Title: Reflection on data set and robustness check appendix for "On the same page? Experts are mostly, but not always aligned about disinformation in times of generative AI"
Authors: Teresa Weikmann (1), Ferre Wouters (2), Marina Tulin (1), Michael Hameleers (1), Claes de Vreese (1), Brahim Zarouali (2), Michaël Opgenhaffen (2)
Date: March 3rd, 2026
Note: The material contained herein is supplementary to the article named in the title and published in the Harvard Kennedy School (HKS) Misinformation Review.

Appendix A: Reflection on data set and robustness check

The dataset used in this study is based on a hard-to-reach expert sample and is characterized by uneven group sizes. Assembling the journalist subsample proved particularly challenging. Our aim was not to include just any journalists, but specifically those with dedicated experience in covering or investigating disinformation, a relatively small population, which resulted in only *n* = 16 in this expert group. To account for this limitation, we conducted Welch's ANOVA paired with Games-Howell post hoc tests to examine significant group differences. This approach performs particularly well under such conditions, as demonstrated by Delacre et al. (2019). Moreover, as Sauder and DeMars (2019) argue, the Games-Howell post hoc test performs particularly well with respect to controlling for Type I and Type II errors. A disadvantage of these tests, however, is that they are not specifically tailored to small sample sizes. In such cases, a nonparametric Kruskal-Wallis test, followed by a Dunn's post hoc test, is preferred (see Chan & Walmsley, 1997). Therefore, we conducted Kruskal-Wallis tests in addition to the Welch's ANOVA to ensure the robustness of our findings. This was confirmed, as shown in the table below.

Table 1. *Group comparisons with post-hoc significance tests (Games-Howell, based on Welch's ANOVA and Dunn's test, based on Kruskal-Wallis Test) for perceived competence, risks, and responsibility in addressing AI-generated disinformation.*

Dependent Variable	***Comparison***	***Mean 1***	***Mean 2***	***Mean Diff***	***95% CI***	***Games-Howell p***	***Dunn's p***
Perceived competence							
	Academics - Fact checkers	5.27	6.05	0.78	[0.34, 1.21]	.0002*	.0002*
	Academics - Journalists	5.27	5.10	-0.17	[-1.04, 0.71]	.883	1.000
	Fact checkers - Journalists	6.05	5.10	-0.94	[-1.82, -0.06]	.034*	.0044*
Perceived risks: ***Moral panic***							
	Academics - Fact checkers	4.47	4.07	-0.40	[-1.28, 0.49]	0.528	0.378
	Academics - Journalists	4.47	3.94	-0.53	[-1.45, 0.39]	0.347	0.224
	Fact checkers - Journalists	4.07	3.94	-0.13	[-1.13, 0.87]	0.945	0.957
Long-standing issue							
	Academics - Fact checkers	4.68	4.31	-0.37	[-1.20, 0.46]	0.534	0.480
	Academics - Journalists	4.68	4.12	-0.56	[-1.75, 0.64]	0.488	0.290
	Fact checkers - Journalists	4.31	4.12	-0.19	[-1.44, 1.06]	0.928	0.973

Dependent Variable	*Comparison*	*Mean 1*	*Mean 2*	*Mean Diff*	*95% CI*	*Games-Howell p*	*Dunn's p*
Seeing is no longer believing							
	Academics - Fact checkers	4.85	4.62	-0.23	[-1.14, 0.68]	0.814	0.884
	Academics - Journalists	4.85	5.81	0.96	[0.16, 1.76]	0.015*	0.048*
	Fact checkers - Journalists	4.62	5.81	1.19	[0.24, 2.15]	0.011*	0.025*
Undermine democratic debate							
	Academics - Fact checkers	4.79	5.31	0.52	[-0.19, 1.23]	0.185	0.048(*)
	Academics - Journalists	4.79	4.88	0.09	[-0.77, 0.94]	0.965	0.275
	Fact checkers - Journalists	5.31	4.88	-0.44	[-1.39, 0.51]	0.506	0.275
Confusion real/fake							
	Academics - Fact checkers	5.87	6.38	0.51	[-0.08, 1.09]	0.102	0.029(*)
	Academics - Journalists	5.87	5.94	0.07	[-0.73, 0.86]	0.977	1.000
	Fact checkers - Journalists	6.38	5.94	-0.44	[-1.28, 0.39]	0.403	0.210
Disrupt political decision making							
	Academics - Fact checkers	4.47	4.83	0.36	[-0.41, 1.12]	0.498	0.431
	Academics - Journalists	4.47	4.81	0.34	[-0.44, 1.13]	0.536	1.000
	Fact checkers - Journalists	4.83	4.81	-0.02	[-0.91, 0.88]	0.999	1.000
Attribution of responsibility ***VLOPs***							
	Academics - Fact checkers	6.23	6.69	0.46	[0.05, 0.87]	0.026*	0.020*
	Academics - Journalists	6.23	6.56	0.33	[-0.39, 1.05]	0.501	0.052
	Fact checkers - Journalists	6.69	6.56	-0.13	[-0.84, 0.58]	0.894	1.000
Journalists							
	Academics - Fact checkers	5.57	5.66	0.08	[-0.53, 0.69]	0.945	0.846
	Academics - Journalists	5.57	6.12	0.55	[-0.13, 1.23]	0.127	0.044(*)
	Fact checkers - Journalists	5.66	6.12	0.47	[-0.32, 1.26]	0.326	0.169

Dependent Variable	*Comparison*	*Mean 1*	*Mean 2*	*Mean Diff*	*95% CI*	*Games-Howell p*	*Dunn's p*
Fact checkers							
	Academics - Fact checkers	5.38	6.07	0.69	[0.07, 1.30]	0.025*	0.019*
	Academics - Journalists	5.38	6.19	0.81	[-0.01, 1.62]	0.053	0.017*
	Fact checkers - Journalists	6.07	6.19	0.12	[-0.71, 0.94]	0.933	0.969
Government/political actors							
	Academics - Fact checkers	5.94	5.79	-0.14	[-0.89, 0.60]	0.888	1.000
	Academics - Journalists	5.94	6.06	0.13	[-0.77, 1.02]	0.934	1.000
	Fact checkers - Journalists	5.79	6.06	0.27	[-0.75, 1.29]	0.796	1.000
Experts/academics							
	Academics - Fact checkers	5.43	5.24	-0.18	[-0.80, 0.43]	0.753	0.881
	Academics - Journalists	5.43	5.75	0.32	[-0.54, 1.19]	0.618	0.283
	Fact checkers - Journalists	5.24	5.75	0.51	[-0.41, 1.43]	0.373	0.154
News users							
	Academics - Fact checkers	4.74	3.62	-1.12	[-1.97, -0.28]	0.006*	0.006*
	Academics - Journalists	4.74	4.06	-0.68	[-1.88, 0.52]	0.345	0.235
	Fact checkers - Journalists	3.62	4.06	0.44	[-0.84, 1.73]	0.675	0.577

Note: Significant results marked with an asterisk in parentheses – (*) – are based on non-significant omnibus tests and should be interpreted with caution, as they are not considered meaningful. Hence, they are not mentioned in the paper.