performed in addition to ANOVA to ensure the results are robust.

1. Perceived credibility:

The ANOVA revealed a significant main effect of condition on perceived credibility when the reverse-coded item "not believable" was included F (1, 800) = 49.29, $p < 0.001$, and when it was not included F (1, 800) = 39.6, $p < 0.001$. This holds for the Kruskal-Wallis-Test: H (3) = 50.04, $p < 0.001$ (three-item measure) and H (3) = 44.32, $p < 0.001$ (two-item measure).

To further examine the pairwise differences between the groups, post-hoc pairwise t-tests with a Bonferroni adjustment were conducted. The results of the pairwise comparisons for perceived credibility **including** the reverse-coded item "unbelievable" are as follows:

Group 1 vs. Group 2: ***p* = 0.23423**
Group 1 vs. Group 3: $p = 0.02386$
Group 1 vs. Group 4: $p < 0.00000000002$
Group 2 vs. Group 3: $p = 1.00000$
Group 2 vs. Group 4: $p = 0.00000220209$
Group 3 vs. Group 4: $p = 0.00032$

The Kruskal-Wallis/Dunn Test revealed similar results:

Group 1 vs. Group 2: ***p* = 0.1308**
Group 1 vs. Group 3: $p = 0.0187$
Group 1 vs. Group 4: $p = 0.0000$
Group 2 vs. Group 3: $p = 1.0000$
Group 2 vs. Group 4: $p = 0.0000$
Group 3 vs. Group 4: $p = 0.0001$

The results of the pairwise comparisons for the two-item measure perceived **credibility (new)** are as follows:

Group 1 vs. Group 2: ***p* = 0.05898**
Group 1 vs. Group 3: $p = 0.01532$
Group 1 vs. Group 4: $p < 0.0000000012$
Group 2 vs. Group 3: $p = 1.00000$
Group 2 vs. Group 4: $p = 0.00048$
Group 3 vs. Group 4: $p = 0.00594$

Similar significant differences were found in the Kruskal-Wallis-Test/Dunn Test when using the item credibility (new), with the exception that the difference between Group 1 vs Group 2 (Authentic vs. Decontextualized video) is significant ($p = 0.0454$):

Group 1 vs. Group 2: ***p* = 0.0454**
Group 1 vs. Group 3: $p = 0.0073$
Group 1 vs. Group 4: $p = 0.0000$
Group 2 vs. Group 3: $p = 1.0000$
Group 2 vs. Group 4: $p = 0.0001$
Group 3 vs. Group 4: $p = 0.0019$

We decided to report the ANOVA using the variable **credibility (new)** in our text, as this variable is most reliable. As the analysis of variance shows, participants who were exposed to a deepfake ($M = 2.37$, $SD = 1.69$) found this video to be significantly less credible compared to a cheapfake ($M = 3.28$, $SD = 1.66$), decontextualized video ($M = 3.36$, $SD = 1.55$) or authentic video ($M = 3.78$, $SD = 1.61$). Participants who saw the cheapfake perceived it as significantly less credible than the authentic video, but not less credible than a decontextualized video. The difference in credibility perception between an authentic video and a decontextualized video is approaching significance. **Overall, the findings are robust across different tests and with different items included in the perceived credibility measurement.**

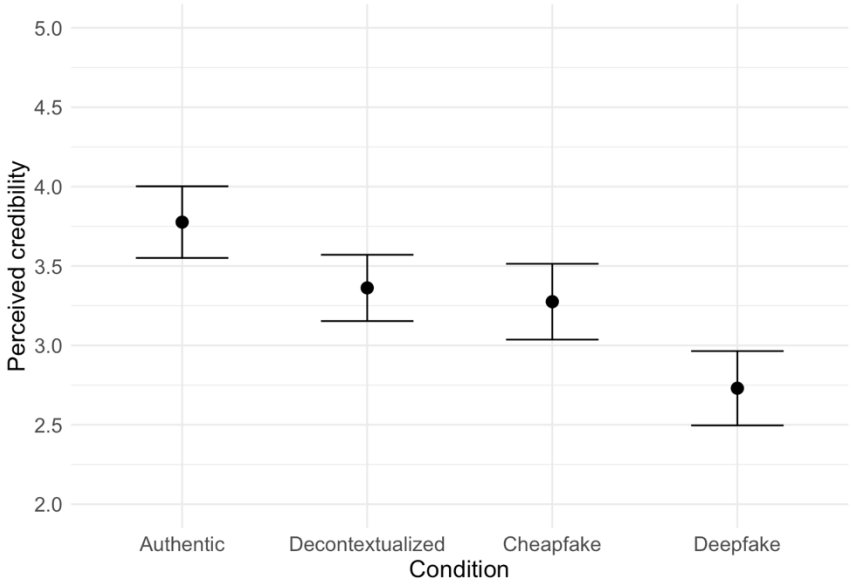


Figure 4. Interval plot for the variable perceived credibility.

2. Misperceptions:

The ANOVA revealed a significant main effect of condition on misperceptions regarding Ursula von der Leyen, F (1, 800) = 11.33, p = 0.000799. Also, the Kruskal-Wallis-Test was significant H (3) = 18.70, $p < 0.001$. The results of the post-hoc pairwise t-tests with Bonferroni corrections are as follows:

Group 1 vs. Group 2: p = 0.17737
Group 1 vs. Group 3: p = 0.00026
Group 1 vs. Group 4: p = 0.02101
Group 2 vs. Group 3: p = 0.25029
Group 2 vs. Group 4: p = 1.00000
Group 3 vs. Group 4: p = 1.00000

Results for Kruskal-Wallis/Dunn-Test:

Group 1 vs. Group 2: p = 0.0950
Group 1 vs. Group 3: p = 0.0001
Group 1 vs. Group 4: p = 0.0106
Group 2 vs. Group 3: p = 0.0916
Group 2 vs. Group 4: p = 1.0000
Group 3 vs. Group 4: p = 0.5343

This shows that participants who were exposed to a deepfake ($M = 4.01$, $SD = 2.00$) believed more strongly that Ursula von der Leyen is in favour of giving preference to Christian refugees in the EU compared to participants who saw an authentic video ($M = 3.47$, $SD = 1.68$). In addition, people who saw a cheapfake ($M = 4.25$, $SD = 1.91$) held a significantly stronger misperception than participants who saw an authentic video. There are no significant differences between those who saw a decontextualized video ($M = 3.87$, $SD = 1.88$) and any of the other groups.

Regarding the other misperception items ("In the EU, church and politics are clearly separated", "Most EU politicians reject the admission of Muslim refugees", "Refugees from Christian countries enjoy special advantages according to EU directives") no significant differences between groups were found.

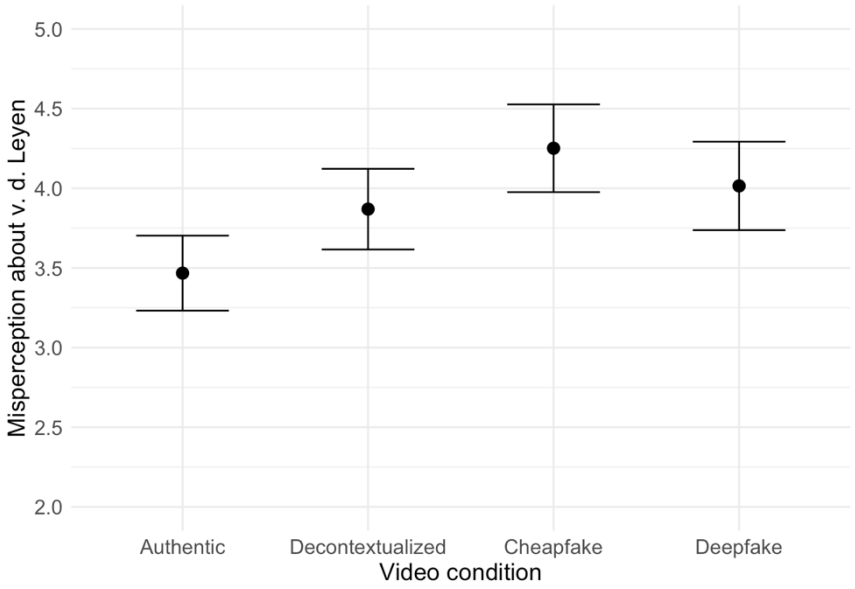


Figure 5. Interval plot for the variable misperception about Ursula von der Leyen.

3. Perception of von der Leyen:

The analysis revealed a significant main effect of condition on perceptions of von der Leyen F (1, 800) = 9.231, $p = 0.00246$ (same for Kruskal-Wallis-Test: H (3) = 8.40, $p < 0.038$). Post-hoc pairwise t-tests with Bonferroni corrections revealed that participants who saw a deepfake of von der Leyen ($M = 3.54$, $SD = 1.75$) had significantly more negative attitudes towards here than those who saw an authentic video ($M = 4.06$, $SD = 1.75$). There were no significant differences in regard to those who saw a decontextualized video ($M = 3.85$, $SD = 1.69$) or cheapfake $M = 3.75$, $SD = 1.76$) and any of the other conditions.

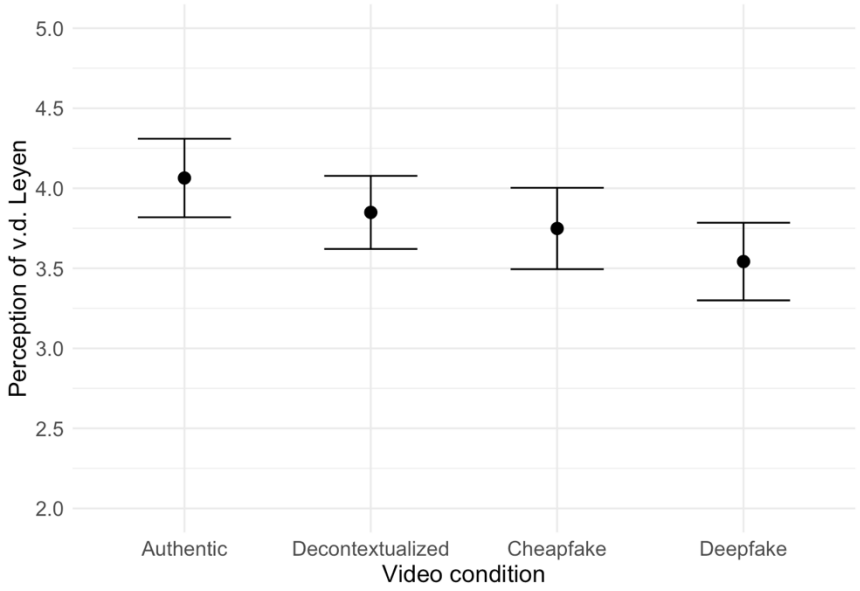


Figure 6. Interval plot for the variable perception of Ursula von der Leyen.

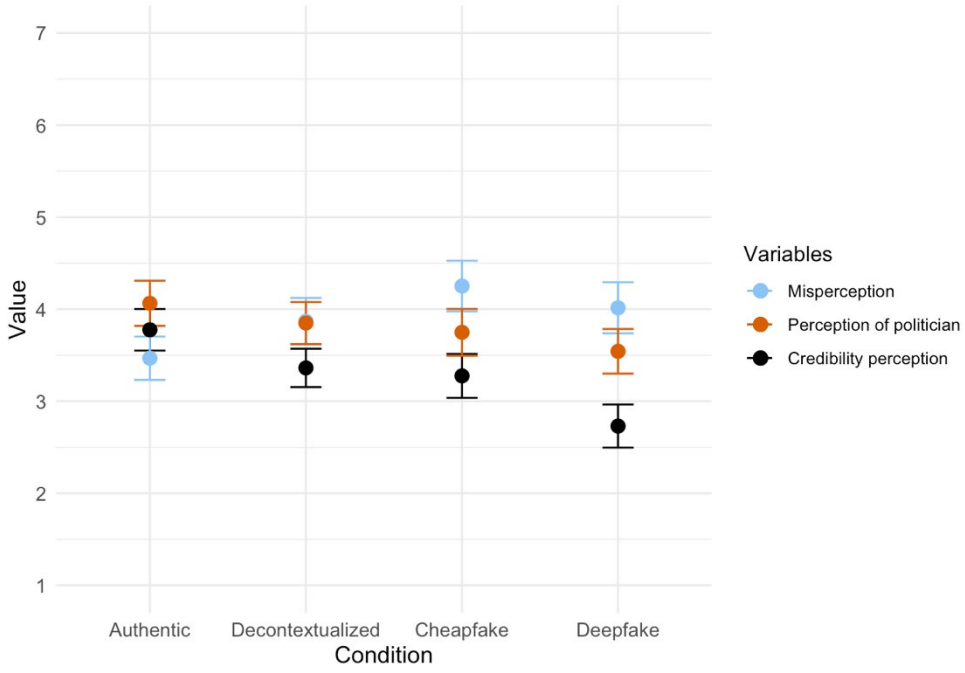


Figure 7. Combined interval plots ranging from 1-7 scale points for misperception about Ursula von der Leyen, perceptions of her integrity and perceived credibility.

Robustness check: Recognition of Ursula von der Leyen

83.54% (n = 670) of participants correctly identified Ursula von der Leyen as a politician, 16.46% (n = 132) did not.

1. Data analysis for participants who **did** recognize Ursula von der Leyen

We subset our sample to only those participants that correctly sorted von der Leyen into the category politician and ran our analyses again (n = 670). Randomization for this subsample was successful. The findings are largely the same for this subsample with one exception: the difference between the authentic condition (M = 3.89, SD = 1.58) and those who saw a decontextualized video (M = 3.43, SD = 1.55) is statistically significant for the dependent variable perceived credibility (post-hoc test result: p = 0.04779).

2. Data analysis for participants who **did not** recognize Ursula von der Leyen

Next, we subset the sample to participants who did not sort von der Leyen correctly (n = 132). Randomization for this subsample was successful. Notably, this sample is much smaller than the sample size determined in the power analysis. No significant main effects were found for perceived credibility, misperceptions regarding von der Leyen, attitudes towards the politician and general credibility ascribed to visuals.