

The Manufactured Reality and the Emerging Mind

The Impact of AI-Generated Visual Content on the Cognitive
and Social Development of Generations Z and Alpha
— Toward an Ethical UI/UX Design Framework

Multi-Source Comparative Analysis · Research Synthesis Paper

Clark & Lewandowsky (2026) · Mastercard Institute (2026) · Batool & Bashir (2026) · European Parliament (2025)

SOURCES
4 Peer-Reviewed

METHOD
Comparative Matrix

FOCUS
Gen Z & Gen Alpha

OUTPUT
Ethical UI/UX Models

Abstract

The proliferation of AI-generated visual content — particularly deepfake videos — presents an unprecedented challenge to the cognitive and social development of Generations Z (born 1997–2012) and Alpha (born 2013–present). This paper synthesises four peer-reviewed and institutional sources to construct an original **Four-Dimensional Comparative Matrix**, spanning: *Digital Awareness*, *Behavioural Engagement*, *Cognitive Dependency*, and *Critical Resilience*. Drawing on experimental psychology (Clark & Lewandowsky, 2026), socio-cultural trend analysis (Mastercard Institute, 2026), digital literacy research (Batool & Bashir, 2026), and legislative policy review (European Parliament / EPRS, 2025), the paper maps differential vulnerability profiles across both generations. Converging evidence demonstrates that cognitive awareness of fabricated content is insufficient to neutralise its influence — a structural phenomenon termed herein the **Awareness–Action Gap**. In response, this paper proposes six **Ethical UI/UX Design Models** that operationalise transparency, informed consent, and age-sensitive disclosure architecturally within digital interfaces. Findings carry direct implications for platform designers, educators, policymakers, and AI regulators.

Keywords: Deepfake · Generative AI · Generation Z · Generation Alpha · Cognitive Dependency · Critical Resilience · Ethical UI/UX · Digital Literacy · Comparative Matrix · Transparency Warnings

Table of Contents

1	Introduction & Theoretical Framework	3
2	Literature Review	3–5
2.1	Clark & Lewandowsky (2026) — Deepfake Influence Despite Warnings	3
2.2	Mastercard Institute (2026) — Gen Z & Alpha: From Scrolling to Shaping	4
2.3	Batool & Bashir (2026) — AI Media Perception & Digital Literacy in Gen Z	4
2.4	European Parliament/EPRS (2025) — Children and Deepfakes	5
2.5	Cross-Source Synthesis	5
3	Methodology: The Four-Dimensional Comparative Matrix	5
4	Analysis & Findings	6–7
4.1	Generation Z — Vulnerability Profile	6
4.2	Generation Alpha — Vulnerability Profile	6
4.3	Cross-Generational Comparative Analysis	7
5	Ethical UI/UX Design Framework: Six Proposed Models	7–8
6	Discussion, Policy Implications & Recommendations	9
7	Conclusion	10
	References	11–12

1. Introduction & Theoretical Framework

The contemporary digital landscape is undergoing a fundamental transformation: content production has migrated from exclusively human domains to algorithmic generative systems. AI-generated video — once a technical curiosity — has emerged as an active cultural force reshaping public perception and redefining the parameters of social interaction. This shift imposes existential cognitive challenges upon Generations Z and Alpha, who inhabit a reality in which the authentic and the synthetic are increasingly indistinguishable.

1.1 Research Gap

Prior scholarship has concentrated predominantly on technical detection capabilities or isolated ethical dimensions of AI-generated content. A critical gap persists at the intersection of behavioural shifts — from passive consumption to active cultural co-creation — and the cognitive and social developmental outcomes arising from sustained exposure to generative deepfake content. The absence of a critical resilience framework commensurate with the pace of algorithmic video evolution renders current protective measures structurally inadequate.

1.2 Research Questions

RQ1	How do generative AI technologies affect perceptual and discriminatory cognitive capacities in Generation Z and Generation Alpha?
RQ2	What is the interactive relationship between digital awareness and cognitive dependency under persistent deepfake influence, despite explicit transparency warnings?
RQ3	How can Ethical UI/UX interfaces be designed to strengthen critical resilience among emerging digital users?

1.3 Key Conceptual Definitions

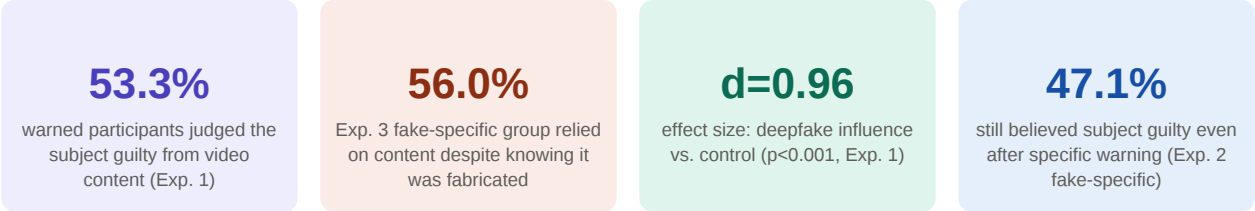
Term	Operational Definition
Generative AI Videos	Visual content created or modified by AI algorithms, including deepfake technology, without necessarily relying on pre-existing footage.
Generation Z	Individuals born 1997–2012; digital natives who came of age during the full transition to mobile-first internet culture.
Generation Alpha	Individuals born 2013 onward; the first 'AI-native' generation, raised in an environment where AI is a default condition of existence.
Critical Resilience	Cognitive capacity to evaluate the credibility of digital content and resist the manipulative influence of AI-generated material.
Awareness–Action Gap	The empirically observed divergence between knowing content is fabricated and being cognitively unaffected by its influence.

2. Literature Review

2.1 Clark & Lewandowsky (2026) — The Continued Influence of Deepfake Videos

SOURCE: Clark, S. & Lewandowsky, S. (2026). *Communications Psychology*, 4, 13.

Published in *Communications Psychology* (Nature Portfolio), this three-experiment pre-registered study (N = 175, 275, 223) is the most rigorous empirical examination of post-identification deepfake influence to date. The study produced original deepfake videos — eliminating prior-exposure confounds — and assessed perception of guilt and fakeness using validated 7-point Likert scales. Three warning conditions were tested: no warning, generic warning, and specific warning.



CORE FINDING: Participants who explicitly stated they believed the warning and knew the video was fake still relied on its content when forming judgements. This is the empirical foundation of the Awareness-Action Gap — the central contribution of this synthesis.

Specific warnings reduced influence but did not eliminate it; generic warnings introduced 'generalised uncertainty', eroding trust in authentic video material. The study also identifies the 'liar's dividend': the strategic potential for malicious actors to label genuine footage as deepfakes to evade accountability. The authors conclude that transparency-based regulation alone is structurally insufficient and recommend content removal as a complementary tool.

2.2 Mastercard Institute (2026) — Gen Z & Alpha: From Scrolling to Shaping

SOURCE: Mastercard Institute (2026). *The Next Growth Generation*. 18 markets, N=5,300.

This large-scale mixed-methods report surveyed approximately 3,500 Gen Z respondents (ages 17–29) and ~1,800 parents of Gen Alpha children (ages 11–16) across 18 global markets. Its central thesis identifies a fundamental paradigm shift: both generations are transitioning from passive content consumers to active cultural co-creators. Gen Alpha is defined as the first 'AI-native' generation — born into an environment where AI is not a novelty but a default condition.

65%

of Gen Z identify as content creators building personal brands

59%

of Gen Z comfortable using AI tools for career advice

40%

of 12–15 year-olds practice intentional digital self-care breaks

\$33T

projected Gen Z earning potential by 2030

PARADIGM SHIFT: Generation Alpha is the first 'AI-native' generation — the first cohort to grow up in an environment where AI is a default condition of existence, fundamentally restructuring their relationship with knowledge, identity, and creativity. Reality is no longer split between digital and physical — the two seamlessly interweave.

The report documents the fragmentation of mass culture into hyper-niche sub-cultures algorithmically curated for individuals. The majority of Gen Z respondents affirm that a shared cultural 'mainstream' no longer exists. AI chatbots function as co-pilots for major life decisions, and digital and physical realities interweave seamlessly in daily experience. Today's most successful young creators turn micro-communities into monetisable enterprises, exemplifying co-authored storytelling between brands and audiences.

2.3 Batool & Bashir (2026) — AI Media Perception & Digital Literacy in Gen Z

SOURCE: Batool, M. & Bashir, S. (2026). *Academia Journal of Psychology*, 5(3). Pakistan (N=300).

This quantitative correlational study ($N = 300$; $M_{\text{age}} = 21.95$, $SD = 2.41$) surveyed Gen Z university students in Pakistan, using standardised scales and Hayes' PROCESS macro for mediation analysis. It tested whether digital literacy mediates the relationship between AI-generated media perception and attitudes toward AI.

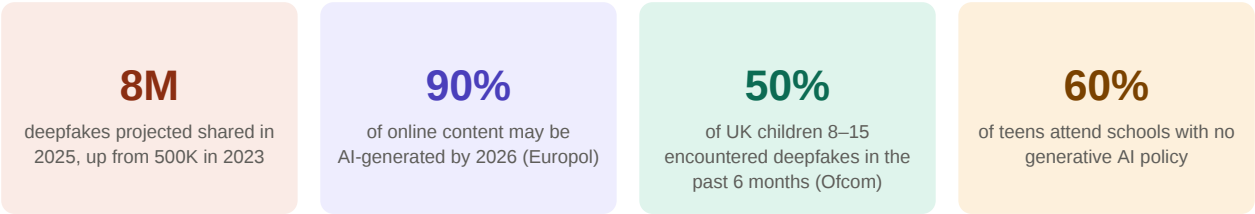
COUNTERINTUITIVE FINDING: Digital literacy was positively associated with AI attitudes, yet AI-generated media perception did NOT significantly predict digital literacy — indicating a direct rather than mediated pathway. This directly parallels Clark & Lewandowsky's gap between knowing content is fake and being cognitively unaffected by it.

This finding challenges conventional AI-literacy curricula premised on the assumption that exposure to AI-generated content naturally develops critical discernment. It implies that intentional, structured digital literacy education is necessary — and that the Awareness–Action Gap cannot be closed by content exposure alone.

2.4 European Parliament / EPRS (2025) — Children and Deepfakes

SOURCE: Negreiro, M. (2025). EPRS Briefing PE 775.855. European Parliament, July 2025.

This institutional briefing synthesises data from Europol, Ofcom, UNICEF, the Internet Watch Foundation, and national data protection authorities to map the deepfake crisis for children. It covers technological evolution (GANs, autoencoders, diffusion models), legislative frameworks (AI Act, DSA, Directive 2024/1385), and educational imperatives.



CRITICAL ALERT — THE CONFIDENCE PARADOX: Only 9% of adults feel confident identifying deepfakes, versus 20% of children aged 8–15. Yet both groups perform near chance level (~50%) in detection tasks. Children's overconfidence in an environment of genuine incapacity constitutes a compounded cognitive vulnerability. The EPRS calls for a triad: Technology + Legislation + Education.

The report further reveals that only 37% of parents whose children use AI tools are aware of this, and nearly 25% incorrectly believe their children do not use such tools. AI Act Article 50 imposes labelling obligations but provides no definitive guidelines on implementation and includes exceptions. EU-level monitoring indicators for children's online use are currently non-existent.

2.5 Cross-Source Synthesis: Convergences and Divergences

Dimension	Convergence	Tension / Complement
Awareness vs. Effect	All four sources confirm awareness alone is insufficient	Clark: awareness ≠ immunity; Batool: literacy does not mediate perception
Age Vulnerability	All sources identify younger cohorts as higher-risk	EPRS: child data; Mastercard: active agency despite structural risk
Education	Universal call for AI literacy education	EPRS: 60% lack school policies; Batool: effect is direct, not mediated
Legislation	Existing frameworks are inadequate	Clark: transparency insufficient; EPRS: AI Act labelling has exceptions
Cultural Agency	Gen Z/Alpha as active cultural shapers	Mastercard: creators; Clark: creators equally susceptible to manipulation

SYNTHESIS CONCLUSION: The four sources converge on a single structural insight — the Awareness-Action Gap is not a peripheral anomaly but the defining mechanism through which AI-generated content achieves influence. It persists across warned, aware, and digitally literate populations, demanding systemic rather than individual-level responses.

3. Methodology: The Four-Dimensional Comparative Matrix

This paper adopts a Comparative Matrix methodology — an original analytical framework designed to map source findings across four interdependent cognitive-behavioural dimensions. Unlike single-axis analyses, the matrix enables simultaneous tracking of overlapping and competing influences across generational cohorts and evidentiary sources. Each dimension is operationally defined, sourced, and assigned a risk classification.

Dimension	Operational Definition	Primary Source(s)	Risk Level
Digital Awareness	Individual capacity to distinguish authentic from AI-generated content at point of consumption	Clark & Lewandowsky; EPRS (2025)	HIGH
Behavioural Engagement	Patterns of content consumption, sharing, remix, and active co-creation across platforms	Mastercard Institute (2026)	MEDIUM-HIGH
Cognitive Dependency	Degree of reliance on AI-generated content in shaping beliefs, identity, and epistemic reference	Mastercard; Batool & Bashir	HIGH
Critical Resilience	Cognitive proficiency in evaluating content credibility and resisting manipulative AI material	All four sources	CRITICAL

Table 1. The Four-Dimensional Comparative Matrix — analytical framework for this study.

4. Analysis & Findings

4.1 Generation Z — Vulnerability Profile

Generation Z occupies a paradoxical position: highly digitally literate, yet empirically susceptible to AI-generated content influence. Mastercard data reveals that 65% identify as content creators and 59% utilise AI for career counsel — reflecting high Behavioural Engagement and Cognitive Dependency. Yet Clark & Lewandowsky's experimental data demonstrate that Digital Awareness does not translate to Critical Resilience: even warned participants who knew content was fake continued to be influenced by deepfake material.

Batool & Bashir's findings further nuance this: digital literacy positively predicts AI attitudes but does not mediate the perceptual impact of AI media — confirming that for Gen Z, the Awareness–Action Gap is a structural, not individual, phenomenon. It cannot be solved by simply increasing digital literacy scores.

Matrix Dimension	Gen Z Profile	Evidence	Direction
Digital Awareness	Moderate–High self-reported; lower actual performance in detection tasks	Clark & Lewandowsky (2026)	↑ Conf., ↓ Acc.
Behavioural Engagement	Active creator-consumer hybrid; 65% creator-identified	Mastercard (2026)	↑↑ High
Cognitive Dependency	59% use AI for major life decisions; AI as co-pilot	Mastercard (2026)	↑ High
Critical Resilience	Insufficient — awareness does not confer cognitive immunity	Clark; Batool & Bashir	↓ Low

Table 2. Generation Z profile across the Four-Dimensional Comparative Matrix.

4.2 Generation Alpha — Vulnerability Profile

Generation Alpha presents a qualitatively distinct and more acute risk profile. As the first AI-native generation (Mastercard, 2026), they are immersed in generative AI environments before developing stable epistemological frameworks. The EPRS (2025) reports that 50% of children aged 8–15 have encountered deepfakes, while the confidence paradox — where children report 20% detection confidence against near-chance actual performance — signals a structural vulnerability that awareness campaigns cannot address.

Their cognitive abilities are, by developmental definition, still forming — making them disproportionately vulnerable to the continued-influence effects documented by Clark & Lewandowsky. Moreover, 60% attend schools with no generative AI policy, and only 37% of their parents are aware they use AI tools — creating an environment of triple exposure: technological, institutional, and parental.

Matrix Dimension	Gen Alpha Profile	Evidence	Direction
Digital Awareness	Overconfident (20% self-rated) vs. ~50% actual detection chance	EPRS (2025)	↑↑ False conf.
Behavioural Engagement	Higher intensity than adults; AI-native from birth	Mastercard; EPRS	↑↑ Highest
Cognitive Dependency	Structural — AI is default environment, not optional tool	Mastercard (2026)	↑↑ Critical
Critical Resilience	Severely underdeveloped — cognitive faculties still forming	EPRS; Clark & Lewandowsky	↓↓ Critical

Table 3. Generation Alpha profile across the Four-Dimensional Comparative Matrix.

4.3 Cross-Generational Comparative Analysis

THE AWARENESS–ACTION GAP (Confirmed): Across both generations, the most consistent and consequential finding is the empirical divergence between knowing content is fabricated and being cognitively unaffected by it. The gap is wider for Generation Alpha due to developmental factors, and persists for Generation Z despite higher baseline digital literacy. Neither awareness, nor warnings, nor literacy alone close this gap.

While Gen Z demonstrates adaptive agency — reorienting from pure consumption to co-creation (Mastercard, 2026) — this agency does not insulate them from perceptual manipulation. Gen Alpha faces structural vulnerability: their AI-native status means deepfake content is not an intrusion into their media environment but a constitutive part of it. Both generations require systemic rather than individual-level interventions.

	Generation Z	Generation Alpha
Digital Awareness	Moderate–High (inflated)	Overconfident (20% vs. 50% actual)
Behavioural Engagement	High — active co-creation	Very High — AI-native default
Cognitive Dependency	High — AI as co-pilot	Critical — AI as environment
Critical Resilience	Low — Awareness≠Immunity	Critical — still developing
Key Risk	Awareness–Action Gap	Confidence Paradox + Cognitive Immaturity

Table 4. Cross-generational comparative analysis across the Four-Dimensional Matrix.

5. Ethical UI/UX Design Framework: Six Proposed Models

The consistent failure of awareness-based interventions across the evidence base establishes a compelling case for design-level solutions. If cognitive warnings are insufficient, the interface itself must encode transparency, friction, and informed consent architecturally. Drawing on the principle that *design is an ethical act*, the following six models operationalise the Four-Dimensional Matrix as interface obligations, not optional features.

M-1	AI Content Labelling System	A dynamic, contextual labelling architecture beyond static watermarks. Labels adapt by content risk level, audience age, and platform context. Grounded in AI Act Article 50 and Clark & Lewandowsky's finding that specific warnings outperform generic ones. Critically, labels must be visually prominent, not buried in metadata.
M-2	Contextual Disclosure UI	An interface layer providing contextual AI-generation information before and during content viewing — not only as post-hoc metadata. Addresses the temporal limitation of after-the-fact fact-checking identified in Clark & Lewandowsky's experimental design. Disclosure must precede the 'seeing is believing' moment.
M-3	Progressive Trust Indicator	A visual trust-gradient system showing degrees of content authenticity rather than binary real/fake labels. Responds to the nuanced perception-of-fakeness distributions (7-point Likert) in Clark & Lewandowsky (2026). Gradients reduce the generalized uncertainty produced by blunt generic warnings.
M-4	Embedded Verification Widget	An inline tool enabling users to query content provenance, generation metadata, and source credibility without leaving the viewing interface. Reduces cognitive friction of external fact-checking — directly addressing the behavioural engagement pattern of continuous, frictionless scrolling identified in Mastercard (2026).
M-5	Children's Digital Awareness Dashboard	An age-appropriate interface layer for platforms accessed by Gen Alpha, providing simplified AI-literacy indicators, parental visibility features, and 'safe-by-design' content categorisation. Responds directly to EPRS (2025) recommendations on child protection and the non-existence of EU-level monitoring indicators.
M-6	Informed Consent Flow	A structured consent architecture requiring active user acknowledgement prior to AI-generated content consumption. Distinguishes itself from passive labels by demanding cognitive engagement — a direct design response to the 'seeing-is-believing' heuristic and the Awareness–Action Gap confirmed across all four source studies.

DESIGN IMPERATIVE: These six models are not cosmetic improvements to existing interfaces. They represent a fundamental reconceptualisation of the designer's ethical responsibility in the age of generative AI. The interface is not neutral — it is the primary site at which the Awareness–Action Gap can be structurally narrowed.

5.1 Model-to-Matrix Alignment

UI/UX Model	Primary Matrix Dimension Addressed	Source Foundation
M-1: AI Content Labelling System	Digital Awareness	Clark & Lewandowsky; AI Act Art. 50
M-2: Contextual Disclosure UI	Digital Awareness + Critical Resilience	Clark & Lewandowsky (2026)
M-3: Progressive Trust Indicator	Critical Resilience	Clark & Lewandowsky (2026)
M-4: Embedded Verification Widget	Behavioural Engagement	Mastercard Institute (2026)
M-5: Children's Digital Awareness Dashboard	Cognitive Dependency (Gen Alpha)	EPRS (2025)
M-6: Informed Consent Flow	All Four Dimensions	All four sources

Table 5. Alignment of Ethical UI/UX Models with the Four-Dimensional Comparative Matrix.

6. Discussion, Policy Implications & Recommendations

The convergent findings across four independent sources produce a coherent and urgent picture. The Awareness–Action Gap is not a peripheral phenomenon but the central mechanism through which AI-generated deepfake content achieves influence. Its persistence across warned, aware, and digitally literate populations (Clark & Lewandowsky; Batool & Bashir) disqualifies purely educational or transparency-based responses as standalone solutions.

For Generation Alpha specifically, the structural dimension of this vulnerability demands systemic rather than individual intervention. As the first AI-native generation, they are not simply encountering AI content — they are being cognitively constituted within it. The EPRS (2025) establishes that legislative frameworks lag behind technological proliferation, and EU-level monitoring indicators for children's online behaviour are non-existent. This represents a governance emergency requiring urgent multilateral action.

Domain	Recommendation	Evidence Base
Platform Design	Mandate Ethical UI/UX frameworks (Models M-1 to M-6) as regulatory compliance standards, not voluntary design choices	Clark & Lewandowsky (2026); EPRS (2025)
Education Policy	Embed AI literacy as a core competency from primary school; update cyberbullying and conduct codes to explicitly include deepfakes	EPRS (2025); Batool & Bashir (2026)
Legislation	Supplement AI Act Article 50 labelling with enforceable removal provisions; establish EU-level child deepfake monitoring indicators	Clark & Lewandowsky (2026); EPRS (2025)
Industry / Tech	Move from transparency-only to 'safe-by-design' generative AI architectures; enforce age-appropriate access controls	Mastercard (2026); EPRS (2025)
Parental Awareness	Develop AI literacy campaigns targeting parents; only 37% are currently aware their children use AI tools (EPRS, 2025)	EPRS (2025)
Research	Develop longitudinal studies on Gen Alpha cognitive development under AI-native conditions; empirically test the six Ethical UI/UX models	All four sources

Table 6. Recommendations by domain, with evidence base.

POLICY NEXUS: Regulatory frameworks must move beyond transparency mandates to encompass architectural design obligations. The interface is the primary site at which the Awareness–Action Gap can be closed — not through user willpower, but through design that makes transparent engagement the default, not the exception.

6.1 Limitations of this Synthesis

This paper is constrained by the scope of its four source documents. Clark & Lewandowsky's experiments were conducted with adult populations (ages 19–76), limiting direct generalisability to Generation Alpha. Batool & Bashir's sample is geographically specific (Pakistan) and may not represent Gen Z globally. The Mastercard report uses survey methodology subject to social desirability bias. Future research should include longitudinal developmental studies specifically targeting children aged 5–15 under controlled AI-native exposure conditions.

7. Conclusion

This paper has synthesised four peer-reviewed and institutional sources to construct a comprehensive analytical account of how AI-generated visual content affects the cognitive and social development of Generations Z and Alpha. Through the Four-Dimensional Comparative Matrix — spanning Digital Awareness, Behavioural Engagement, Cognitive Dependency, and Critical Resilience — a coherent and troubling picture emerges: the threat is neither uniform across generations nor addressable through individual awareness alone.

The Awareness–Action Gap emerges as the defining empirical phenomenon: the consistent, experimentally verified divergence between knowing content is fabricated and being cognitively resistant to its influence. For Generation Z, this gap is navigable with the right structural supports. For Generation Alpha, it represents a developmental emergency — a generation forming its epistemic frameworks within an environment of proliferating synthetic reality, with inadequate institutional or technological protection.

The six Ethical UI/UX models proposed herein translate this analytical picture into actionable design principles. They represent not merely aesthetic improvements to existing interfaces but a fundamental reconceptualisation of the designer's ethical responsibility in the age of generative AI. Design, in this paradigm, is not merely aesthetic — it is epistemically, developmentally, and politically consequential.

CLOSING THESIS: In an environment where AI-generated content is indistinguishable from authentic reality, where warnings are insufficient to neutralise influence, and where the first AI-native generation is forming its worldview, the interface becomes the last line of cognitive defence. Ethical UI/UX is not a feature — it is a fundamental right.

References

- [01] Batool, M., & Bashir, S. (2026). AI-generated media perception and AI attitudes among Gen Z: The mediating role of digital literacy. *Academia Journal of Psychology*, 5(3). [https://doi.org/10.63056/academia.5.3\(b\).2026.1769](https://doi.org/10.63056/academia.5.3(b).2026.1769)
- [02] Chesney, B., & Citron, D. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107, 1753–1820.
- [03] Clark, S., & Lewandowsky, S. (2026). The continued influence of AI-generated deepfake videos despite transparency warnings. *Communications Psychology*, 4, 13. <https://doi.org/10.1038/s44271-025-00381-9>
- [04] Clayton, K., Nyhan, B., Porter, E., & Wood, T. (2020). Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags. *Political Behaviour*, 42, 1073–1095.
- [05] Ecker, U. K. H., Lewandowsky, S., & Tang, D. T. (2010). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition*, 38, 1087–1100.
- [06] European Parliament Research Service [EPRS]. (2025). *Children and deepfakes* (Briefing PE 775.855). Author: Mar Negreiro. [https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI\(2025\)775855](https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2025)775855)
- [07] European Union. (2024). Artificial Intelligence Act, Article 50. Transparency obligations for certain AI systems. <https://artificialintelligenceact.eu/article/50/>
- [08] Europol. (2022). *Facing reality? Law enforcement and the challenge of deepfakes*. Publications Office of the European Union.
- [09] Groh, M., Epstein, Z., Firestone, C., & Picard, R. (2022). Deepfake detection by human crowds, machines, and machine-informed crowds. *Proceedings of the National Academy of Sciences*, 119, e2110013119.
- [10] Internet Watch Foundation. (2024). *Annual report on AI-generated child sexual abuse material*. IWF Publications.
- [11] Köbis, N. C., Doležalová, B., & Soraperra, I. (2021). Fooled twice: People cannot detect deepfakes but think they can. *iScience*, 24. <https://doi.org/10.1016/j.isci.2021.103364>
- [12] Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13, 106–131.
- [13] Mastercard Institute. (2026). *The next growth generation: Gen Z and Gen Alpha shift from scrolling to shaping culture*. Mastercard Incorporated.
- [14] Negreiro, M. (2025, February). *Children and generative AI*. EPRS Briefing. European Parliament.
- [15] Ofcom. (2024). *Children's media use and attitudes report 2024*. Office of Communications, United Kingdom.
- [16] Schiff, K. J., Schiff, D. S., & Bueno, N. S. (2025). The liar's dividend: Can politicians claim misinformation to evade accountability? *American Political Science Review*, 119, 71–90.
- [17] UNICEF. (2024). *AI and children: Policy brief on generative AI literacy*. United Nations Children's Fund.
- [18] Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6. <https://doi.org/10.1177/2056305120903408>
- [19] Wittenberg, C., Epstein, Z., Péloquin-Skulski, G., Berinsky, A. J., & Rand, D. G. (2025). Labeling AI-generated media online. *PNAS Nexus*, 4, pgaf170.