Discussion Paper Series

IZA DP No. 18715

June 2026

Context and the Perception of Harm: Evidence from Online and In-Person Sexual Harassment

Alessandra Foresta
University of Southampton

Valentina Tonei
University of Southampton
and IZA@LISER

Martina Vecchi
University of Southampton

IZA DP No. 18715 June 2026

Context and the Perception of Harm: Evidence from Online and In-Person Sexual Harassment*

Abstract

We provide causal evidence that the context in which harmful conduct occurs shapes how it is evaluated. We study this using a vignette-based survey experiment that holds behaviour constant while randomising whether an incident of sexual harassment occurs online or in person. Online settings generate a systematic discount in perceived seriousness (7% of the mean), with larger effects on willingness to report (13% of the mean) and preferred sanctions. The discount is concentrated in image-based harassment and larger among male respondents. In addition, it is not moderated by either direct or vicarious experience of harassment, suggesting that it may reflect normative perceptions of online harm as less serious rather than lack of exposure alone. These context-dependent distortions have implications for the enforcement of emerging legal protections, victim support, and the design of public communication around digital abuse.

JEL classification

D83, J16, C93, K42, D91

Keywords

sexual harassment, misperceptions, perceived harm, social norms, digital environments

Corresponding author

Valentina Tonei
v.tonei@soton.ac.uk

* The authors gratefully acknowledge helpful comments and suggestions from Pamela Giustinelli, Johanna Rickne, Michael Vlassopoulos, seminar participants at the C-BEAR seminar at the University of Southampton, and participants at the 15th meeting of the Behavioral and Experimental Economics Network (BEEN). All errors are our own.

1 Introduction

Digital communication has fundamentally transformed how individuals interact, shifting a growing share of social and economic activity from physical to online environments. As a result, behaviours that traditionally occurred in person increasingly take place in digital settings, where visibility, social cues, and accountability differ. A central implication of this transformation is that identical actions may be evaluated differently depending on the context in which they occur. If perceptions of harm depend on context rather than on the underlying conduct, this may distort reporting, sanctioning, and support for regulation. So we ask: do individuals assign the same severity to harmful behaviour when it occurs online as when it occurs in person?

We study this question in the context of sexual harassment, defined as behaviours that humiliate or demean people based on their sex or gender (Berdahl, 2007), whose associated economic costs and the policy relevance are well established. Sexual harassment affects labour supply, job mobility, and earnings, and contributes to persistent gender inequalities in the workplace (Folke and Rickne, 2022; Adams-Prassl et al., 2024; Dahl and Knepper, 2026; Coly and Suteau, 2025), especially in sectors where digital presence is important to advancement (Hegewisch et al., 2021).

Sexual harassment increasingly occurs in online environments (Office for National Statistics, 2023), prompting both domestic and international policy responses: the UK Online Safety Act has recently introduced new criminal offences targeting online gendered abuse, and online sexual harassment is now recognised internationally as a form of technology-facilitated gender-based violence, a public health and human rights priority (United Nations Population Fund, 2024). Recent studies document that online sexual harassment can have negative psychological and behavioural consequences, such as risk of depression, anxiety, and trauma symptoms, comparable to those of in-person harassment (Benítez-Hidalgo et al., 2024; Iroegbu et al., 2024; McLaughlin et al., 2017; Ståhl and Dennhag, 2021). However, it remains unclear whether comparable behaviours are perceived differently across online and in-person contexts and how such evaluations vary across individuals and incident characteristics.

Previous work in psychology and economics on perceptions of sexual harassment has focused mainly on in-person incidents, particularly in workplace settings. Antecol

and Cobb-Clark (2006), for example, show that changes over time in the willingness to label unwanted behaviour as sexual harassment are associated with broader shifts in attitudes. Rhodes et al. (2025) instead study the difference between how individuals perceive in-person sexual harassment compared to their perceived social norm. Folke et al. (2026) show, in a field experiment among Norwegian military recruits, that providing information about peers' beliefs regarding what constitutes sexual harassment improves knowledge and shifts tolerant attitudes, pointing to perceived social norms as a key driver of how harassment is categorised. To the best of our knowledge, Lam and Mesch (2025) are the only authors to examine directly the difference in perceptions between online and in-person sexual harassment. Using survey data on Israeli adult internet users, they find that online sexual harassment is perceived as less serious and less harmful than in-person harassment, with gender as a significant predictor. However, their study is observational and focuses on three closely related perceptions—seriousness, harmfulness, and wrongfulness—without distinguishing between types of harassment. It therefore cannot separate the effect of the online setting from differences in the content of the incidents compared, nor identify how any discount varies across harassment types or respondent characteristics.

We answer this question by conducting a vignette-based survey experiment on a representative sample of UK adults. We randomly vary the setting (online versus in-person), the victim's gender, and the type of harassment (verbal versus image-based), while holding the underlying behaviour constant. We measure perceived severity, blame attribution, and reporting and punishment preferences, alongside incentivised behavioural measures, including donations to a victim support charity and support for stronger regulation. By holding constant the key characteristics of each scenario and using a within-person identification strategy, we isolate the causal effect of context on perceived harm.

We document a robust context effect in the evaluation of harm. Holding behaviour constant, identical harassment is perceived as significantly less serious when it occurs online than when it occurs in person, indicating a systematic "online discount" in perceived severity, amounting to approximately 0.4 points on a 7-point scale (one-third of a standard deviation). This discount is not uniform across incident types: it is concentrated in image-based harassment, with the online context sharply compressing

the severity gap between harassment types, while the discount for verbal harassment is small and largely insignificant. The online discount is also significantly larger among male respondents and is not moderated by prior personal or vicarious experience of sexual harassment. These perception gaps are also accompanied by differences in stated behavioural intentions: individuals are less willing to support reporting and prefer less severe punishment when the same conduct is framed as occurring online. Incentivised behavioural measures suggest that support for stronger regulation is near-universal and more strongly associated with political ideology than with personal experience of harassment. At the same time, both regulatory support and donations are significantly related to respondents' self-reported attitudes.

Our study relates first to the literature on misperceptions, which examines how individuals form beliefs under imperfect information and how these beliefs shape behaviour (Bénabou and Tirole, 2016; Haaland et al., 2023; Jensen, 2010; Wiswall and Zafar, 2015; Kuziemko et al., 2015). In this literature, individuals rely on prior beliefs and noisy signals rather than directly observing relevant outcomes, generating systematic distortions. A closely related strand examines misperceptions about others' attitudes specifically, showing that perceived social norms can diverge substantially from actual peer attitudes, with direct consequences for behaviour (Bursztyn et al., 2020; Bursztyn and Yang, 2022). We apply this framework to the evaluation of harmful behaviour, arguing that online environments may provide less informative or less salient signals about harm, thereby increasing reliance on priors and generating an "online discount" in perceived severity even when underlying conduct is comparable.

Second, we contribute to a literature showing that beliefs about harm are endogenous to the informational environment in which they are formed. Perceptions of crime, for example, respond to the salience of crime-related information rather than only to underlying crime itself (Mastrorocco and Minale, 2018), and subjective beliefs about risk and enforcement are shaped by both personal experience and informational frictions (Lochner, 2007; Salm and Vollaard, 2021). More broadly, recent work shows that disclosure and framing can affect attitudes through belief updating and salience channels (Keita et al., 2024). Our contribution is to show that such distortions can arise even when underlying conduct is fixed: the context in which harm occurs—online rather than in-person—can itself alter perceived severity.

Third, we contribute to the economics of sexual harassment. Existing work shows that harassment affects occupational sorting, pay inequality, and job mobility (Laband and Lentz, 1998; Folke and Rickne, 2022; Hersch, 2011), and that reporting decisions depend on expectations about enforcement, retaliation, and corroboration, as well as on social norms and individual characteristics (Antecol and Cobb-Clark, 2006; Cheng and Hsiaw, 2022; Mukherjee and Dasgupta, 2022; Levy and Mattsson, 2022; Batut et al., 2026). A smaller literature further examines how sexual harassment is perceived, documenting gaps between individual and social evaluations and their variation across groups (Rhodes et al., 2025; Folke et al., 2026). We complement this work by identifying a systematic distortion in how harassment is evaluated across contexts, and by showing that this distortion is not uniform but varies predictably across individuals and situations. In particular, if online harassment is perceived as less severe, individuals may be less willing to report, sanction, or support regulation—especially among male respondents and for the most severe forms of online harassment—thereby weakening enforcement and potentially perpetuating inequality.

The rest of the paper is organised as follows. Section 2 describes the conceptual framework and states the hypotheses we test in our experiment, while Section 3 describes the experimental design and data. Section 5 presents the results, and Section 6 concludes.

2 Conceptual Framework and Hypotheses

Sexual harassment encompasses behaviours that humiliate or demean individuals on the basis of their sex or gender (Berdahl, 2007). Many individuals fail to recognise certain conduct as harassment. While coercive or physically intrusive conduct is widely identified as such, recognition is considerably less consistent for verbal and image-based forms (Cortina and Areguin, 2021; Roehling et al., 2022), with direct consequences for reporting and bystander intervention (Cheung et al., 2018). We focus on two of its most prevalent non-contact forms in the UK, verbal harassment and image-based harassment (Office for National Statistics, 2023; Bleakley et al., 2025).

Consider an individual i evaluating an incident j of sexual harassment. We define true harm H_j as the psychological, social, and economic consequences of incident j; for comparable behaviours, H_j is similar across settings (Benítez-Hidalgo et al., 2024; Iroegbu et al., 2024; Ståhl and Dennhag, 2021). Individuals do not observe H_j directly but form a subjective perception $\hat{H}_{ij}$, which depends on the setting $M_j \in \{\text{online}, \text{in-person}\}$ in which the incident occurs and on individual-specific beliefs and information. We define the *online discount* as the expected difference in perceived harm across settings,

$$\Delta_i \equiv \mathbb{E}[\hat{H}_{ij} \mid M_j = \text{online}] - \mathbb{E}[\hat{H}_{ij} \mid M_j = \text{in-person}]. \tag{1}$$

Hypothesis 1 (Setting effect). Identical conduct is perceived as less harmful when it occurs online than in person, so that the online discount is negative ($\Delta_i < 0$).

Since perceived harm shapes reporting decisions, sanctioning preferences, and support for regulation (Cheung et al., 2018; Antecol and Cobb-Clark, 2006), a negative discount would translate into weaker institutional and social responses to equally harmful conduct. Two mechanisms can generate a negative discount, with distinct implications for how it varies with individual experience.

The first is an *information friction* mechanism. In in-person settings, harassment is more socially visible, emotionally immediate, and frequently discussed, generating relatively precise signals about its consequences. In online settings, harassment is more diffuse, often anonymous, and routinely normalised within digital environments, producing noisier signals. Formally, consistent with Bayesian belief formation, perceived harm can be written as a weighted average of a signal s_m, capturing the information available to the individual about the consequences of the incident, and a context-specific prior μ_m, with weight λ_m on the signal,

$$\hat{H}_{ij} = \lambda_m s_m + (1 - \lambda_m)\mu_m, \quad m \in \{\text{online}, \text{in-person}\}. \tag{2}$$

If online environments generate lower signal precision ($\lambda_{\text{online}} < \lambda_{\text{in-person}}$), individuals rely more heavily on prior beliefs, and perceived severity falls if those priors understate true harm.

The second is a *normative categorisation* mechanism. Rather than reflecting imperfect signal quality, a negative discount may arise from a stable categorical belief

that online environments are spaces of lower accountability and less serious conduct. Under this mechanism, the prior μ_{online} is socially determined, shaped by cultural narratives, perceived social norms, and the collective framing of digital spaces, and is downward-biased relative to its in-person counterpart ($\mu_{\text{online}} < \mu_{\text{in-person}}$). Distinct from information friction, the prior under this mechanism is not merely noisy but anchored to a social consensus that direct experience does not update. Evidence that perceived social norms diverge substantially from actual peer attitudes, sustaining behaviours that most individuals privately oppose, is provided by Bursztyn et al. (2020) and reviewed by Bursztyn and Yang (2022); the collective framing of digital spaces as less serious may persist precisely because individuals misperceive it as the social consensus, rather than as a genuinely shared belief. Folke et al. (2026) show that targeting information about peers' beliefs regarding what constitutes harassment improves knowledge and shifts tolerant attitudes, pointing to perceived social norms as a key driver of how harassment is categorised.

Hypothesis 2a (Information friction). The online discount is smaller among respondents with direct experience of online sexual harassment, since personal exposure generates more precise signals and reduces reliance on prior beliefs.

Hypothesis 2b (Normative categorisation). The online discount is invariant to prior experience of online sexual harassment, since μ_{online} is socially anchored and not updated through individual exposure.

H2a and H2b are mutually exclusive predictions about the same empirical pattern, tested directly in Section 5.3.

Hypothesis 3 (Heterogeneity). The online discount varies systematically across respondent characteristics and incident attributes. Male respondents are expected to exhibit a larger discount than non-male respondents, consistent with evidence that men perceive a wider range of behaviours as acceptable and rate harassment as less serious (Rotundo et al., 2001; Lam and Mesch, 2025). Image-based harassment is expected to attract a larger discount than verbal harassment, since unsolicited explicit images are routinely normalised in online settings in ways that verbal conduct is not (Bleakley et al., 2025). How the discount varies with victim gender is examined as an exploratory dimension in Section 5.2.

3 Experimental Design and Data

3.1 Sample and Recruitment

Data were collected through an online survey conducted in the United Kingdom between September and November 2025. Participants were recruited via Prolific, and eligibility was restricted to UK residents aged 18 and above. The median completion time was 9 minutes, and the base payment was £1.50. Data quality was ensured through CAPTCHA verification, a self-reported attention screen before the vignette section, and an instructional attention check toward the end of the survey. The experimental design, empirical analysis, and hypothesis were pre-registered in the AEA RCT Registry (AEARCTR-0016406) and received ethical approval from the University of Southampton before the data collection started.

The final sample includes 602 respondents (Table 1). The mean age of respondents is 44 years (SD = 13.45). In terms of gender, 48% identify as male, 50% as female, and a small share identify as non-binary or prefer not to disclose their gender. The sample is highly educated, with 56% holding a bachelor's degree or higher, and predominantly in full-time employment (59%). In terms of ethnicity, 89% of respondents identify as White. Regarding prior exposure, 36% report personal or vicarious experience of online sexual harassment, and 51% report personal or vicarious experience of in-person sexual harassment.[1]

3.2 Experimental Design

Our objective is to understand the determinants of individuals' attitudes toward sexual harassment, particularly whether they depend on the context (online vs in-person) in which these episode occurs. Observational data are not well suited to answering this question because incidents in digital environments differ systematically from those in-person in ways that are difficult to control for, including anonymity, the relationships between the individuals involved, and the severity of the behaviour.

[1] A comparison with UK population benchmarks (Census 2021) is in Table A.8. The sample closely matches the UK adult population in terms of gender and regional distribution, but over-represents degree-holders (56% versus 34%) and full-time workers (59% versus 40%), is younger on average (44 versus 49 years), and is somewhat less ethnically diverse (89% White versus 82%). These patterns are typical of online survey panels.

Table 1. Demographic characteristics of participants

	Mean	SD
Age	44.06	(13.45)
Gender:		
Male	0.48	(0.50)
Female	0.50	(0.50)
Non-binary/Third gender	0.00	(0.06)
Prefer not to say	0.01	(0.10)
Harassment experience:		
Online	0.36	(0.48)
In person	0.52	(0.50)
Gross annual income:		
<£20,000	0.14	(0.35)
£20k–39,999	0.32	(0.47)
£40k–59,999	0.22	(0.41)
£60k–99,999	0.26	(0.44)
>£100k	0.06	(0.24)
Employment:		
Full-time	0.59	(0.49)
Part-time	0.17	(0.38)
Other (unemployed/student/etc.)	0.24	(0.43)
Education:		
No degree	0.43	(0.50)
Bachelor or higher	0.56	(0.50)
Prefer not to say	0.01	(0.10)
Political ideology:		
Left	0.35	(0.48)
Centre	0.49	(0.50)
Right	0.15	(0.36)
N	602	

Note: all variables are dummies with the exception of *Age*.

To isolate the role of context itself, we use a survey-based vignette experiment that holds the interaction's content constant while varying only the environment in which it takes place.

Survey experiments are increasingly used in economics to study beliefs, perceptions, and normative evaluations that cannot be easily observed through revealed choices (Hainmueller et al., 2015; Stantcheva, 2023; Haaland et al., 2023). In this approach, respondents evaluate hypothetical scenarios in which key attributes are randomly assigned, allowing causal identification of how each attribute affects judgments. Factorial vignette designs are particularly useful when the situations of interest are rare, sensitive, or difficult to observe in naturally occurring data (Auspurg and Hinz, 2015; Wallander, 2009).

Each respondent evaluated four randomly assigned hypothetical scenarios (e.g., the vignettes) describing incidents of sexual harassment. In addition, respondents evaluate one neutral filler vignette, used to reduce demand effects and priming. The use of multiple scenarios per respondent generates within-individual variation in incident attributes while keeping cognitive burden limited.

Each vignette varies along three experimentally manipulated dimensions: the setting in which the incident occurs, the victim's gender and the type of harassment (Table 2). The first dimension captures the interaction environment. In the *in-person* condition, the incident occurs in a physical public space and is described as taking place "while walking in the street". In the *online* condition, the same interaction occurs through a digital platform and is described as taking place "in a direct message on a social media platform". The content of the interaction, the victim's age, and the anonymity of the perpetrator are held constant across conditions, so that the only difference between the two scenarios is whether the incident occurs online or in-person. Victim gender varies across three categories: *man*, *woman*, and *non-binary person*. Finally, we distinguish between *verbal harassment*, defined as unsolicited explicit sexual comments, and *image-based harassment*, defined as unsolicited explicit sexual images or exposure. These two harassment types were selected to represent the most prevalent forms of non-contact sexual harassment across both settings. Verbal harassment, comprising unsolicited explicit sexual comments, is the most widely reported type of sexual harassment in England and Wales, identified as the single most

common form and experienced by 3% of adults in the past year, a figure rising to 23% among women aged 16 to 24 (Office for National Statistics, 2023). Image-based harassment, comprising unsolicited explicit images online and physical exposure in person, is among the most prevalent forms of digital sexual harassment, with around 40% of women aged 18 to 34 in the UK reporting having received unsolicited explicit images (Bleakley et al., 2025) and more than one in eight women in England and Wales reporting experience of indecent exposure (Office for National Statistics, 2023). Combining these dimensions yields 12 distinct vignette types (2 settings × 3 victim genders × 2 harassment types).

Table 2. Experimental design

Dimension	**Levels**
Setting	In-person: "while walking in the street" Online: "in a direct message on a social media platform"
Victim gender	Man Woman Non-binary person
Type of harassment	Image-based: "has a stranger expose themselves"; "receives unsolicited explicit sexual images from an anonymous user" Verbal: "receives unsolicited explicit sexual comments from a stranger"; "receives unsolicited explicit sexual comments from an anonymous user"

The wording of vignettes was kept constant across conditions except for the manipulated dimensions, so that differences in responses can be attributed to the experimental factors rather than to narrative framing. For example, the same behaviour is described either as occurring "while walking in the street" or "in a direct message on a social media platform", while all other elements remain unchanged. For image-based harassment, the in-person condition describes physical exposure by a stranger in a public space and the online condition describes receiving unsolicited explicit images

in a direct message. While the physical medium necessarily differs, both conditions share the key characteristics of the harassment: it is unsolicited, sexually explicit, and perpetrated by a stranger. The design therefore constructs the closest real-world equivalents of image-based sexual harassment in each setting, consistent with the paper's goal of comparing contextual analogues rather than physically identical acts, which would be impossible across online and in-person environments by definition.

Each respondent observes four vignettes randomly drawn from the full set of possible scenarios, and the order of presentation is randomised across respondents. Random assignment of vignette attributes ensures that the experimental factors are orthogonal to respondents' characteristics, providing internally valid estimates of the effect of setting, victim gender, and harassment type on perceived harm.

Respondents were compensated with a base payment for participation and were eligible for additional performance-based bonuses through incentivised tasks embedded in the survey. These tasks included a coordination game used to elicit social norms and a donation decision linked to a real charity supporting victims of online harassment.

The full instructions shown to participants, together with a detailed description of the experimental design, are reported in Appendix D.

3.3 Outcome Variables

After each vignette, respondents evaluate the incident along several dimensions intended to capture different aspects of perceived harm and moral evaluation. The set of outcomes is designed to distinguish between perceived severity, emotional responses, willingness to report the incident, responsibility attribution, perceived social norms and preferred sanctions. Collecting multiple measures allows us to assess whether the experimental manipulations affect all evaluation dimensions equally or operate through specific channels.

Perceived severity is measured by asking respondents how serious they consider the situation described in the vignette using a seven-point scale ranging from "not serious at all" to "very serious". Emotional reactions are measured on the same scale by asking how angry the respondent feels toward the perpetrator. To assess the likelihood of formal intervention, we measure willingness to report the incident by

asking respondents how likely they would be to report it to someone in a position of authority (1 = extremely unlikely; 7 = extremely likely). Responsibility attribution is elicited separately for the two parties involved in the incident. Respondents indicate how much responsibility they assign to the targeted person (victim blame) and to the perpetrator (perpetrator responsibility), with higher values indicating greater responsibility.

Perceived social norms are measured by asking respondents how "most people" would evaluate the behaviour described in the vignette on a five-point scale ranging from "very socially appropriate" to "very socially inappropriate". Following Krupka and Weber (2013), this question is incentivised via a coordination-game mechanism in which respondents receive a bonus for matching the modal response, shifting responses toward beliefs about prevailing norms rather than private moral evaluations.

We measure sanction preferences by asking respondents to choose the most appropriate response from six ordered options with increasing levels of severity (no punishment; verbal warning; formal police warning; restorative justice meeting; community service; prison sentence), which we treat as an ordered indicator of punitiveness.

In addition to vignette-level evaluations, two outcomes are measured once per respondent. Participants were asked to indicate whether they would support a petition calling for stronger legal protections against online sexual harassment. Petition-signing decisions are commonly used in survey experiments to assess whether stated preferences translate into real behavioural choices (Haaland et al., 2023; Grigorieff et al., 2020). Following Dechezleprêtre et al. (2022), respondents are told that the results of the petition will be communicated to the relevant authorities (the UK government), making the decision closer to a real policy expression rather than a purely hypothetical response. Finally, we elicit costly support for victims through an incentivised allocation decision. Respondents enter a lottery for £100 and indicate how much of the potential prize (0–100) they wish to donate to Victim Support UK, a charity assisting victims of crime and harassment.[2] Monetary allocation tasks of this type are standard in experimental economics to measure prosocial preferences and charitable giving (Andreoni, 1989; Andreoni and Miller, 2002; DellaVigna et al., 2012; Exley, 2016). Because the decision affects a real payment, this measure captures

[2]https://www.victimsupport.org.uk/

willingness to incur a monetary cost rather than purely stated attitudes.

Table 3 presents descriptive statistics for the main outcome variables. On average, respondents evaluate the incidents as highly harmful: perceived seriousness, anger toward the perpetrator, and willingness to report the incident all have mean values close to 6 on the 1–7 scale. Responsibility attributed to the targeted person is very low (around 1), while responsibility attributed to the perpetrator is near the top of the scale.

In terms of preferred sanctions, the most frequently selected option is a formal police warning (47%), followed by a verbal warning (24%). Support for stronger legal protection against online sexual harassment is high, with 89% of respondents indicating that they would support the petition. In the incentivised donation task, participants allocate on average £22.28 to Victim Support.

3.3.1 Factor Analysis

We use factor analysis to construct a single summary index of the attitudinal outcomes (perceived seriousness, anger, willingness to report, victim blame, perpetrator blame, social norm, and preferred sanction), which reduces dimensionality and mitigates concerns about multiple testing across correlated measures. The model is estimated via maximum likelihood on complete cases (n = 2,438 vignettes, filler excluded).

Only Factor 1 has an eigenvalue greater than one (2.458) and is retained (Kaiser, 1960; Cattell, 1966), accounting for 96% of the variance. Table 4 reports the factor loadings, uniqueness, and scoring coefficients. Loadings are positive for all variables except victim blame, with the strongest associations for perceived seriousness, anger, willingness to report, and preferred sanction, consistent with these outcomes capturing a common dimension of perceived harm. We extract the Factor 1 score and include it in the main analysis as *Index*.

Table 3. Descriptive statistics of dependent variables

	Mean	SD	Min	Max
Perceived seriousness	5.86	(1.18)	1.00	7.00
Anger	5.72	(1.33)	1.00	7.00
Willingness to report	5.02	(1.90)	1.00	7.00
Blame:				
Victim	1.23	(0.88)	1.00	7.00
Perpetrator	6.84	(0.81)	1.00	7.00
Social norm	4.70	(0.73)	1.00	5.00
Preferred sanction:				
No punishment	0.06	(0.24)	0.00	1.00
Verbal warning	0.24	(0.43)	0.00	1.00
Formal police warning	0.47	(0.50)	0.00	1.00
Restorative justice	0.04	(0.20)	0.00	1.00
Community service	0.09	(0.29)	0.00	1.00
Prison sentence	0.09	(0.29)	0.00	1.00
Petition support†	0.89	(0.31)	0.00	1.00
Donation to Victim Support (£)†	22.28	(23.78)	0.00	100.00
No. vignettes		7,224		
No. respondents		602		

Notes: Perceived seriousness, anger, and willingness to report are measured on a 7-point Likert scale (1 = not at all; 7 = very much). Victim and perpetrator blame are measured on the same 7-point scale (1 = no responsibility; 7 = full responsibility). Social norm is measured on a 5-point scale (1 = very socially appropriate; 5 = very socially inappropriate), elicited via an incentivised coordination game following Krupka and Weber (2013). Preferred sanction categories are reported as proportions summing to one. All variables except petition support and donation are measured at the vignette level.

† Measured once per respondent ($N = 602$). Petition support is an indicator equal to one if the respondent would sign a petition calling for stronger legal protections against online sexual harassment. Donation is the amount (£0–£100) allocated to Victim Support UK in an incentivised lottery task.

Table 4. Factor loadings

	Factor 1	Uniqueness	Scoring Coefficients
	(1)	(2)	(3)
Perceived seriousness	0.836	0.301	0.362
Angry	0.807	0.349	0.299
Willingness to report	0.747	0.442	0.231
Blame: victim	-0.174	0.970	-0.038
Blame: perpetrator	0.079	0.994	0.026
Socially Norm	0.323	0.896	0.054
Preferred sanction	0.640	0.590	0.156

Notes: Factor analysis is used to create an aggregate measure derived from the individual outcome variables. Entries represent maximum likelihood estimates from a factor analysis model restricted to one factor.

4 Empirical Strategy

We use four estimating equations, each corresponding to a specific set of hypotheses or outcomes. All specifications are estimated by OLS with standard errors clustered at the respondent level.

To test Hypothesis 1 (Setting Effect), we estimate the following equation for each vignette-level outcome:

$$Y_{ij} = \alpha + \beta \, Online_{ij} + \varepsilon_{ij} \tag{3}$$

where Y_{ij} is a vignette-level outcome for respondent i evaluating scenario j, and $Online_{ij}$ is a binary indicator equal to one if the scenario describes online sexual harassment and zero if it describes in-person harassment. The parameter β captures the average difference in perceived harm between otherwise identical online and offline incidents. For outcomes where larger values indicate stronger condemnation (perceived seriousness, anger, willingness to report, perpetrator blame, social inappropriateness, preferred sanction), H1 predicts $\beta < 0$. For victim blame, where larger values indicate greater victim responsibility, H1 predicts $\beta > 0$.

To test the incident-level component of Hypothesis 3 (Heterogeneity), let T_{ij} be an

indicator for image-based harassment ($T_{ij} = 0$ for verbal, the omitted category) and let $\mathbf{Z}_{ij} = (Woman_{ij},\ NonBinary_{ij})'$ be a vector of victim-gender indicators (omitted category: male victim). We first estimate separate models for each incident dimension, then verify findings in a joint fully-interacted specification:

$$\begin{aligned} Y_{ij} = &\alpha + \beta\, Online_{ij} + \theta_T T_{ij} + \boldsymbol{\theta}_Z' \mathbf{Z}_{ij} + \boldsymbol{\psi}'(T_{ij} \times \mathbf{Z}_{ij}) \\ &+ \delta_T(Online_{ij} \times T_{ij}) + \boldsymbol{\delta}_Z'(Online_{ij} \times \mathbf{Z}_{ij}) + \boldsymbol{\phi}'(Online_{ij} \times T_{ij} \times \mathbf{Z}_{ij}) + \varepsilon_{ij} \end{aligned} \tag{4}$$

The scalar θ_T and vector $\boldsymbol{\theta}_Z$ capture the main effects of harassment type and victim gender on in-person baseline evaluations. δ_T captures how the online discount varies with harassment type (for the baseline victim gender), and $\boldsymbol{\delta}_Z$ captures how it varies with victim gender (for the baseline harassment type). $\boldsymbol{\psi}$ captures the interaction between victim gender and harassment type in the in-person baseline. The vector $\boldsymbol{\phi}$ contains three-way interactions testing whether the online discount by harassment type differs across victim-gender group.

To test the respondent-level component of Hypothesis 3 and to discriminate between Hypotheses 2a and 2b, let N_i be an indicator for non-male respondents ($N_i = 0$ for male, the omitted category) and $\mathbf{E}_i = (E_{O,i},\ E_{P,i})'$ be a vector of prior harassment experience indicators, where $E_{O,i}$ denotes experience of online harassment and $E_{P,i}$ denotes experience of in-person harassment (each measured as personal and/or vicarious exposure within the respective setting). We first estimate separate models for each respondent dimension, then verify findings in a joint fully-interacted specification:

$$\begin{aligned} Y_{ij} = &\alpha + \beta\, Online_{ij} + \theta_N N_i + \boldsymbol{\theta}_E' \mathbf{E}_i \\ &+ \delta_N(Online_{ij} \times N_i) + \boldsymbol{\delta}_E'(Online_{ij} \times \mathbf{E}_i) \\ &+ \boldsymbol{\psi}'(N_i \times \mathbf{E}_i) + \boldsymbol{\phi}'(Online_{ij} \times N_i \times \mathbf{E}_i) + \varepsilon_{ij} \end{aligned} \tag{5}$$

N_i equals one if the respondent identifies as a woman or non-binary person.[3] The component of $\boldsymbol{\delta}_E$ corresponding to $E_{O,i}$ is the key test discriminating between H2a and H2b: H2a (information friction) predicts this coefficient is negative, so that the discount narrows among respondents with prior online experience; H2b (normative

[3]Non-binary respondents are grouped with women due to the small cell size.

categorisation) predicts it equals zero, so that the discount is invariant to experience. The vector $\boldsymbol{\phi}$ tests whether the experience moderation differs between male and non-male respondents.

The two respondent-level behavioural outcomes petition support (P_i) and donation to Victim Support (D_i), are each measured once per respondent and are not linked to the vignette randomisation. We estimate descriptive OLS regressions of the form:

$$Y_i = \alpha + \mathbf{X}_i'\gamma + \varepsilon_i \tag{6}$$

where $\mathbf{X}_i$ is a vector of respondent characteristics including demographics and, in a second specification, respondent-level summaries of vignette attitudes (average condemnation of online harassment and the online–offline norm gap). The γ coefficients are descriptive associations rather than causal effects.

Because vignette attributes are randomly assigned within respondents, $Online_{ij}$ is orthogonal to both observed and unobserved respondent characteristics, and OLS consistently estimates the causal effect of setting without further controls. Including individual fixed effects would additionally absorb respondent-level differences in scale use, improving precision, but is not required for identification. We nonetheless adopt OLS without individual fixed effects as the primary specification for two reasons. First, the respondent-level outcomes in Equation 6 cannot be estimated with respondent fixed effects. Second, Equation 5 includes time-invariant respondent characteristics whose main effects are collinear with individual fixed effects. However, we report OLS with individual fixed effects as a robustness check in Appendix C. As additional robustness checks we estimate Romano-Wolf stepdown adjusted p-values, clustered at the respondent level, to address multiple hypothesis testing across outcome families (Appendix B).

5 Results

5.1 The Online Discount

Our results reveal a consistent and precisely estimated "online discount" across all attitudinal outcomes (Figure 1 and Table C.1).[4] Perceived seriousness falls by 0.394 points on the 1–7 scale, equivalent to 6.7% decrease. Anger declines by a similar magnitude (7.6%). Both effects are statistically significant at the 1% level and point to a systematic dampening of affective responses to identical behaviour when it is framed as occurring online.

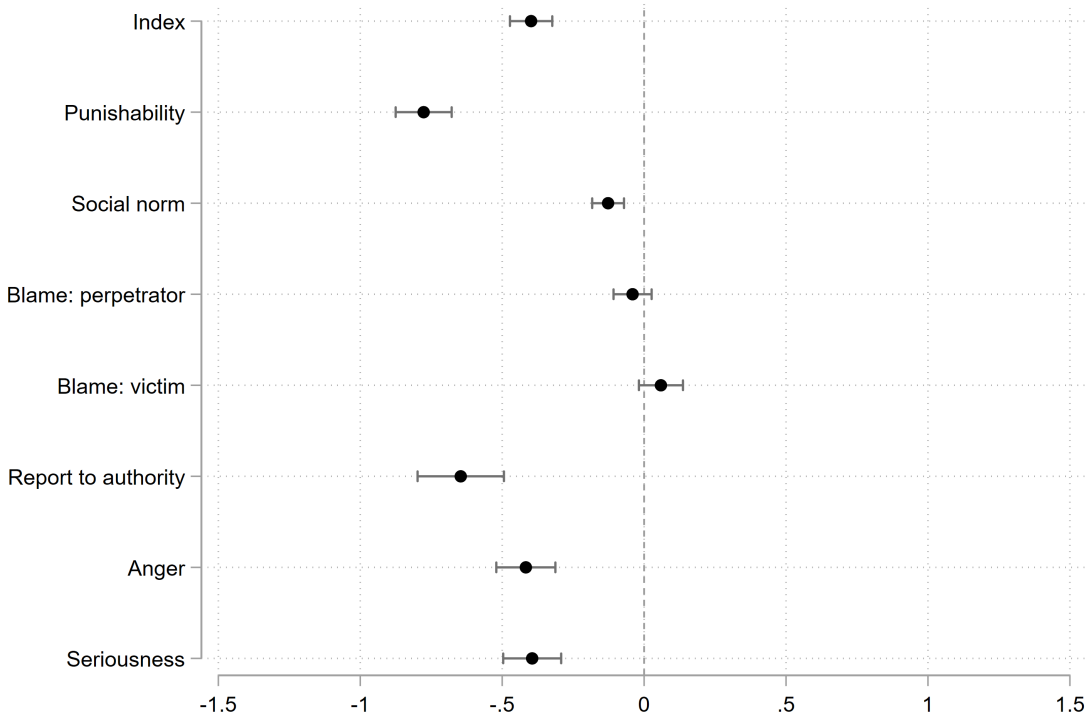
The largest effect concerns reporting behaviour. Willingness to report the incident to an authority decreases by 0.668 points, a 13.3% relative decrease. This is nearly twice the proportional effect on perceived seriousness, indicating that the online context does not merely reduce perceived harm but substantially lowers the propensity to seek formal intervention—a channel directly relevant to enforcement. Social norm evaluations shift in the same direction but to a much more modest extent: online incidents are rated 0.131 points less socially inappropriate (2.8% of the mean of 4.70). The divergence between the effect on reporting and on social norms suggests that the reduction in formal intervention in the online setting is not driven purely by a shift in perceived social appropriateness; additional mechanisms, such as perceived barriers to reporting, lower efficacy beliefs, or uncertainty about jurisdiction, likely play a role.

Two further results deserve attention. Neither victim blame nor perpetrator blame differs significantly across settings: responsibility attributed to the targeted person is virtually unchanged, consistent with the discount not operating through victim-blaming, while responsibility attributed to the perpetrator is marginally lower online (-0.036), suggesting a slight tendency to hold perpetrators less accountable in digital environments.

The online discount extends to sanction preferences, with Figure 2 decomposing the shift across six categories. Online incidents are significantly more likely to attract no punishment (+3.8 pp) or a verbal warning (+14.6 pp), and significantly less likely

[4]Results are robust to Romano-Wolf multiple hypothesis testing corrections applied to the fully interacted specifications (Appendix B). Heterogeneity results in Sections 5.2 and 5.3 are also robust to the inclusion of individual fixed effects (Appendix C).

Figure 1. Effect of online vs in-person sexual harassment



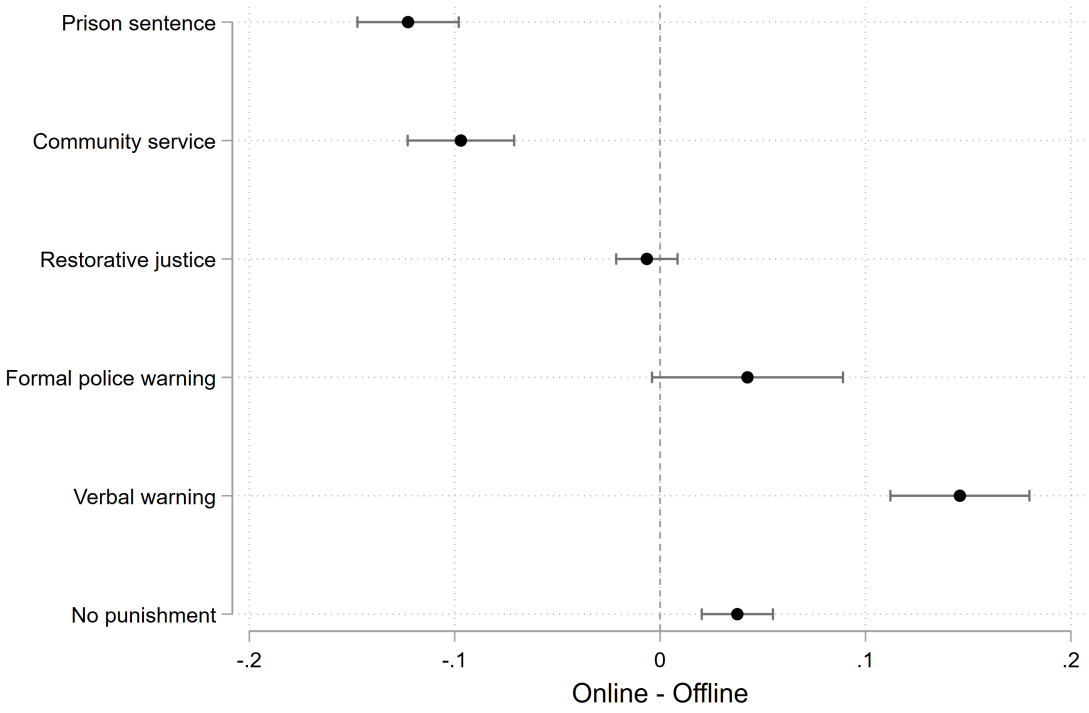
Notes: OLS estimates. Points show the estimated online discount for each outcome; bars indicate 95% confidence intervals with standard errors clustered at the respondent level. Outcome variables are described in Section 3. The corresponding regression estimates are reported in Table C.1.

to result in community service (−9.7 pp) or imprisonment (−12.3 pp). There is a smaller and only marginally significant increase in the likelihood of a formal police warning (+4.3 pp). Together, these estimates point to a wholesale shift in preferred sanctions toward the lenient end of the scale: identical conduct, when framed as occurring online, is treated as warranting at most an informal response.

5.2 Heterogeneity by Incident Characteristics

The online discount is not uniform across incident characteristics. It is concentrated almost entirely in image-based harassment and does not vary significantly by victim gender. These findings emerge from separate interaction models for each incident dimension (Tables C.2 and C.3), and are confirmed in a fully interacted

Figure 2. Effect of online vs in-person harassment on preferred punishment



Notes: OLS estimates. Points show the estimated shift in preferred sanction category when the incident occurs online; bars indicate 95% confidence intervals with standard errors clustered at the respondent level.

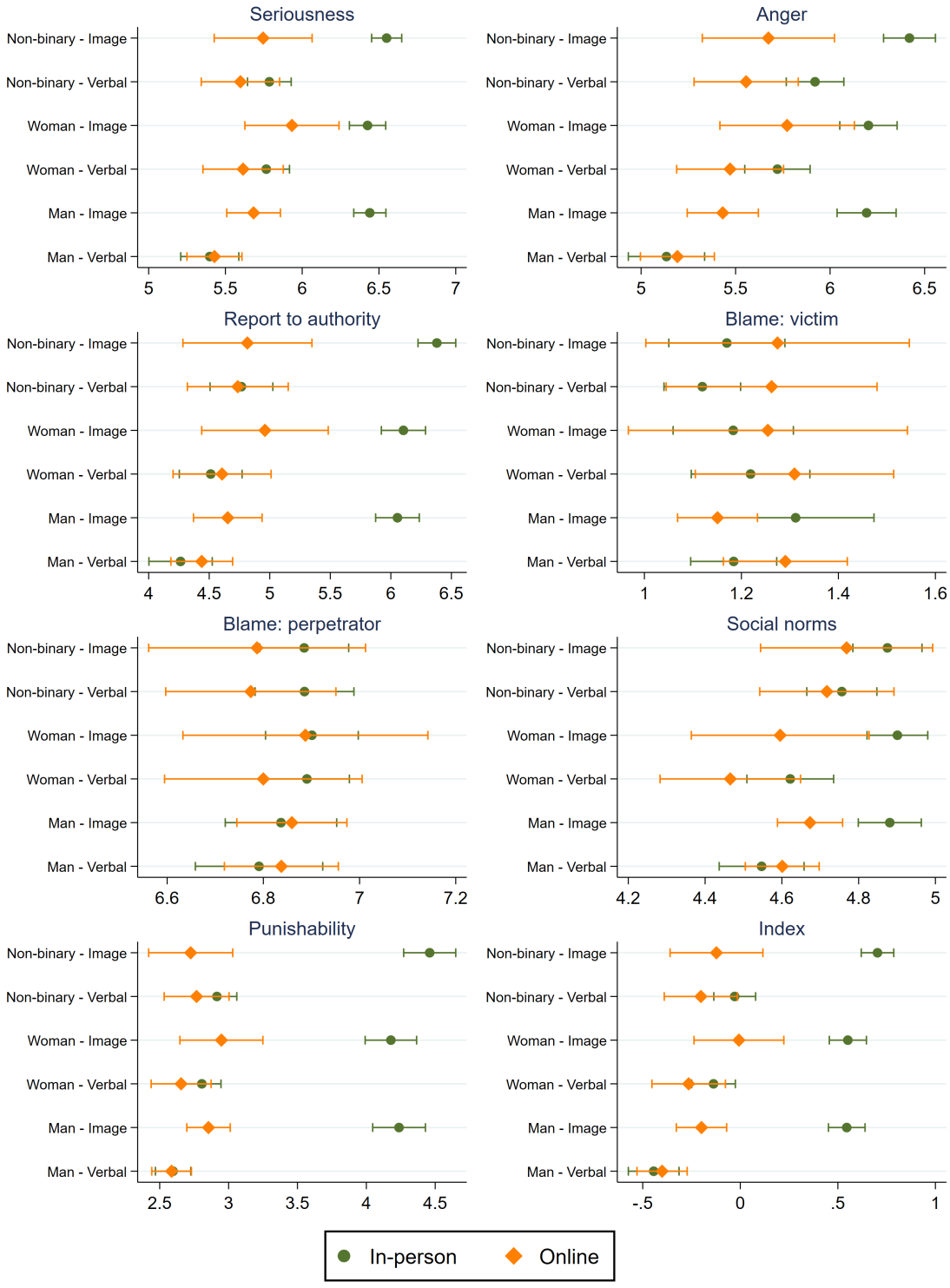
specification (Table A.4; Figure 3).

More specifically, we find that image-based harassment is evaluated far more seriously than verbal harassment in the in-person baseline, with perceived seriousness 0.810 points higher, anger 0.601 points higher, and reporting intentions 1.620 points higher (Table C.2). The Online × Image interaction is −0.589 for seriousness and −1.292 for reporting ($p < 0.01$), implying a total online effect of −0.678 on seriousness (11.6% of the mean) and −1.294 on reporting (25.8% of the mean). This is consistent with the normative categorisation mechanism: image-based conduct is the form most specifically associated with the online environment (Bleakley et al., 2025), making it most susceptible to downgrading when the online setting activates a more permissive normative reference class.

The online–offline gap does not vary significantly by victim gender (Table C.3).

The interactions Online × Woman and Online × Non-binary are statistically indistinguishable from zero across all main outcomes. In terms of level differences, incidents targeting women and non-binary victims are rated significantly more seriously than those targeting men across all outcomes, but this severity ranking does not translate into differential online discounting.

Figure 3. Perceived harmfulness of online vs. in-person sexual harassment by incident characteristic



Notes: OLS estimates. Points show estimated online discounts by incident type and victim gender; bars indicate 95% confidence intervals with standard errors clustered at the respondent level. Outcome variables are described in Section 3. The corresponding regression estimates are reported in Table A.4.

5.3 Heterogeneity by Respondent Characteristics

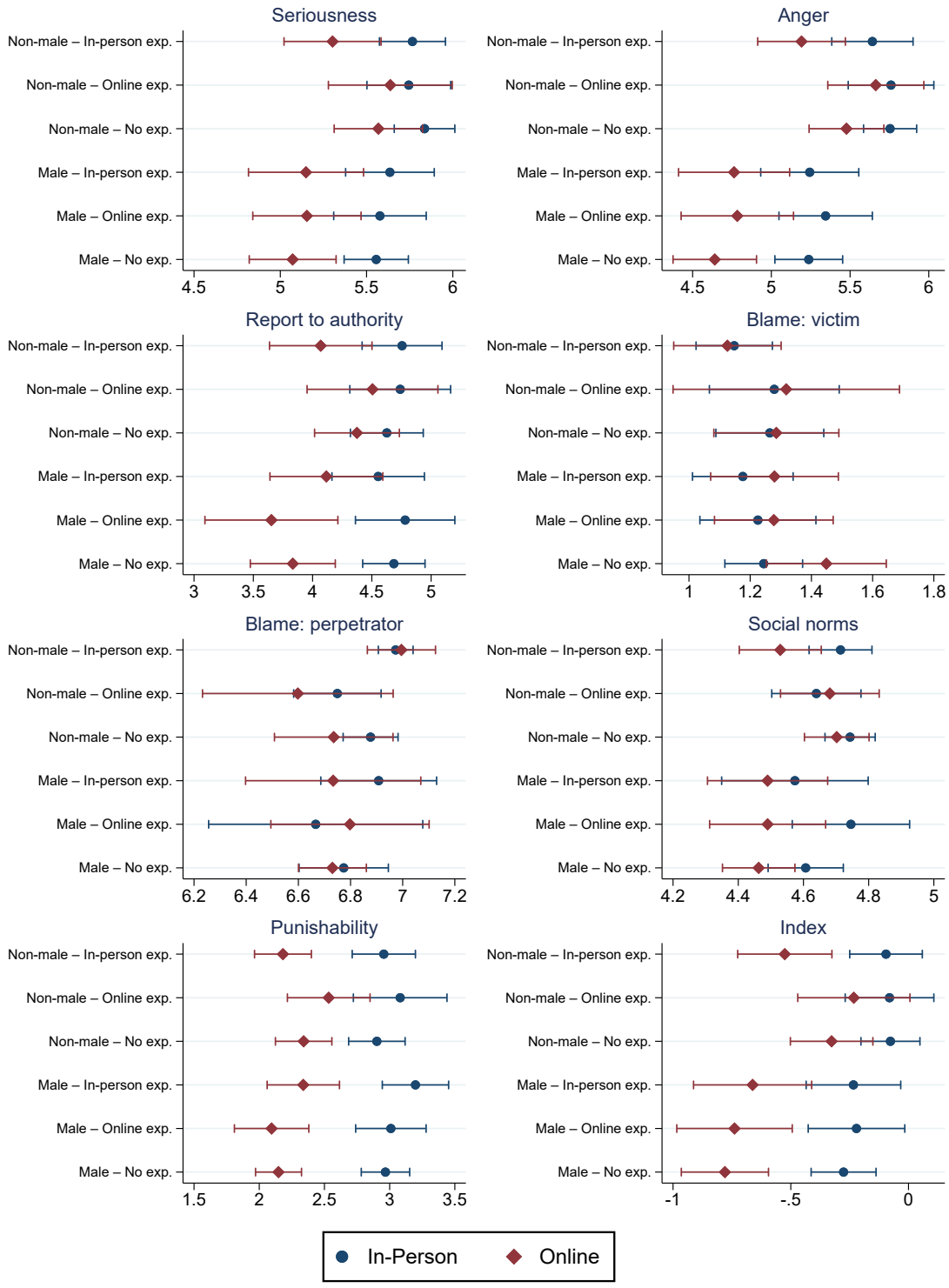
The online discount varies substantially by respondent gender but is invariant to prior harassment experience.

Male respondents evaluate sexual harassment as systematically less severe than non-male respondents. In-person seriousness is 0.29 points lower among men and anger 0.53 points lower, corresponding to differences of 5.0% and 9.3% of the respective outcome means; reporting intentions do not differ significantly by gender at the in-person baseline (Table A.7). Sanction preferences follow the same pattern, with male respondents more likely to select no punishment and less likely to recommend a formal police warning.

Beyond these level differences, the online discount is substantially larger among men. The Online × Non-male interaction is positive and significant for anger (+0.32), reporting (+0.62), and the Index (+0.26) (Table A.7; Figure 4), indicating that non-male respondents discount online incidents less strongly. The online reduction in reporting intentions reaches −0.85 for men, while for non-male respondents reporting intentions are broadly similar across the online and in-person contexts. The online discount in reporting behaviour is therefore largely a male-respondent phenomenon, consistent with evidence that men apply more permissive normative standards to sexual harassment in less formalised contexts (Rotundo et al., 2001; Lam and Mesch, 2025).

Prior harassment experience does not moderate the online discount. Our experience measures capture both personal and vicarious exposure within each setting. Even so, the interaction terms between Online and prior experience are small and statistically indistinguishable from zero across all main outcomes (Table A.7), including when prior exposure is measured specifically as online harassment. These findings do not reconcile with the information-friction mechanism (H3a), which predicts that the discount should narrow with direct online experience. Instead, they are more consistent with normative categorisation (H3b), whereby the discount reflects a stable social representation of online harm as less serious that persists regardless of exposure.

Figure 4. Perceived harmfulness of online vs. in-person harassment by respondent characteristic



Notes: OLS estimates. Each point shows the predicted level of the outcome for a given respondent type and setting; bands indicate ±1 standard error clustered at the respondent level. Outcome variables are described in Section 3. The corresponding regression estimates are reported in Table A.7.

5.4 Behavioural Outcomes

The attitudinal and reporting outcomes examined above exploit within-respondent experimental variation in vignette assignment. We complement them with two incentivised behavioural measures collected once per respondent: willingness to sign a petition calling for stronger legal protections against online sexual harassment, and the share of a potential £100 lottery prize donated to Victim Support UK. Because these outcomes are not linked to the vignette randomisation, the estimates in Table 5 are descriptive associations rather than causal effects; they nonetheless allow us to characterise which respondents are less willing to support action against online sexual harassment in both low-cost and costly domains.

Panel A shows that petition support is driven primarily by political ideology rather than demographics. Left-leaning respondents are 12.5 percentage points more likely to sign the petition than right-leaning respondents, and centrist respondents are 8.3 percentage points more likely. Prior exposure to online sexual harassment is also positively associated with petition support (+7.0 pp). The donation regression displays a similar ideological gradient, with left-leaning respondents giving more; age is positively associated with giving, while prior in-person harassment is negatively associated with donations to support victims of online sexual harassment. Gender, employment, income, and education are not significantly related to either behavioural outcome.

Panel B relates both outcomes to respondent-level summaries of vignette attitudes. The online–offline norm gap is not significantly associated with petition support or donation once controls are included. By contrast, average condemnation of online harassment is positively and significantly associated with both outcomes: respondents who express lower condemnation in the vignette task are also less likely to sign the petition and donate less. Downstream support for action against online sexual harassment is therefore better predicted by general condemnation than by the gap between own and perceived social norms.

Table 5. Behavioural outcomes and norm-gap links

Panel A. Behavioural outcomes

	(1) Petition (AME)	(2) Donation (OLS)
Age	-0.001	0.271***
	(0.001)	(0.077)
Gender:		
Non-male	0.021	2.457
	(0.025)	(2.092)
Employment:		
Full-time	-0.037	1.643
	(0.035)	(2.638)
Other	0.010	4.237
	(0.035)	(3.023)
Income:		
60k or more	0.032	-0.285
	(0.027)	(2.156)
Education:		
Bachelor or higher	-0.029	-0.942
	(0.026)	(2.062)
Prior online harassment	0.070**	3.324
	(0.033)	(2.225)
Prior in-person harassment	-0.006	-5.183**
	(0.027)	(2.226)
Left-leaning	0.125***	6.541**
	(0.037)	(2.915)
Centrist	0.083***	3.918
	(0.030)	(2.596)
Observations	590	596

Panel B. Norm-gap links

	(1) Petition (AME)	(2) Donation (OLS)
Online-offline norm gap	-0.013	-0.052
	(0.013)	(1.147)
Average online condemnation	0.074***	4.995***
	(0.012)	(1.086)
Observations	424	426

Notes: Panel A reports associations between respondent characteristics and behavioural outcomes. Employment base category is part-time; income base category is below £60k; education base category is no bachelor's degree; political ideology base category is right-leaning. Panel B relates petition and donation to two respondent-level summaries of vignette attitudes: the online-offline norm gap (mean online norm evaluation minus mean offline norm evaluation) and average condemnation of online harassment (mean vignette condemnation score for online incidents). Petition support is estimated via logit with average marginal effects reported; donation is estimated via OLS. Robust standard errors in parentheses. N varies across panels because each regression uses respondent-level complete cases. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6 Conclusions

This paper shows that context alone can generate systematic distortions in the perceived severity of harm. Holding conduct constant, identical incidents of sexual harassment are evaluated as substantially less serious when they occur online, with the discount concentrated in image-based harassment, larger among male respondents, and invariant to prior experience. These results advance the economics of sexual harassment (Folke and Rickne, 2022; Hersch, 2011; Laband and Lentz, 1998; Rhodes et al., 2025) by identifying a systematic, context-dependent distortion in harm evaluation, and contribute to the literature on misperceptions and context-dependent belief formation (Haaland et al., 2023; Mastrorocco and Minale, 2018; Keita et al., 2024). The null effect of prior, whether personal and/or vicarious and whether online or in person, suggests that the online discount is not simply driven by lack of exposure or correctable information gaps. Instead, it is consistent with normative categorisation of online as less harmful, connecting our findings to the literature on misperceived social norms and their persistence (Bursztyn et al., 2020; Bursztyn and Yang, 2022; Folke et al., 2026).

These findings have direct implications for emerging policy frameworks, including the UK Online Safety Act, which aim to strengthen enforcement and accountability in digital environments. Our results point to a demand-side constraint. Public support for enforcement may be weaker precisely where intervention is most needed, and individuals are less willing to sanction identical conduct when it occurs online. The null effect of prior experience is particularly relevant for policy design. If the discount reflected information gaps, awareness campaigns about the consequences of online harm would be a natural corrective. Its persistence regardless of prior exposure points instead to a stable normative categorisation of digital environments as spaces of lower accountability, one that information provision alone may not shift. Our norm data are consistent with this picture. Respondents believe others condemn online harassment less than they do, suggesting the discount may partly reflect misperceived rather than actual social tolerance and creating a potentially self-reinforcing dynamic between public under-response and institutional inaction. Norm-communication campaigns that correct misperceptions about how seriously others evaluate online harassment may prove more effective, and consistent enforcement by courts, employers, and

platforms can reinforce such corrections by providing credible public signals about what society considers acceptable.

Our study also has limitations. First, while the experimental design allows us to isolate the causal effect of context by holding behaviour constant, it relies on hypothetical scenarios rather than observed real-world behaviour. Although we complement stated evaluations with incentivised measures, responses in survey settings may differ from actions in naturally occurring environments. Second, our sample focuses on a specific population and context, and perceptions may differ across countries or institutional settings with different legal frameworks and social norms. Third, our design cannot directly distinguish between the two mechanisms in the framework: the data are consistent with both information frictions and stable normative categorisation as drivers of the online discount, and both may operate simultaneously.

These limitations suggest several directions for future research. A natural next step is to examine whether providing information about the consequences of online harassment can shift perceptions and behaviour, and whether such interventions differ in effectiveness across groups. Further work could also explore how institutional features, such as platform design, moderation policies, or legal enforcement, interact with perceptions of harm. Finally, extending the analysis to other forms of harmful behaviour in digital environments would help assess the generality of context-dependent distortions in perceived severity.

Taken together, our findings suggest that the expansion of social and economic interactions into digital environments not only changes where harm occurs but also how it is perceived. Understanding and addressing these perception gaps may be key to designing effective policies in increasingly digital societies.

References

Adams-Prassl, A., Huttunen, K., Nix, E., and Zhang, N. (2024). Violence against women at work. *The Quarterly Journal of Economics*, 139(2):937–991.

Andreoni, J. (1989). Giving with impure altruism: Applications to charity and ricardian equivalence. *Journal of Political Economy*, 97(6):1447–1458.

Andreoni, J. and Miller, J. (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, 70(2):737–753.

Antecol, H. and Cobb-Clark, D. (2006). The sexual harassment of female active-duty personnel: Effects on job satisfaction and intentions to remain in the military. *Journal of Economic Behavior & Organization*, 61(1):55–80.

Auspurg, K. and Hinz, T. (2015). *Factorial Survey Experiments*. SAGE, Thousand Oaks, CA.

Batut, C., Coly, C., and Schneider-Strawczynski, S. (2026). It'sa man's world: culture of abuse,# metoo and worker flows. Technical report, CESifo Working Paper.

Bénabou, R. and Tirole, J. (2016). Mindful economics: The production, consumption, and value of beliefs. *Review of Economic Studies*, 83(4):1241–1276.

Benítez-Hidalgo, V., Henares-Montiel, J., Ruiz-Pérez, I., and Pastor-Moreno, G. (2024). Cyber sexual harassment against women and impact on health. a cross-sectional study in a representative population sample. *Journal of Public Health*, 46(1):3–11.

Berdahl, J. L. (2007). The sexual harassment of uppity women. *Journal of Applied Psychology*, 92(2):425–437.

Bleakley, P., Kelly, D., and Short, E. (2025). What's the harm? Gender and age differences in acceptance of cyberflashing, reporting intentions, and perceived prevalence. *Current Issues in Criminal Justice*.

Bursztyn, L., González, A. L., and Yanagizawa-Drott, D. (2020). Misperceived social norms: Women working outside the home in Saudi Arabia. *American Economic Review*, 110(10):2997–3029.

Bursztyn, L. and Yang, D. Y. (2022). Misperceptions about others. *Annual Review of Economics*, 14(1):425–452.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, 1(2):245–276.

Cheng, I.-H. and Hsiaw, A. (2022). Reporting sexual misconduct. *American Economic Journal: Microeconomics*, 14(2):242–277.

Cheung, H. K., Goldberg, C. B., King, E. B., and Magley, V. J. (2018). Are they true to the cause? Beliefs about organizational and unit commitment to sexual harassment awareness training. *Group & Organization Management*, 43(4):531–560.

Coly, C. and Suteau, M. (2025). The economic effects of sexual harassment in the workplace. *LSE Public Policy Review*, 3(4).

Cortina, L. M. and Areguin, M. A. (2021). Putting people down and pushing them out: Sexual harassment in the workplace. *Annual Review of Organizational Psychology and Organizational Behavior*, 8(1):285–309.

Dahl, G. B. and Knepper, M. (2026). Why is workplace sexual harassment underreported? the value of outside options amid the threat of retaliation. *American Economic Review*, 116(3):897–933.

Dechezleprêtre, A., Fabre, A., Kruse, T., Planterose, B., Sanchez Chico, A., and Stantcheva, S. (2022). Fighting climate change: International attitudes toward climate policies. *American Economic Review*, 112(12):3855–3895.

DellaVigna, S., List, J. A., and Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *Quarterly Journal of Economics*, 127(1):1–56.

Exley, C. L. (2016). Excusing selfishness in charitable giving: The role of risk. *The Review of Economic Studies*, 83(2):587–628.

Folke, O., Kosadam, A., Hanson, T., Rickne, J., and Johnsen, Å. A. (2026). Targeting attitudes to combat sexual harassment: A randomized intervention in the Norwegian military. Discussion Paper 18637, IZA Institute of Labor Economics.

Folke, O. and Rickne, J. (2022). Sexual harassment and gender inequality in the labor market. *The Quarterly Journal of Economics*, 137(4):2163–2212.

Grigorieff, A., Roth, C., and Ubfal, D. (2020). Does information change attitudes

toward immigrants? representative evidence from survey experiments. *American Economic Journal: Applied Economics*, 12(4):1–29.

Haaland, I., Roth, C., and Wohlfart, J. (2023). Measuring what people care about: Using surveys to study preferences, beliefs, and policy views. *Journal of Economic Literature*, 61(3):892–958.

Hainmueller, J., Hopkins, D. J., and Yamamoto, T. (2015). Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. *Political Analysis*, 23(1):1–30.

Hegewisch, A., Forden, J., and Mefferd, E. (2021). Paying today and tomorrow: Charting the financial costs of workplace sexual harassment.

Hersch, J. (2011). Compensating differentials for sexual harassment. *American Economic Review*, 101(3):630–634.

Iroegbu, M., O'Brien, F., Muñoz, L. C., and Parsons, G. (2024). Investigating the psychological impact of cyber-sexual harassment. *Journal of Interpersonal Violence*, 39(15–16):3424–3445.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics*, 125(2):515–548.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and psychological measurement*, 20(1):141–151.

Keita, S., Renault, T., and Valette, J. (2024). The usual suspects: Offender origin, media reporting and natives' attitudes towards immigration. *The Economic Journal*, 134(657):322–362.

Krupka, E. L. and Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3):495–524.

Kuziemko, I., Norton, M. I., Saez, E., and Stantcheva, S. (2015). How elastic are preferences for redistribution? evidence from randomized survey experiments. *American Economic Review*, 105(4):1478–1508.

Laband, D. N. and Lentz, B. F. (1998). The effects of sexual harassment on job

satisfaction, earnings, and turnover among female lawyers. *ILR Review*, 51(4):594–607.

Lam, I. and Mesch, G. S. (2025). Gender differences in public perceptions of the seriousness of offline and online sexual harassment. *Psychology, Crime & Law*, pages 1–24.

Levy, R. and Mattsson, M. (2022). The effects of social movements: Evidence from# metoo. Technical Report 3496903, SSRN.

Lochner, L. (2007). Individual perceptions of the criminal justice system. *American Economic Review*, 97(1):444–460.

Mastrorocco, N. and Minale, L. (2018). News media and crime perceptions: Evidence from a natural experiment. *Journal of Public Economics*, 165:230–255.

McLaughlin, H., Uggen, C., and Blackstone, A. (2017). The economic and career effects of sexual harassment on working women. *Gender & Society*, 31(3):333–358.

Mukherjee, A. and Dasgupta, S. (2022). "He Says, She Says": Sexism and Sexual Harassment in Higher Educational Institutions of India. *Journal of Economic Issues*, 56(2):408–415.

Office for National Statistics (2023). Experiences of harassment in england and wales: December 2023. Statistical Bulletin. Released 7 December 2023.

Rhodes, L. A., Sánchez, G. E., Espinoza-Velasteguí, N., and Borja, V. (2025). Assessing the gap between social and individual perceptions of sexual harassment. *International Review of Economics*, 72(2):28.

Roehling, M. V., Wu, D., Choi, M. G., and Dulebohn, J. H. (2022). The effects of sexual harassment training on proximal and transfer training outcomes: A meta-analytic investigation. *Personnel Psychology*, 75(1):3–31.

Rotundo, M., Nguyen, D.-H., and Sackett, P. R. (2001). A meta-analytic review of gender differences in perceptions of sexual harassment. *Journal of Applied Psychology*, 86(5):914–922.

Salm, M. and Vollaard, B. (2021). The dynamics of crime risk perceptions. *American law and economics review*, 23(2):520–561.

Ståhl, S. and Dennhag, I. (2021). Online and offline sexual harassment: Associations

of anxiety and depression in an adolescent sample. *Nordic Journal of Psychiatry*, 75(5):330–335.

Stantcheva, S. (2023). How to run surveys: A guide to creating your own identifiable variation and revealing the invisible. *Annual Review of Economics*, 15:205–234.

United Nations Population Fund (2024). Technology-facilitated gender-based violence. United Nations Population Fund.

Wallander, L. (2009). 25 years of factorial surveys in sociology: A review. *Social Science Research*, 38(3):505–520.

Wiswall, M. and Zafar, B. (2015). How do college students respond to public information about earnings? *Journal of Human Capital*, 9(2):117–169.

Appendix A Additional Results

A.1 Heterogeneity by Incident Characteristics

Table A.1. Effect of online vs in-person harassment on evaluations

	(1) Serious	(2) Angry	(3) Report to Authority	(4) Victim Blame	(5) Perpetrator Blame	(6) Social Norm	(7) Punish	(8) Index
Online	-0.39***	-0.42***	-0.65***	0.06	-0.04	-0.13***	-0.78***	-0.40***
	(0.05)	(0.05)	(0.08)	(0.04)	(0.03)	(0.03)	(0.05)	(0.04)
Mean	5.86	5.72	5.02	1.23	6.84	4.70	3.15	-0.00
SD	1.18	1.33	1.90	0.88	0.81	0.73	1.31	0.92
N	2408	2408	2408	2408	2408	2408	2408	2408

Notes: OLS estimates. Outcome variables are described in Section 3.

Standard errors clustered at the respondent level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.2. Effect of image-based vs verbal harassment on evaluations

	(1) Serious	(2) Angry	(3) Report to Authority	(4) Victim Blame	(5) Perpetrator Blame	(6) Social Norm	(7) Punish	(8) Index
Online	-0.10	-0.19**	0.08	0.11***	-0.05	-0.05	-0.10*	-0.09*
	(0.07)	(0.07)	(0.10)	(0.04)	(0.03)	(0.04)	(0.06)	(0.05)
Image-based	0.82***	0.68***	1.66***	0.05	0.02	0.24***	1.52***	0.80***
	(0.06)	(0.07)	(0.10)	(0.05)	(0.04)	(0.04)	(0.08)	(0.05)
Online x Image	-0.58***	-0.46***	-1.45***	-0.11*	0.02	-0.16***	-1.34***	-0.62***
	(0.09)	(0.09)	(0.14)	(0.06)	(0.05)	(0.05)	(0.09)	(0.06)
Mean	5.86	5.72	5.02	1.23	6.84	4.70	3.15	-0.00
SD	1.18	1.33	1.90	0.88	0.81	0.73	1.31	0.92
N	2408	2408	2408	2408	2408	2408	2408	2408

Notes: OLS estimates without. Outcome variables are described in Section 3.

Omitted categories: in-person setting, verbal harassment.

Standard errors clustered at the respondent level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.3. Effect of image-based vs verbal harassment on evaluations

	(1) Serious	(2) Angry	(3) Report to Authority	(4) Victim Blame	(5) Perpetrator Blame	(6) Social Norm	(7) Punish	(8) Index
Online	-0.37***	-0.35***	-0.62***	-0.03	0.03	-0.08*	-0.70***	-0.35***
	(0.08)	(0.09)	(0.12)	(0.06)	(0.05)	(0.04)	(0.08)	(0.06)
Non-binary victim	0.18**	0.30***	0.15	-0.05	0.08	0.05	0.07	0.16***
	(0.07)	(0.08)	(0.12)	(0.06)	(0.06)	(0.05)	(0.09)	(0.06)
Woman victim	0.25***	0.50***	0.41***	-0.10**	0.07	0.10**	0.27***	0.28***
	(0.07)	(0.08)	(0.12)	(0.05)	(0.04)	(0.04)	(0.09)	(0.05)
Online x Non-binary	0.01	-0.10	-0.03	0.16*	-0.11	-0.04	0.01	-0.04
	(0.10)	(0.11)	(0.16)	(0.09)	(0.07)	(0.07)	(0.11)	(0.08)
Online x Woman	-0.09	-0.09	-0.06	0.10	-0.12*	-0.11	-0.24**	-0.10
	(0.09)	(0.11)	(0.16)	(0.07)	(0.07)	(0.07)	(0.11)	(0.07)
Mean	5.86	5.72	5.02	1.23	6.84	4.70	3.15	-0.00
SD	1.18	1.33	1.90	0.88	0.81	0.73	1.31	0.92
N	2408	2408	2408	2408	2408	2408	2408	2408

Notes: OLS estimates. Outcome variables are described in Section 3.

Omitted categories: in-person setting, male victim.

Standard errors clustered at the respondent level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.4. Online vs in-person harassment by incident characteristics

	(1) Serious	(2) Angry	(3) Report to Authority	(4) Victim Blame	(5) Perpetrator Blame	(6) Social Norm	(7) Punish	(8) Index
Online	0.03	0.06	0.17	0.11	0.05	0.05	-0.01	0.04
	(0.12)	(0.13)	(0.17)	(0.07)	(0.07)	(0.06)	(0.08)	(0.08)
Image-based	1.04***	1.06***	1.79***	0.13	0.05	0.33***	1.64***	0.99***
	(0.10)	(0.12)	(0.15)	(0.09)	(0.07)	(0.07)	(0.11)	(0.08)
Non-binary victim	0.37***	0.59***	0.25	0.03	0.10	0.07	0.21**	0.31***
	(0.11)	(0.12)	(0.17)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)
Woman victim	0.39***	0.79***	0.50***	-0.06	0.09	0.21***	0.32***	0.42***
	(0.11)	(0.11)	(0.17)	(0.05)	(0.07)	(0.07)	(0.08)	(0.07)
Online x Image	-0.79***	-0.82***	-1.58***	-0.27**	-0.02	-0.26***	-1.37***	-0.79***
	(0.15)	(0.17)	(0.23)	(0.11)	(0.09)	(0.08)	(0.15)	(0.11)
Online x Non-binary	-0.22	-0.42**	-0.20	0.04	-0.16	-0.09	-0.14	-0.22**
	(0.16)	(0.17)	(0.23)	(0.12)	(0.10)	(0.10)	(0.12)	(0.11)
Online x Woman	-0.18	-0.31**	-0.08	-0.02	-0.14	-0.21**	-0.14	-0.17*
	(0.14)	(0.15)	(0.23)	(0.08)	(0.10)	(0.09)	(0.12)	(0.10)
Image x Non-binary	-0.38***	-0.58***	-0.20	-0.16	-0.04	-0.05	-0.27*	-0.30***
	(0.13)	(0.15)	(0.20)	(0.13)	(0.09)	(0.09)	(0.14)	(0.10)
Image x Woman	-0.28**	-0.56***	-0.18	-0.08	-0.05	-0.22**	-0.10	-0.26***
	(0.13)	(0.16)	(0.20)	(0.10)	(0.09)	(0.09)	(0.15)	(0.10)
Online x Image x Non-binary	0.45**	0.64***	0.34	0.25	0.10	0.11	0.29	0.36**
	(0.21)	(0.23)	(0.30)	(0.17)	(0.14)	(0.13)	(0.20)	(0.15)
Online x Image x Woman	0.17	0.44**	0.04	0.23*	0.04	0.19*	-0.21	0.14
	(0.18)	(0.21)	(0.29)	(0.14)	(0.14)	(0.12)	(0.19)	(0.14)
Mean	5.86	5.72	5.02	1.23	6.84	4.70	3.15	-0.00
SD	1.18	1.33	1.90	0.88	0.81	0.73	1.31	0.92
Observations	2408	2408	2408	2408	2408	2408	2408	2408

Notes: OLS estimates. Outcome variables are described in Section 3.

Omitted categories: in-person setting, verbal harassment, male victim.

Standard errors clustered at the respondent level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

A.2 Heterogeneity by Respondent Characteristics

Table A.5. Heterogeneity on evaluations by respondent gender

	(1) Serious	(2) Angry	(3) Report to Authority	(4) Victim Blame	(5) Perpetrator Blame	(6) Social Norm	(7) Punish	(8) Index
Online	-0.462***	-0.542***	-0.771***	0.128**	-0.043	-0.152***	-0.859***	-0.477***
	(0.081)	(0.084)	(0.117)	(0.065)	(0.061)	(0.048)	(0.074)	(0.060)
Non-male	0.169**	0.429***	0.104	-0.023	0.092	0.056	-0.045	0.164***
	(0.078)	(0.094)	(0.127)	(0.058)	(0.065)	(0.051)	(0.100)	(0.063)
Online x Non-male	0.128	0.225**	0.250	-0.123	0.001	0.045	0.162	0.148*
	(0.105)	(0.105)	(0.155)	(0.080)	(0.069)	(0.058)	(0.101)	(0.076)
Mean	5.86	5.72	5.03	1.23	6.84	4.70	3.15	0.00
SD	1.19	1.33	1.90	0.88	0.81	0.72	1.31	0.92
N	2400	2400	2400	2400	2400	2400	2400	2400

Notes: OLS estimates. Outcome variables are described in Section 3. Non-male includes women and non-binary respondents. Standard errors clustered at the respondent level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.6. Effect of prior harassment experience on evaluations

	(1) Serious	(2) Angry	(3) Report to Authority	(4) Victim Blame	(5) Perpetrator Blame	(6) Social Norm	(7) Punish	(8) Index
Online	-0.40***	-0.49***	-0.63***	0.13*	-0.07	-0.11***	-0.72***	-0.41***
	(0.09)	(0.09)	(0.11)	(0.07)	(0.07)	(0.04)	(0.07)	(0.06)
Prior online harassment.	-0.03	0.06	0.13	0.00	-0.12	0.01	0.13	0.03
	(0.09)	(0.10)	(0.14)	(0.06)	(0.09)	(0.06)	(0.11)	(0.07)
Prior in-person harassment	0.06	0.08	0.04	-0.09	0.14	-0.01	0.13	0.06
	(0.08)	(0.11)	(0.14)	(0.06)	(0.08)	(0.06)	(0.11)	(0.07)
Online x Prior online haras.	0.11	0.14	-0.11	-0.06	0.07	-0.01	-0.04	0.05
	(0.11)	(0.10)	(0.18)	(0.07)	(0.07)	(0.06)	(0.12)	(0.08)
Online x Prior in-person haras.	-0.06	0.04	0.05	-0.10	0.02	-0.03	-0.09	-0.01
	(0.11)	(0.11)	(0.18)	(0.08)	(0.08)	(0.06)	(0.11)	(0.08)
Mean	5.86	5.72	5.02	1.23	6.84	4.70	3.15	-0.00
SD	1.18	1.33	1.90	0.88	0.81	0.73	1.31	0.92
N	2408	2408	2408	2408	2408	2408	2408	2408

Notes: OLS estimates. Outcome variables are described in Section 3.
Standard errors clustered at the respondent level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A.7. Online vs in-person harassment by respondent characteristics

	(1) Serious	(2) Angry	(3) Report to Authority	(4) Victim Blame	(5) Perpetrator Blame	(6) Social Norm	(7) Punish	(8) Index
Online	-0.48***	-0.59***	-0.85***	0.21**	-0.04	-0.14**	-0.81***	-0.50***
	(0.12)	(0.12)	(0.16)	(0.10)	(0.09)	(0.06)	(0.10)	(0.09)
Respondent: Non-male	0.29**	0.53***	-0.05	0.02	0.10	0.13**	-0.07	0.21**
	(0.12)	(0.13)	(0.20)	(0.11)	(0.10)	(0.06)	(0.15)	(0.09)
Prior online harassment	0.04	0.11	0.12	-0.02	-0.11	0.14	0.07	0.07
	(0.14)	(0.16)	(0.22)	(0.09)	(0.18)	(0.09)	(0.16)	(0.11)
Prior in-person harassment	0.09	0.03	-0.13	-0.07	0.13	-0.03	0.23	0.05
	(0.14)	(0.17)	(0.21)	(0.09)	(0.14)	(0.11)	(0.15)	(0.11)
Online × Respondent non-male	0.22	0.32*	0.62***	-0.19	-0.10	0.10	0.27*	0.26**
	(0.17)	(0.17)	(0.22)	(0.15)	(0.12)	(0.08)	(0.15)	(0.12)
Online × Prior online har.	0.05	0.03	-0.29	-0.16	0.17	-0.11	-0.11	-0.02
	(0.15)	(0.16)	(0.27)	(0.10)	(0.12)	(0.11)	(0.17)	(0.12)
Online × Prior in-person har.	0.01	0.12	0.44*	-0.10	-0.13	0.07	-0.01	0.09
	(0.15)	(0.16)	(0.25)	(0.12)	(0.13)	(0.11)	(0.16)	(0.12)
Non-male × Prior online har.	-0.13	-0.12	-0.01	0.02	-0.02	-0.23**	0.11	-0.07
	(0.18)	(0.21)	(0.29)	(0.12)	(0.19)	(0.11)	(0.23)	(0.14)
Non-male × Prior in-person har.	-0.14	-0.12	0.28	-0.05	-0.04	0.01	-0.14	-0.05
	(0.18)	(0.21)	(0.30)	(0.13)	(0.15)	(0.12)	(0.22)	(0.14)
Online × Non-male × Prior online h.	0.10	0.15	0.27	0.19	-0.18	0.20	0.13	0.11
	(0.22)	(0.21)	(0.37)	(0.14)	(0.15)	(0.13)	(0.23)	(0.16)
Online × Non-male × Prior in-person h.	-0.22	-0.31	-0.89**	0.06	0.29*	-0.22*	-0.24	-0.28*
	(0.23)	(0.22)	(0.35)	(0.17)	(0.16)	(0.13)	(0.23)	(0.17)
Mean	5.86	5.72	5.03	1.23	6.84	4.70	3.15	0.00
SD	1.19	1.33	1.90	0.88	0.81	0.72	1.31	0.92
Observations	2400	2400	2400	2400	2400	2400	2400	2400

Notes: OLS estimates without individual fixed effects. Outcome variables are described in Section 3. Non-male includes women and non-binary respondents. Standard errors clustered at the respondent level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8. Sample characteristics and UK population benchmarks (Census 2021)

	Sample	UK population
Age		
Mean (years)	44.1	49.1
Gender (%)		
Male	48.2	49.0
Female	50.2	51.0
Non-binary / other	1.3	≈0.5
Education (%)		
Bachelor or higher	56.4	33.8
Below bachelor	43.6	66.2
Employment status (%)		
Full-time	58.8	39.9
Part-time	17.1	15.8
Other	23.8	44.3
Ethnicity (%)		
White	88.6	81.7
Asian	5.2	9.3
Black	3.2	4.0
Mixed / multiple	2.9	2.9
Other	0.2	2.1
Region (%)		
North East	4.3	4.0
North West	10.1	11.1
Yorkshire and the Humber	9.0	8.2
East Midlands	10.3	7.3
West Midlands	9.0	8.9
East of England	9.0	9.5
London	11.1	13.1
South East	15.4	13.9
South West	8.3	8.5
Wales	3.3	4.6
Scotland	7.8	8.1
Northern Ireland	2.3	2.8
N	602	—

Notes: Population benchmarks are from the 2021 Census for England and Wales (Office for National Statistics, conducted 21 March 2021), except for *Region* (ONS mid-year population estimates 2021, Nomis dataset NM_2002_1; reference date mid-2021). All tables accessed via Nomis (https://www.nomisweb.co.uk). Income is not collected in Census 2021 and is excluded from the comparison. Non-binary / other is from table TS078 (gender identity, England and Wales); gender identity is recorded separately from sex and the two are not directly comparable. The figure (« 0.5%) is approximate.

Appendix B Additional Results with Multiple Hypothesis Testing Corrections

Table B.1. Online vs in-person harassment by incident characteristics with multiple hypothesis testing corrections

	(1) Serious	(2) Angry	(3) Report to Authority	(4) Victim Blame	(5) Perpetrator Blame	(6) Social Norm	(7) Punish
Online	0.031	0.058	0.175	0.107	0.046	0.054	-0.011
	(0.792)	(0.651)	(0.296)	(0.111)	(0.501)	(0.404)	(0.898)
	[0.909]	[0.834]	[0.532]	[0.163]	[0.764]	[0.711]	[0.909]
Image-based	1.043***	1.059***	1.791***	0.128	0.046	0.334***	1.641***
	(0.000)	(0.000)	(0.000)	(0.161)	(0.493)	(0.000)	(0.000)
	[0.001]	[0.001]	[0.001]	[0.121]	[0.383]	[0.001]	[0.001]
Non-binary victim	0.368***	0.587***	0.249	0.035	0.100	0.075	0.209***
	(0.001)	(0.000)	(0.153)	(0.650)	(0.218)	(0.325)	(0.014)
	[0.003]	[0.001]	[0.184]	[0.542]	[0.248]	[0.301]	[0.006]
Woman victim	0.388***	0.786***	0.502***	-0.065	0.095	0.209***	0.318***
	(0.000)	(0.000)	(0.003)	(0.152)	(0.199)	(0.003)	(0.000)
	[0.002]	[0.001]	[0.003]	[0.105]	[0.105]	[0.003]	[0.002]
Online x Image	-0.788***	-0.819***	-1.576***	-0.268***	-0.024	-0.262***	-1.373***
	(0.000)	(0.000)	(0.000)	(0.019)	(0.791)	(0.002)	(0.000)
	[0.001]	[0.001]	[0.001]	[0.004]	[0.734]	[0.001]	[0.001]
Online x Non-binary	-0.219	-0.422***	-0.205	0.036	-0.158	-0.093	-0.137
	(0.162)	(0.011)	(0.383)	(0.760)	(0.115)	(0.351)	(0.249)
	[0.219]	[0.005]	[0.492]	[0.677]	[0.140]	[0.492]	[0.353]
Online x Woman	-0.182	-0.308**	-0.081	-0.016	-0.137	-0.210**	-0.140
	(0.186)	(0.041)	(0.721)	(0.837)	(0.174)	(0.021)	(0.236)
	[0.249]	[0.039]	[0.861]	[0.861]	[0.249]	[0.018]	[0.249]
Image x Non-binary	-0.383***	-0.577***	-0.199	-0.164	-0.035	-0.055	-0.268**
	(0.004)	(0.000)	(0.326)	(0.217)	(0.707)	(0.557)	(0.063)
	[0.001]	[0.001]	[0.429]	[0.273]	[0.667]	[0.667]	[0.042]
Image x Woman	-0.279***	-0.559***	-0.177	-0.077	-0.046	-0.215***	-0.096
	(0.027)	(0.000)	(0.381)	(0.458)	(0.594)	(0.022)	(0.509)
	[0.009]	[0.001]	[0.609]	[0.628]	[0.628]	[0.009]	[0.628]
Online x Image x Non-binary	0.446**	0.639***	0.339	0.249	0.101	0.112	0.293
	(0.033)	(0.006)	(0.264)	(0.134)	(0.470)	(0.375)	(0.135)
	[0.012]	[0.002]	[0.320]	[0.171]	[0.388]	[0.388]	[0.171]
Online x Image x Woman	0.172	0.437**	0.041	0.229*	0.037	0.194*	-0.215
	(0.351)	(0.037)	(0.884)	(0.092)	(0.785)	(0.093)	(0.267)
	[0.472]	[0.021]	[0.932]	[0.094]	[0.932]	[0.094]	[0.347]
Mean	5.86	5.72	5.02	1.23	6.84	4.70	3.15
SD	1.18	1.33	1.90	0.88	0.81	0.73	1.31
Observations	2408	2408	2408	2408	2408	2408	2408

Notes: OLS estimates without individual fixed effects. Outcome variables are described in Section 3.

Omitted categories: in-person setting, verbal harassment, male victim.

Round parentheses report conventional p-values; square brackets report Romano-Wolf stepdown adjusted p-values, clustered at the respondent level.

Stars are based on Romano-Wolf adjusted p-values. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table B.2. Online vs in-person harassment by respondent characteristics with multiple hypothesis testing corrections

	(1) Serious	(2) Angry	(3) Report to Authority	(4) Victim Blame	(5) Perpetrator Blame	(6) Social Norm	(7) Punish
Online	-0.478***	-0.594***	-0.850***	0.207**	-0.042	-0.143**	-0.815***
	(0.000)	(0.000)	(0.000)	(0.043)	(0.655)	(0.024)	(0.000)
	[0.001]	[0.001]	[0.001]	[0.021]	[0.612]	[0.012]	[0.001]
Respondent: Non-male	0.288**	0.533***	-0.054	0.024	0.104	0.135**	-0.066
	(0.014)	(0.000)	(0.782)	(0.824)	(0.278)	(0.035)	(0.650)
	[0.012]	[0.001]	[0.916]	[0.916]	[0.471]	[0.034]	[0.887]
Prior online harassment	0.037	0.112	0.122	-0.018	-0.106	0.144	0.072
	(0.796)	(0.496)	(0.586)	(0.836)	(0.547)	(0.127)	(0.653)
	[0.924]	[0.838]	[0.891]	[0.924]	[0.891]	[0.298]	[0.891]
Prior in-person harassment	0.089	0.031	-0.125	-0.068	0.135	-0.033	0.226
	(0.515)	(0.851)	(0.556)	(0.448)	(0.338)	(0.763)	(0.137)
	[0.773]	[0.899]	[0.773]	[0.773]	[0.667]	[0.899]	[0.309]
Online x Resp non-male	0.219	0.321*	0.620***	-0.189	-0.099	0.104	0.267*
	(0.198)	(0.054)	(0.005)	(0.203)	(0.425)	(0.178)	(0.076)
	[0.245]	[0.070]	[0.005]	[0.245]	[0.308]	[0.245]	[0.096]
Online x Prior online har.	0.048	0.035	-0.285	-0.157	0.173	-0.115	-0.112
	(0.753)	(0.831)	(0.295)	(0.105)	(0.165)	(0.310)	(0.517)
	[0.859]	[0.859]	[0.495]	[0.195]	[0.298]	[0.495]	[0.727]
Online x Prior in-person har.	0.012	0.121	0.442	-0.100	-0.130	0.066	-0.011
	(0.935)	(0.452)	(0.082)	(0.394)	(0.317)	(0.562)	(0.944)
	[0.990]	[0.698]	[0.156]	[0.698]	[0.651]	[0.793]	[0.990]
Non-male x Prior online har.	-0.129	-0.120	-0.013	0.017	-0.021	-0.232*	0.111
	(0.465)	(0.569)	(0.965)	(0.882)	(0.910)	(0.041)	(0.626)
	[0.833]	[0.933]	[0.997]	[0.997]	[0.997]	[0.075]	[0.933]
Non-male x Prior online har.	-0.144	-0.118	0.281	-0.053	-0.038	0.011	-0.141
	(0.412)	(0.575)	(0.345)	(0.684)	(0.804)	(0.924)	(0.528)
	[0.806]	[0.902]	[0.708]	[0.921]	[0.949]	[0.949]	[0.902]
Online x Non-male x Prior online h.	0.103	0.150	0.272	0.189	-0.184	0.196	0.125
	(0.638)	(0.484)	(0.456)	(0.163)	(0.206)	(0.140)	(0.592)
	[0.709]	[0.697]	[0.697]	[0.294]	[0.344]	[0.271]	[0.709]
Online x Non-male x Prior in-person h.	-0.223	-0.311	-0.885**	0.064	0.293	-0.219	-0.236
	(0.331)	(0.159)	(0.013)	(0.709)	(0.074)	(0.091)	(0.306)
	[0.404]	[0.191]	[0.023]	[0.632]	[0.112]	[0.112]	[0.404]
Mean	5.86	5.72	5.03	1.23	6.84	4.70	3.15
SD	1.19	1.33	1.90	0.88	0.81	0.72	1.31
Observations	2400	2400	2400	2400	2400	2400	2400

Notes: OLS estimates without individual fixed effects. Outcome variables are described in Section 3. Non-male includes women and non-binary respondents. Round parentheses report conventional p-values; square brackets report Romano-Wolf stepdown adjusted p-values, clustered at the respondent level. Stars are based on Romano-Wolf adjusted p-values. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix C Additional Results with Fixed Effects

C.1 Heterogeneity by Incident Characteristics

Table C.1. Effect of online vs in-person harassment on evaluations

	(1) Serious	(2) Angry	(3) Report to Authority	(4) Victim Blame	(5) Perpetrator Blame	(6) Social Norm	(7) Punish
Online	-0.394***	-0.434***	-0.668***	-0.011	-0.036*	-0.131***	-0.790***
	(0.038)	(0.039)	(0.062)	(0.032)	(0.018)	(0.026)	(0.046)
Mean	5.86	5.72	5.02	1.23	6.84	4.70	3.15
SD	1.18	1.33	1.90	0.88	0.81	0.73	1.31
N	2408	2408	2408	2408	2408	2408	2408

Notes: OLS estimates with individual fixed effects. Outcome variables are described in Section 3.
Standard errors clustered at the respondent level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table C.2. Effect of image-based vs verbal harassment on evaluations

	(1) Serious	(2) Angry	(3) Report to Authority	(4) Victim Blame	(5) Perpetrator Blame	(6) Social Norm	(7) Punish
Online	-0.089**	-0.234***	-0.002	0.036	-0.064***	-0.038	-0.106**
	(0.045)	(0.049)	(0.067)	(0.030)	(0.024)	(0.036)	(0.045)
Image-based	0.810***	0.601***	1.620***	0.063	-0.031	0.236***	1.480***
	(0.051)	(0.057)	(0.090)	(0.049)	(0.032)	(0.035)	(0.067)
Online x Image	-0.589***	-0.383***	-1.292***	-0.093	0.057	-0.180***	-1.332***
	(0.063)	(0.069)	(0.104)	(0.060)	(0.038)	(0.051)	(0.078)
Mean	5.86	5.72	5.02	1.23	6.84	4.70	3.15
SD	1.18	1.33	1.90	0.88	0.81	0.73	1.31
N	2408	2408	2408	2408	2408	2408	2408

Notes: OLS estimates with individual fixed effects. Outcome variables are described in Section 3.
Omitted categories: in-person setting, verbal harassment.
Standard errors clustered at the respondent level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table C.3. Effect of victim gender on evaluations

	(1) Serious	(2) Angry	(3) Report to Authority	(4) Victim Blame	(5) Perpetrator Blame	(6) Social Norm	(7) Punish
Online	-0.364***	-0.418***	-0.620***	-0.041	-0.003	-0.101***	-0.739***
	(0.060)	(0.066)	(0.097)	(0.051)	(0.029)	(0.038)	(0.072)
Non-binary victim	0.211***	0.321***	0.267**	0.019	0.032	0.066	0.038
	(0.064)	(0.063)	(0.112)	(0.056)	(0.030)	(0.042)	(0.079)
Woman victim	0.306***	0.514***	0.513***	-0.085	0.057**	0.102**	0.236***
	(0.066)	(0.067)	(0.111)	(0.052)	(0.029)	(0.042)	(0.082)
Online x Non-binary	0.017	-0.020	-0.005	0.066	-0.075	-0.017	0.070
	(0.078)	(0.081)	(0.133)	(0.064)	(0.047)	(0.056)	(0.092)
Online x Woman	-0.105	-0.028	-0.139	0.023	-0.024	-0.071	-0.221**
	(0.085)	(0.088)	(0.136)	(0.071)	(0.040)	(0.064)	(0.096)
Mean	5.86	5.72	5.02	1.23	6.84	4.70	3.15
SD	1.18	1.33	1.90	0.88	0.81	0.73	1.31
N	2408	2408	2408	2408	2408	2408	2408

Notes: OLS estimates with individual fixed effects. Outcome variables are described in Section 3.

Omitted categories: in-person setting, male victim.

Standard errors clustered at the respondent level in parentheses. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table C.4. Online vs in-person harassment by incident characteristics with participant fixed effects

	(1) Serious	(2) Angry	(3) Report to Authority	(4) Victim Blame	(5) Perpetrator Blame	(6) Social Norm	(7) Punish
Online	0.023	-0.062	0.120	0.043	-0.044	0.027	0.054
	(0.073)	(0.075)	(0.108)	(0.056)	(0.043)	(0.056)	(0.063)
Image-based	1.008***	0.942***	1.828***	0.134	-0.046	0.350***	1.644***
	(0.076)	(0.085)	(0.129)	(0.094)	(0.060)	(0.059)	(0.096)
Non-binary victim	0.381***	0.593***	0.407***	0.066	0.037	0.146**	0.209***
	(0.084)	(0.077)	(0.130)	(0.059)	(0.039)	(0.065)	(0.071)
Woman victim	0.443***	0.763***	0.708***	-0.025	0.027	0.197***	0.336***
	(0.080)	(0.088)	(0.128)	(0.047)	(0.036)	(0.057)	(0.070)
Online x Image	-0.732***	-0.681***	-1.392***	-0.165	0.081	-0.243***	-1.513***
	(0.100)	(0.115)	(0.164)	(0.114)	(0.069)	(0.074)	(0.116)
Online x Non-binary	-0.166	-0.305***	-0.166	0.037	-0.063	-0.085	-0.245***
	(0.110)	(0.108)	(0.165)	(0.086)	(0.075)	(0.082)	(0.092)
Online x Woman	-0.166	-0.204*	-0.194	-0.060	0.004	-0.109	-0.234**
	(0.101)	(0.109)	(0.156)	(0.084)	(0.050)	(0.082)	(0.092)
Image x Non-binary	-0.311***	-0.521***	-0.224	-0.094	-0.011	-0.151**	-0.292**
	(0.105)	(0.103)	(0.172)	(0.118)	(0.062)	(0.076)	(0.114)
Image x Woman	-0.286***	-0.510***	-0.405**	-0.122	0.060	-0.193**	-0.206*
	(0.105)	(0.116)	(0.181)	(0.104)	(0.059)	(0.083)	(0.124)
Online x Image x Non-bin.	0.290**	0.506***	0.179	0.055	-0.024	0.111	0.507***
	(0.142)	(0.147)	(0.228)	(0.135)	(0.095)	(0.100)	(0.149)
Online x Image x Woman	0.124	0.362**	0.101	0.170	-0.054	0.078	0.020
	(0.130)	(0.147)	(0.218)	(0.141)	(0.070)	(0.114)	(0.154)
Mean	5.86	5.72	5.02	1.23	6.84	4.70	3.15
SD	1.18	1.33	1.90	0.88	0.81	0.73	1.31
Observations	2408	2408	2408	2408	2408	2408	2408

Notes: OLS estimates with individual fixed effects. Outcome variables are described in Section 3.

Omitted categories: in-person setting, verbal harassment, male victim.

Standard errors clustered at the respondent level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix D Experimental Instructions and Design

This appendix reproduces the instructions shown to participants and describes the experimental design in detail. The experiment was implemented as an online survey using Qualtrics. Participants were recruited through Prolific. To be eligible, participants had to be 18 or above and reside in UK.

Participants evaluated a series of short scenarios describing interpersonal interactions and answered questions after each scenario. Participants who did not consent or failed quality checks were excluded. All wording below reflects the text shown to participants, with formatting adjustments for readability.[5]

All payments were made through Prolific. Participants received a fixed participation payment. In addition, participants could receive:

- A £1 bonus for correctly predicting the most common answer in one randomly selected social appropriateness question.
- Entry into a lottery to win £100. Participants were informed that one participant would be randomly selected to receive the £100 prize. If selected, participants could choose to donate any portion of the prize to a UK charity supporting victims of harassment.

D.1 Consent and participation

Participants first saw the following information screen:

> **Welcome and Thank You for Participating!**
>
> This is a survey conducted for academic research purposes by researchers from the University of Southampton. We expect the survey to take about 10 minutes. Your responses are **anonymous** and will be used for **research purposes only**. We will **not collect any identifying personal information.**
>
> You will receive compensation **if you complete the survey and pass our standard quality checks**, which help us detect rushed or inconsistent responses. You will also be **entered into a lottery to win £100.** It

[5]The full online version is available at http://www.studypreview.com

is very important for the validity of our research that you answer honestly and read the questions carefully before answering.

What is this survey about?

We are studying people's perceptions and attitudes toward different kinds of social interactions in various settings, including online and offline environments. You will be presented with brief descriptions of interpersonal situations. We will ask you to share your honest opinions and feelings about these scenarios. At the end of the survey, you may also be invited to support a social cause.

If you have any questions or concerns about this survey or its outcomes, please contact Martina Vecchi at m.vecchi@soton.ac.uk.

This study has been reviewed and approved by the Ethics and Research Governance body of the University of Southampton. If you wish to make a formal complaint, you can contact the Head of Research Integrity and Governance at: Email: rgoinfo@soton.ac.uk, phone: +44 2380 595058. Please quote the study number 107127.

By clicking the button below and taking the survey, you agree to participate in the study.

Thank you again for your participation!

- I consent, and I am willing to participate in the study.
- I do not consent, and I am not willing to participate in the study.

Participants had to select a consent option to proceed.
Participants were then asked to confirm the following:

Important Note:

Because this research addresses an important topic, your responses matter! To ensure we collect quality data, we ask that you **read all the information provided carefully** and give your **best and truthful answer** to each question in the survey.

Do you commit to carefully reading and providing thoughtful and accurate answers to the questions in this survey?

- I will read carefully and provide my best answers.
- I will not promise to read carefully and provide my best answers.

Participants also completed a CAPTCHA and entered their Prolific ID.

D.2 General instructions

Participants were shown the following instructions:

In what follows, you will be asked to consider **five different situations involving interpersonal behaviour. The short description at the start of each situation will change**, but the same set of questions will be asked after each one. We will ask you to share your **honest opinions and feelings** about these scenarios. Your responses are completely confidential, and there are no right or wrong answers: **we want to hear your genuine perspective!**

Among the questions, we will also ask you to evaluate how **socially appropriate** certain behaviours are in those situations. In each case, you must indicate whether the behaviour would be "socially appropriate" or "socially inappropriate." There will be five possible responses, as shown below, from which you must select exactly one:

- Very socially appropriate
- Somewhat socially appropriate
- Neither inappropriate nor appropriate
- Somewhat socially inappropriate
- Very socially inappropriate

By socially appropriate, we mean behaviour that you think most people would agree is the "right" thing to do. Another way to

think about this is that, if someone behaved in a socially inappropriate way, other people might become upset or angry. Note that what is considered socially appropriate is not necessarily made explicit, supported by laws, or enforced through legal sanctions. Therefore, an action may be "appropriate" even if it is not legal, or "inappropriate" even if it is not illegal.

After all responses to the survey are collected, we will randomly select one of the situations we asked you about. We will look at your answer to the question asking how socially appropriate the behaviour in that situation was. To reward you, **if your answer to this question is the same as the answer provided by the highest number of participants in this survey, and if you are selected as eligible for bonus payment, you will receive a £1 bonus.** All participants in this survey are from the UK and are recruited online.

Please note: **the bonus is based only on your answer to the appropriateness question.**

D.3 Experimental task

Participants evaluated five situations in total:

- one filler situation unrelated to harassment,
- four situations describing sexual harassment.

Each harassment vignette was drawn from a pool of possible scenarios varying along three dimensions:

- setting (online / in person)
- gender of the target (man / woman / non-binary)
- type of harassment (verbal / image-based)

Each participant saw four different combinations drawn without replacement. The order of situations was randomized.

In addition, all participants first saw the following situation (filler vignette), which was included to reduce demand effects:

> A 25-year-old man is standing in line at the supermarket checkout when a stranger walks past him and cuts to the front of the queue without saying anything.

D.4 List of harassment vignettes

All harassment vignettes were shown in a large bold font at the top of the screen and followed the template:

> A 25-year-old [gender] receives [type of behaviour] in a [setting].

The full set of possible vignettes is shown in Table D.1.

Table D.1. Vignettes combination

Gender	Setting	Type	Text
Man	Online	Image	receives unsolicited explicit sexual images in a direct message
Man	Online	Verbal	receives unsolicited explicit sexual comments in a direct message
Man	In person	Image	has a stranger expose themselves in the street
Man	In person	Verbal	receives explicit sexual comments from a stranger in the street
Woman	Online	Image	receives unsolicited explicit sexual images in a direct message
Woman	Online	Verbal	receives unsolicited explicit sexual comments in a direct message
Woman	In person	Image	has a stranger expose themselves in the street
Woman	In person	Verbal	receives explicit sexual comments from a stranger in the street
Non-binary	Online	Image	receives unsolicited explicit sexual images in a direct message
Non-binary	Online	Verbal	receives unsolicited explicit sexual comments in a direct message
Non-binary	In person	Image	has a stranger expose themselves in the street
Non-binary	In person	Verbal	receives explicit sexual comments from a stranger in the street

As an attention and quality check, after the first filler vignette, participants were asked: "In your honest opinion, should we keep your responses, or discard them because you were not fully attentive?"

At the end of the five vignettes, they were then asked: Please choose "yellow" from the options below. If you do not select yellow in this question, you will be forced to leave the survey.

- Blue
- Yellow
- Red
- I do not have a preferred color

D.5 Questions after each vignette

After each situation participants answered the following questions.

- **Seriousness**: How serious or not serious do you think the situation described is? (1-7 scale: Not serious at all; Moderately not serious; Slightly not serious; Neither serious nor not serious; Slightly serious; Moderately serious; Very serious.)
- **Reporting**: If you were aware this had happened, how likely would you be to report it? (7-point scale: Extremely unlikely; Moderately unlikely; Slightly unlikely; Neither likely nor unlikely; Slightly likely; Moderately likely; Extremely likely).
- **Anger**: How angry does this situation make you feel? (7-point scale: Not angry at all; Moderately not angry; Slightly not angry; Neither angry nor not angry; Slightly angry; Moderately angry; Very angry.)
- **Social appropriateness**: Would most people consider the behaviour in this situation to be socially appropriate or inappropriate? By socially appropriate, we mean behaviour that most people would consider the right thing to do.

You may receive a £1 bonus if your answer matches the most common answer given by other participants.(5-point scale: Very socially appropriate; Somewhat socially appropriate; Neither appropriate nor inappropriate; Somewhat socially inappropriate; Very socially inappropriate.)

- **Punishment**: In your opinion, what would be the most appropriate response to this kind of behaviour? Please select the one option you think is most appropriate.(No punishment; Verbal warning from someone in a position of authority; Formal police warning; Restorative justice meeting between the victim and the offender (facilitated by trained professionals); Community service; Prison sentence).

- **Responsibility (target)**: To what extent is the person targeted responsible for what happened? (7-point scale: Not at all responsible; Mostly not responsible; Slightly not responsible; Neither responsible nor not responsible; Slightly responsible; Mostly responsible; Definitely responsible).

- **Responsibility (perpetrator)**: To what extent is the person who carried out the behaviour responsible for what happened? (7-point scale: Not at all responsible; Mostly not responsible; Slightly not responsible; Neither responsible nor not responsible; Slightly responsible; Mostly responsible; Definitely responsible).

D.6 Donation decision

Participants entered a lottery to win £100. They were asked:

> By taking this survey, **you are automatically entered into a lottery to win £100**. In a few days, you will know whether you have been selected as a winner. The payment will be made to you in the same way as your compensation for this survey, so no further action is required on your part.
>
> If you win, **you can donate a portion of your £100 prize to Victim Support**, an independent charity dedicated to supporting victims of crime

and traumatic incidents in England and Wales. Any donations from this study will go specifically toward their work to **support victims of online harassment and ending online abuse.**

Should you win the lottery, how much of your prize would you like to donate to Victim Support?

Enter your donation amount (£) using the slider below.

£0 _____ £100

D.7 Background questionnaire

Participants reported:

- Age
- Gender
- Region
- Education
- Employment
- Income
- Household composition
- Ethnicity
- Political orientation
- Experience with harassment
- Internet use
- Social media use

We then asked participants what they thought the survey was about and invited them to provide feedback on the clarity and flow of the survey.

> The survey is nearing completion. You can now enter any comments, thoughts, or suggestions in the field below.
>
> How clear were the survey instructions?
>
> - Very clear
> - Mostly clear
> - Somewhat unclear
> - Very unclear
>
> What do you think this study was mainly about? Please select the option that best matches your impression.
>
> - How people decide whether to support social causes
> - People's opinions about different kinds of social interactions in everyday life
> - Judgments about fairness and responsibility in social situations
> - Views about crime and punishment
> - Attitudes toward online and offline harassment
> - General personality and attitudes
> - Something else. Please specify:
>
> Do you have any suggestions to improve the clarity or flow of the survey?

D.8 Petition task

Participants were shown the following screen:

Finally, we ask whether you are willing to **support a petition to "stand up against online sexual harassment."**

As soon as the survey is complete, **we will send the results to the UK Home Office,** informing them of the share of people who took this survey who were willing to support the following petition:

> "I agree that stronger protections against online sexual harassment are urgently needed. Now is the time to ensure that harmful online behaviour, including unwanted sexual messages and explicit images, is recognised and addressed as a serious offence. Digital spaces should be safe for everyone, regardless of gender or age. I am adding my voice to the call for our leaders to take meaningful legal action to criminalise online sexual harassment and protect people from abuse in digital environments."

Please note: this petition is not affiliated with any political party or organisation.

Do you support this petition? (you will not be asked to sign it, only your answer here is required and remains anonymous)?

- Yes
- No

D.9 Randomization

The experiment involved the following randomizations:

- harassment vignettes drawn from the full set without replacement
- vignette order randomized
- order of questions randomized
- lottery winner selected randomly