Measuring User Responses to Online Age Verification Mechanisms Through A Deceptive Experiment

Yanzi Lin, Cheng Zhang, Madelyne Xiao*, Lorrie Faith Cranor, Sarah Scheffler
{*veronic2, chengz3, lorrie, sschefffl*}*@andrew.cmu.edu, *madelyne@princeton.edu*
*Carnegie Mellon University, *Princeton University*

***Abstract*—The U.S. Supreme Court's 2025 decision in *Free Speech Coalition v. Paxton* established that age verification systems must be "adequately tailored" to avoid undue burdens on adults' First Amendment rights. We conducted an IRB-approved, deceptive web experiment (n = 1635) examining how different age verification methods affect adults' decisions to access R-rated content. Completion rates varied significantly: checkbox self-attestation achieved 99%, government-ID methods only 23–27% regardless of data-handling reassurances, email-based estimation 86%, and AI facial estimation 51%. Follow-up survey responses (n = 884) revealed concerns about privacy, surveillance, and data security. These findings suggest that technically robust verification methods may be ineffective in practice if users systematically decline to comply.**

The U.S. Supreme Court's 2025 decision in *Free Speech Coalition v. Paxton* established that age verification systems must be "adequately tailored" to prevent undue burdens on adults' First Amendment rights [1]. In February 2026, the FTC issued a policy statement announcing it would not pursue COPPA enforcement against operators that collect personal information solely for age verification purposes, provided they meet certain data minimization and security conditions [2]. Current age verification approaches represent distinct technical architectures with different implications for user privacy and compliance. These range from simple checkbox self-declaration to more complex systems involving government-issued identification documents, third-party commercial databases, and emerging biometric technologies like AI-based facial age estimation [3], [4]. Empirical work to date has primarily relied on surveys measuring stated attitudes toward age verification. These studies generally find conditional public support for age checks, with significant privacy concerns [5], [6], [7], [8]. However, this literature largely relies on self-reported preferences rather than observing how individuals behave when actually prompted with age verification requirements.

This study empirically examines how different age verification methods and accompanying data handling disclosures influence user behavior.Specifically, we aim to answer three research questions: **(RQ1)** How do different age assurance methods affect participants' decisions to access R-rated content? **(RQ2)** Within a given age assurance method, do data-handling reassurance statements influence participants' access decisions? **(RQ3)** What are people's general attitudes toward online age assurance?

We conduct an IRB-approved, deceptive web experiment ($n = 1{,}635$) to measure completion rates across randomly assigned age verification conditions, with over 225 participants completing the experiment in each condition. We then followed up with a subset of participants via a survey ($n = 884$) to examine their perceptions and attitudes toward online age verification.[1] These user behaviors should be taken into account when determine what constitutes "adequate tailoring" of age verification systems. Moreover, as many U.S. states are transitioning toward digital ID, these results should be considered for a smooth roll-out of future systems for verifying things other than age as well.

In our study design, to set up a realistic verification context and avoid biasing participants' responses, participants were informed during recruitment that the study concerned R-rated romantic media and would involve viewing short movie clips containing "suggestive scenes, adult themes, and explicit language typical of R-rated films." In the online consent form, we also disclosed that part of the study was purposefully misrepresented for scientific reasons, but that participants would receive a debrief later and could choose to withdraw from the study entirely if they wished.

After obtaining consent, the user would be informed that we must verify their age through a trusted third-party age verification provider before accessing the study content. They would be prompted to complete the process through AgeGuardian, which was a mock service set up by the researchers. AgeGuardian would run the user through a mock (but seemingly real) randomly-selected form of age verification: checkbox self-attestation, taking a picture of a government-issued ID, using AI facial age estimation, or an email-based age estimation using third-party databases. Additional variants tested different "reassurance statements" ac-

1. This study extends a pilot ($n = 99$) conducted in a gambling site context, which found similar patterns of avoidance for government-ID-based conditions [9].

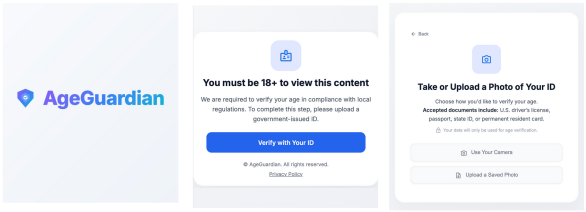

Figure 1. Screenshots of the researcher-controlled mock age verification service, AgeGuardian. Users were deceptively instructed to age-verify with AgeGuardian before conducting a "R-rated romantic media" study. After completing (or not completing) age verification, the user would be directed to (or sent) a debrief informing them that the main purpose of the study was to measure whether or not they would (appear to) verify their age using their assigned age verification condition on AgeGuardian.

companying the government-ID conditions, in which we disclosed information about how collected data would be used and retained. In the Simple Reassurance condition, participants were informed that "Your data will be used only for age verification." In the Compound Reassurance condition, participants were additionally informed that "We will delete your identifying information after access has been granted." (See Figure 1 for some mockup pictures of AgeGuardian.)

The AgeGuardian service was programmed to leave all sensitive information on the user's device; even though it would appear to upload their ID or face picture, no data was ever transmitted over the network or seen by the researchers. We logged only web activity data, including clicks and timestamps. If the user age verified, they would be directed to a debrief and survey. If the user did not complete age verification (closed browser tab or timed out on the webpage) then the user would be automatically sent an email to the address provided in the consent form. The email informed them that they had already completed Part 1 of the study (declining age verification was considered a valid completion of the experiment) and invited them to return to complete a short follow-up survey. The message also disclosed that participation in Part 2 would not require age verification. Participants who completed age verification, clicked on the "Exit Study" button on the study site during age verification, or returned via the custom link were shown a detailed debrief explaining the true purpose of the study and were given the option to withdraw their data entirely, retain their web experiment data without completing the survey, or proceed to the survey portion of the study. In the survey, participants answered Likert-scale and open-ended questions regarding their comfort, perceived risk, effectiveness, and convenience of the age verification process.

To preserve ecological validity and recruit a diverse sample, we advertised across three platforms (Reddit Ads, Meta Ads, and TrafficJunky, an adult content ad network) using keyword- and interest-based targeting.

We found that completion rates differed significantly across age verification conditions (Chi-square test of independence: $\chi^2(6) = 619.82$, $p < .001$). Checkbox self-attestation produced the highest completion rate (99%) and was associated with the highest self-reported comfort, while government-issued ID–based methods exhibited substantially lower completion and comfort. Pairwise comparisons of completion proportions with Holm–Bonferroni correction indicate that all government-ID conditions—no reassurance (23%), simple reassurance (24%), and compound reassurance (27%)—had significantly lower completion rates than checkbox, AI facial age estimation (51%), and email-based age estimation (86%) (all $p < .001$). Within government-ID conditions, adding reassurance statements did not significantly change verification outcomes. Among the non–government-ID methods, email-based age estimation (86%) exhibited significantly higher completion rates than all other methods except checkbox ($p < .001$), while AI facial age estimation (51%) had the third-highest completion rate and differed significantly from the remaining methods ($p < .001$). The government-ID condition with a liveness check had the lowest completion rate overall (18%).

These findings raise important questions about the consumer protection implications of age verification policies. If more intrusive verification methods substantially reduce service access, including among adults legitimately seeking information or services, then such policies risk creating barriers that disproportionately affect certain user populations. Effective age verification systems must balance child safety with adults' ability to access online services.

To better understand additional factors that may influence participants' decisions to complete or decline age verification, we plan to run logistic regression models examining predictors of completion outcomes, with demographic characteristics and prior experiences with online age verification included as covariates. We will also conduct thematic analysis of survey responses to gain qualitative insights into participants' experiences with AgeGuardian and their broader perspectives on service-level age verification. Beyond these analyses, future work could examine how compliance varies across different service contexts (e.g., online dating, social media, gambling) and whether alternative technical solutions (e.g., device-based verified credentials) achieve higher user acceptance.

Qualitative analysis is ongoing, but preliminary findings point to several recurring themes: (1) mixed support for age verification as a general policy approach, (2) concerns about privacy and surveillance, and (3) concerns about security, including confusion about which parties were collecting age-related data and reluctance to share sensitive personal information with unfamiliar entities.

Acknowledgments

This study was funded by NSF grants 2207216 and DGE-2039656.

References

[1] Supreme Court of the United States, "Free speech coalition v. paxton," 2025.

[2] Federal Trade Commission, "FTC issues COPPA policy statement to incentivize the use of age verification technologies to protect children online," February 2026.

[3] E. Rescorla, Z. Arnao, and A. Cooper, "Age assurance online: A technical assessment of current systems and their limitations," Knight-Georgetown Institute, Tech. Rep., January 2026.

[4] S. Liu and S. Scheffler, "Adequately tailoring age verification regulations," *arXiv preprint arXiv:2601.20241*, 2026.

[5] Australian eSafety Commissioner, "Public perceptions of age verification for limiting access to pornography," 2021.

[6] Family Online Safety Institute, "Making sense of age assurance: Cross-national survey results from the u.s., u.k., and france," 2022.

[7] Ofcom, "Adult users' attitudes to age verification on adult sites," 2022.

[8] P. J. Wright and D. Herbenick, "U.S. Attitudes Toward Age-Verification for Online Pornography," *Journal of Sex Research*, 2025, forthcoming.

[9] Y. V. Lin, V. Lieu, C. Zhang, W. Zhang, W. Hu, L. F. Cranor, and S. Scheffler, "Carded by the internet: Measuring user responses to online age assurance mechanisms," *USENIX SOUPS Poster*, 2025.