Title: Supplementary statistical tables for "Experts disagree on how to fight AI disinformation, but agree that health and politics need different solutions"
Authors: Alexander Loth (1), Martin Kappes (1), Marc-Oliver Pahl (2)
Date: July 29th, 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 C: Supplementary statistical tables

***Table C1.*** *Sample characteristics of the expert respondents (N = 54).*

| Characteristic | Value |
|---|---|
| Region — European Union | 23 (42.6%) |
| Region — North America | 19 (35.2%) |
| Region — Non-EU Europe (incl. UK) | 7 (13.0%) |
| Region — Asia-Pacific / Global | 5 (9.3%) |
| Years of professional experience — median | 17 |
| Years — range | 1–65 |
| Self-rated LLM understanding (1–7) — mean | 4.89 |
| Self-rated deepfake familiarity (1–7) — mean | 4.54 |
| Attribution preference — anonymous | 27 (50.0%) |
| Attribution preference — acknowledgment only | 14 (25.9%) |
| Attribution preference — quote-attribution opt-in | 13 (24.1%) |

*Note: Descriptive statistics, distributional diagnostics, recruitment details, and the domain-level policy priority matrix.*

***Table C2.*** *Distribution shape of mitigation effectiveness ratings across strategies (N = 54).*

| | Rating | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Strategy** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **Mean** | ***SD*** | **High %** | **Mid %** | **Low %** |
| Government regulation | 2 | 2 | 5 | 5 | 13 | 12 | 15 | 5.24 | 1.65 | 74.1 | 9.3 | 16.7 |
| Digital watermarking / provenance | 1 | 3 | 2 | 11 | 11 | 13 | 13 | 5.20 | 1.53 | 68.5 | 20.4 | 11.1 |
| Media / public literacy | 0 | 5 | 6 | 9 | 9 | 10 | 15 | 5.07 | 1.67 | 63.0 | 16.7 | 20.4 |
| Platform enforcement | 0 | 4 | 4 | 12 | 15 | 11 | 8 | 4.91 | 1.42 | 63.0 | 22.2 | 14.8 |
| Technical detection | 1 | 3 | 8 | 13 | 13 | 13 | 3 | 4.57 | 1.40 | 53.7 | 24.1 | 22.2 |

*Note: Likert votes (count per rating) per strategy, with summary share of high (≥ 5), mid (= 4), and low (≤ 3) ratings. All five distributions are right-skewed and unimodal. None exhibits the bimodal pattern that would indicate genuine effectiveness polarization. Disagreement in the forced-choice rankings therefore reflects priority-setting rather than effectiveness disagreement.*

***Table C3.*** *Recruitment funnel from candidate identification to valid responses.*

| **Stage** | ***n*** |
|---|---|
| Candidates identified | 516 |
| Directly contacted (any channel) | 454 |
| Channels — Email | 214 |
| Channels — LinkedIn | 95 |
| Channels — Mastodon DM | 37 |
| Channels — Bluesky / X DM | 23 |
| Channels — Mixed / other | 85 |
| Survey responses received | 54 |
| Response rate (responses / contacted) | 11.9% |

*Note: Snowball forwards from contacted candidates to colleagues are not captured in the denominator and are unknown.*

***Table C4.*** *Policy priority matrix showing perceived threat by domain and modality (N = 54).*

| **Domain** | **Primary modality threat (*M*)** | **Priority lever** | **Supporting strategies** |
|---|---|---|---|
| Political | Video deepfake (6.31) | Audiovisual provenance (e.g., C2PA) + platform enforcement | Statutory regulation of election-period synthetic media; rapid-response fact-checking; literacy targeted at electoral periods |
| Health | AI-generated text (5.80) | Labelling of AI-generated health content + medical fact-checking infrastructure | Clinician-facing literacy on LLM hallucinations; platform demotion of unverified medical claims; regulator coordination (e.g., FDA/EMA guidance) |
| Financial | Audio deepfake (5.65) | Voice-authentication standards for high-value transactions + KYC modernization | Sector-specific regulation of synthetic-voice fraud; cross-bank threat-intelligence sharing; staff training |
| Social (cross-cutting) | Mixed (video 6.00 / images 5.76) | Accessibility-first literacy reaching non-technical audiences | Provenance signals embedded in mainstream platforms; transparent platform moderation; community-driven counter-speech |

*Note: A preliminary, operational mapping from the descriptive threat patterns observed in this study to a prioritized mitigation stack per domain. The matrix is intended as a starting point for policymakers and platform operators rather than a closed framework; cells should be re-weighted as the AI capability landscape evolves. Cell content draws on respondent ratings (Findings) and on existing instruments (Bommasani et al., 2024; Chesney & Citron, 2019; Corsi et al., 2024; De Angelis et al., 2023; European Parliament, 2024). The matrix operationalizes the layered, contested-not-polarized mitigation landscape identified in the Implications section.*