

IPOC DATABASE: A FREE DATABASE OF INTERRACIAL POLICE OFFICER-CIVILIAN INTERACTION VIDEOS

Tzipporah P. Dang, Jasmin Cloutier, and Jennifer T. Kubota
University of Delaware

Researchers investigating reactions to interracial police officer-civilian interactions are increasingly interested in utilizing more naturalistic stimuli. To facilitate this research, we introduce the IPOC (interracial police officer-civilian) Database, a freely available set of real-world videos of interracial officer-civilian interactions with norming data. These short, audio-less videos represent a range of interactions focused on the arrest (e.g., traffic stops, foot chases). To create this database, participants evaluated a set of 78 videos on the social category membership (e.g., age, race) of officers and civilians and aspects of the interaction (e.g., aggressiveness) and scene (e.g., neighborhood safety). Videos were then blurred in order to measure officer and civilian aggressiveness in the absence of race cues. Then, 38 videos were equated across multiple dimensions based on civilian race (19 Black, 19 White). The database includes all 78 videos, the subset of 38 equated videos, and ratings for open access use in future research.

Keywords: police, stereotyping and prejudice, arrests, interracial interactions

With the introduction of new media and calls for transparency, the public is increasingly exposed to police officer body-worn camera (BWC), surveillance, and bystander video of police officer-civilian interactions. Even with video footage, the court of public opinion is at times divided in response to these events. But what are the factors that shape divergent perceptions and evaluations of critical police officer-civilian interactions? When presented with the same video, the public, media outlets, and prosecutors interpret the culpability and legitimacy of both the police officer's and civilian's actions in different ways. For example,

We would like to thank Denise Barth, Shawn Cho, Estelle Higgins, and Audrey Shinn for finding the videos, Aisha Suhiaba for data collection, and Jesse Randall for blurring the videos.

Correspondence concerning this article should be addressed to Jennifer T. Kubota, Department of Psychological and Brain Sciences, University of Delaware, 105 The Green, Newark, DE 19716. E-mail: jennifer.kubota@gmail.com.

© 2020 Guilford Publications, Inc.

prosecution used police dash cam footage to contradict the statements of the three police officers acquitted of covering up the killing of Laquan McDonald, a Black 17-year-old, by a White police officer. The judge said that “it is undisputed and undeniable McDonald was an armed assailant” (Lockhart, 2019), while activists say the video shows “McDonald was not attacking or seeking to attack any of the law enforcement officers” (Bosman & Davey, 2019). To better understand what factors shape these divergent views, over the past 20 years there has been a rapid increase in the number of studies investigating interracial police officer-civilian interactions across disciplines (Culhane, Boman, & Schweitzer, 2016; Jennings, Fridell, & Lynch, 2014) as well as calls to understand the potential biases and benefits of using video evidence in courtrooms (Granot, Balcetis, Feigenson, & Tyler, 2018; Lassiter, Diamond, Schmidt, & Elek, 2007). Because it is often difficult for researchers to obtain these types of stimuli, we introduce the IPOC (interracial police officer-civilian) Database, a freely available set of naturalistic, interracial police officer-civilian arrest interactions.

To create this database, a set of 78 short videos were curated and edited to focus on the moment of the arrest and do not provide information regarding the suspected crime. These videos, recorded from a third-person perspective, depict a White, male police officer arresting either a Black or White, male civilian. The videos have no audio and represent a range of perceived aggressiveness. Participants rated the 78 curated videos on a variety of dimensions. Based on these ratings, we equated a subset of 38 videos across a number of dimensions based on civilian race (19 Black, 19 White) for open access use in future research. The resulting stimulus set minimizes the influence of extraneous variables (e.g., time of day, police officer age, etc.) that could potentially impact the perceptions and evaluations of the interactions. We provide instructions for obtaining the database and ratings, which includes all 78 videos as well as the matched 38 videos of Black and White civilians.

AIM 1: GATHER VIDEOS AND RATINGS

The goal of Aim 1 was to collect videos, edit them for presentation, and collect ratings about the physical properties of the videos and participant assessments (e.g., civilian and police officer aggressiveness) that could be used in Aim 2 to equate a subset of Black and White civilian videos.

STIMULI COLLECTION

One hundred and four videos of male police officer and male civilian interactions were collected from the Cops TV series from 1991 (season 4) to 2017 (season 29). Cops is a documentary-based reality TV show that follows police officers as they conduct patrols, respond to calls, make arrests, and so forth. These videos are currently available for public viewing in the original form on the TV show Cops. Individuals on the show gave their permission to be in the videos. These initial videos were a convenience sample after excluding videos that did not

meet basic criteria. We initially selected videos depicting an interaction between police officers that appeared male and White, civilians that appeared male and either White or Black, and ideally had one or a lead police officer and one or a focal civilian. In addition, the interaction had to be in the foreground. Any videos that did not meet these initial criteria as determined by the researcher were immediately excluded. Then, we shortened the videos to focus on the arrest or a central part of the clip (e.g., a police officer pulling a civilian out of a car window) and removed the audio. We cropped the perimeter of the video if there were any embedded logos using iMovie.

Following experimenter selection and editing, we assessed video duration to the second using QuickTime Player (see Supplemental Material Table S1) and obtained subjective ratings of participants' evaluations of the interactions and properties of the videos.

METHOD FOR COLLECTING VIDEO RATINGS

After curating the videos, we collected ratings of variables that may differ between the videos or impact how perceivers evaluate the videos to allow researchers to select those that best fit their research questions. To do so, three stages of participants (total n across the stages = 501) rated the videos on various dimensions (see below; see Supplemental Material for inclusion criteria). Ratings were acquired in three stages for two reasons. First, because we wanted to minimize participant fatigue while rating the videos, participants only made a subset of all possible ratings (i.e., Stages 1 and 2). Second, we wanted to collect ratings of aggressiveness of the interaction after we obscured civilian race with new participants who had not previously seen the videos (Stage 3, see below for more information; see Figure 1 for examples).

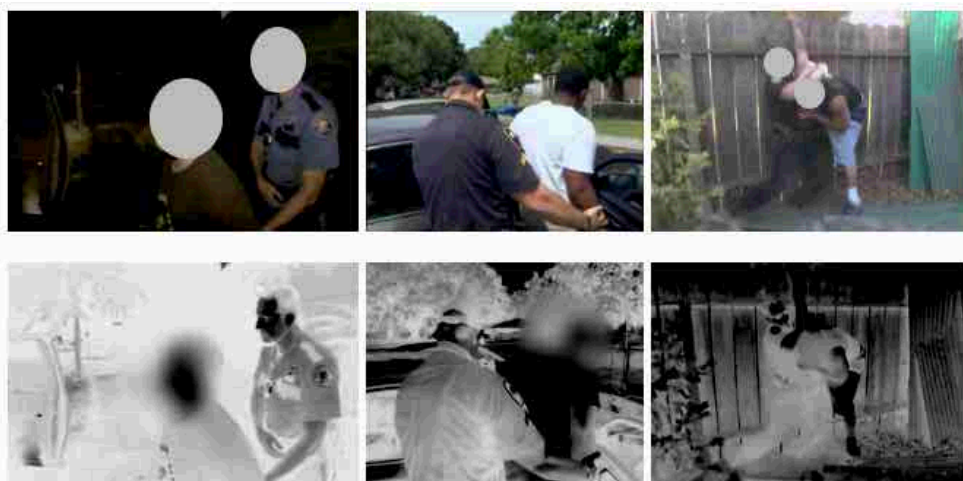


FIGURE 1. Examples of unblurred (top row) and blurred videos (bottom row). Note that for publication, we have put grey dots over the faces of the people in the unblurred videos. These grey dots do not appear in the actual videos.

For all stages, after indicating their consent, participants were introduced to the experiment. They were told they would view short, audio-less videos of police officer-civilian interactions and answer several questions about the videos. All videos were presented using Qualtrics via embedded videos hosted on YouTube. The videos were presented in a random order. Participants were informed that there would be some attention check trials that included explicit instructions on how to respond to make sure they were paying attention. Finally, they were introduced to the ratings.

During the presentation of the videos, participants saw attention check trials that were randomly intermixed with the videos. For attention check trials, participants were shown a police officer-civilian video of non-interest (e.g., the interaction was too far from the camera) and were explicitly told to select a particular answer (e.g., "CATCH TRIAL: SELECT ANSWER #5 FOR THE QUESTION BELOW." See Supplemental Material Figure S1). Following video rating, the experiment completed and participants from the University of Chicago and Amazon Mechanical Turk (MTurk) were compensated with course credit or money, respectively. All participants consented online in accordance with the University of Chicago Institutional Review Board (IRB16-1673) and the Declaration of Helsinki. The University of Chicago IRB and Office of Legal Counsel both determined that the videos could be used for research purposes only. Participants from the University of Chicago only provided demographic ratings (i.e., police officer and civilian race and age) and the first video quality rating. To increase our sample size, we also included MTurk participants who provided the remainder of the ratings. MTurk participants were restricted to those located in the United States who had an Human Intelligence Task (HIT) approval rating greater than or equal to 70%.

VIDEO RATINGS

Video ratings were collected across three stages, each containing new groups of participants (see Table 1). Recall that ratings were collected in stages to (a) not fatigue participants during data collection and (b) measure aggressiveness of the police officers and civilians when participants were not aware of the race of the civilian (see below). Therefore, separate groups of participants rated videos on no more than three dimensions of interest (i.e., degrees of freedom will vary across the dimensions see Table 1).

Our goal was to collect enough participants to allow for variability in the ratings of the videos. For Stage 1, we allowed for a smaller participant sample for demographic ratings because we had a priori constrained the variance in race by selecting our sample of videos based on experimenter ratings of these characteristics. Following Stage 1 of data collection, we collected ratings from larger samples to allow us to capture more inter-individual variance. Importantly for Aim 2, the unit of analysis for equating the videos across civilian race was number of videos (38) and not number of participants who rated each video. For equating, we are powered above 80% to detect differences for all rating dimensions. Additionally,

TABLE 1. Summary of Dimensions Rated in Each Stage and the Number of Participants That Rated Each Dimension. Participants Rated the Same 78 Videos in Each Stage.

Stage	Group	Dimensions	Participant <i>n</i>		
1	1	PO age	34		
		Civilian age			
		PO race			
		Civilian race			
2	2	Aggression of the interaction	90		
		Arousal to the interaction			
		Pleasantness of the neighborhood			
3	3	Safety of the neighborhood	88		
		Time of day			
		Video quality			
		PO aggression [^]			
3	5	Civilian aggression [^]	99		
		6		Blurred civilian race [^]	89
				501	

Note. PO = police officer; [^]Ratings given on blurred videos. Group refers to the manner in which data were collected. Different participants were included in each group to reduce rating fatigue during data collection.

if researchers would like to use participant as the unit of analysis in the future, the participant sample sizes are powered above 80%.

STAGE 1

The goal of Stage 1 ($n = 34$) was to collect participant ratings of police officer and civilian race and age. Stage 1 assessed civilian race to ensure our equated subset of stimuli had an equal distribution of Black and White civilians. Because of the lack of representation of Black police officers in these videos, it was only possible to collect White police officers. We also collected age to equate the police officers and civilians across racial groups. Race was assessed by asking, "What is the police officer's [civilian's] race?" and ratings were categorical (i.e., 1 = Black/African American, 2 = Asian/Asian American, 3 = White/Euro-American, 4 = Latino/Hispanic American, 5 = Middle Eastern/Arab American, 6 = Native American, and 7 = Other). Age was assessed by asking, "What is the police officer's [civilian's] age (in years)?" and the ratings were categorical (i.e., 1 = 18–25, 2 = 26–35, 3 = 36–45, 4 = 46–55, 5 = 56–65, 6 = 66–75, and 7 = 76+ years).

In Stage 1, 34 participants rated all 104 randomly presented videos on several demographic dimensions in a fixed rating order of police officer race, police officer age, civilian race, and civilian age (see Supplemental Material for procedural details and exclusions). From these ratings, we eliminated 26 videos that did not

have at least an 80% consensus on race of the civilians as either White or Black and/or the police officer as White. This resulted in a total of 78 videos (45 videos of White civilians) for further data collection in Stage 2.

STAGE 2

The goal of Stage 2 (total $n = 279$) was to collect new participant ratings of the social interactions and scene quality (i.e., aggression of the interaction, arousal to the interaction, pleasantness of the neighborhood, safety of the neighborhood, time of day, and video quality). Aggressiveness of the interaction was assessed by asking, "Rate the overall aggressiveness of the interaction," and the ratings were provided on a 7-point scale with 1 as "Not at all aggressive" to 7 as "Extremely aggressive." Arousal was assessed by asking, "To you, how arousing is this interaction?" and the ratings were provided on a 7-point scale with 1 as "Not at all arousing" to 7 as "Extremely arousing." Pleasantness was assessed by asking, "Independent of the arrest, how pleasant is the neighborhood?" and the ratings were provided on a 7-point scale with 1 as "Very unpleasant" to 7 as "Very pleasant." Safety was assessed by asking, "Independent of the arrest, how safe is the neighborhood?" and the ratings were provided on a 7-point scale with 1 as "Very unsafe" to 7 as "Very safe." Video quality was assessed by asking, "Rate the video picture quality," and the ratings were provided on a 7-point scale with 1 as "Very poor quality" to 7 as "Very high quality." Time of day was assessed by asking, "At what time of day did this interaction take place?" and the ratings were categorical (1 = Morning, 2 = Early afternoon, 3 = Late afternoon, and 4 = Evening/night).

To reduce participant fatigue during the rating task, data were collected in three groups during Stage 2, with each group rating the videos on different dimensions (see Table 1). Each group gave ratings for all 78 videos remaining after Stage 1 exclusions. Group 2 consisted of 90 MTurk participants who rated the videos on overall aggressiveness of the interaction and arousal in a counterbalanced order (see Supplemental Material for exclusion criteria). Group 3 consisted of 88 MTurk participants who rated the videos on pleasantness, safety, and time of day in a counterbalanced order. Group 4 consisted of 101 MTurk participants who rated the videos on video quality.

STAGE 3

The goal of Stage 3 (total $n = 188$) was to blur the race of the civilian and have new participants rate the police officer and civilian aggression in the absence of race cues. Doing so provides an independent measure of aggression that is not contaminated by stereotypes. This is important because research has shown that regardless of an individual's explicit prejudice, people are aware of the U.S. stereotypes of Black people as criminal and hostile (Devine & Elliot, 1995; Kubota, Banaji, & Phelps, 2012). Additionally, individuals implicitly associate Black people with criminality and guns (Correll, Park, Judd, & Wittenbrink, 2002; Payne, 2001). Therefore, for equating purposes in Aim 2, we wanted to minimize the influence

of racial stereotypes on aggressiveness ratings of these police officer-civilian interactions.

Blurring Process. The goal of blurring was to hide the racial profile of the civilian. We selected a process for blurring race that could be effectively applied across the 78 videos (video duration ranges from 4 to 31 seconds). A contracted professional using Adobe After Effects CC 2014 implemented all blurring procedures.

The first process of blurring was to create a combination of color correction filters to disguise skin tone. First, the footage was desaturated 100%. Second, the contrast was filtered from the footage and then inverted. This was done using a levels adjustment layer that adjusts videos of Black civilians' input to 31, videos of White civilians' input to 206, and videos of White civilians' output to 149. The final stage was to bring back the contrast to make the footage readable again. This was done using a second levels adjustment layer that alters videos of Black civilians' input to 91, videos of White civilians' input to 136, and the gamma to 1.17.

The second process was to blur the civilians' faces to hide identifying features. To do this, a masking selection was made around the civilian's head. The selection then had to be animated to follow the civilian's head throughout each video. This selection was then used to apply the blurring effect for a range of 30 to 100 pixels. The videos were then rendered out with these effects applied for the final result. The H.264 video codec was used for best compression quality and packaged in a compressed folder for delivery (see bottom panel of Figure 1 for examples).

Following blurring, new participants rated the blurred videos on aggressiveness of the police officers and civilians. Aggressiveness of individuals was assessed by asking, "How aggressive is the police officer [civilian]?" and the ratings were provided on a 7-point scale with 1 as "Not at all aggressive" to 7 as "Extremely aggressive." Data were collected with two new groups, with each group rating the videos on different dimensions (see Table 1). Group 5 consisted of 99 MTurk participants who rated the videos on police officer aggressiveness and civilian aggressiveness in a counterbalanced order (see Supplemental Material for participant exclusion criteria). Group 6 consisted of 89 MTurk participants who rated the videos on civilian race. The combination of filters and effects made it difficult to distinguish the civilian's race. Participants were below 50% accuracy in identifying civilian race from the blurred videos (accuracy = 39%). See Table 2 for correlations among all ratings collected across Stages 1, 2, and 3 for the 78 videos and Supplemental Material Table S2 for descriptive statistics of the 78 videos by civilian race.

AIM 2: VIDEO EQUATING AS A FUNCTION OF CIVILIAN RACE

We then wanted to take a subset of the 78 videos and match them on mean ratings by civilian race (Black, White) in order to have two groups of videos that only differed in civilian race. In other words, we selected a group of Black civilian videos that were not rated significantly different (on all dimensions collected for Aim 1) from a group of White civilian videos. We equated a sample of 38 videos: 19 Black

TABLE 2. Correlations of Video ($n = 78$) Ratings (Aim 1)

Dimension	Mean	SD	1	2	3	4	5	6	7	8	9	10	11
Duration (1)	15.410	6.102											
Interaction Aggression (2)	2.908	1.451	0.187										
Civilian Aggression [^] (3)	1.935	1.036	0.174	0.890**									
PO Aggression [^] (4)	2.672	1.554	0.171	0.988**	0.866**								
Arousing (5)	2.582	1.011	0.217	0.993**	0.879**	0.976**							
Pleasant (6)	3.525	0.663	-0.088	-0.487**	-0.455**	-0.483**	-0.484**						
Safe (7)	3.613	0.608	-0.065	-0.534**	-0.491**	-0.527**	-0.535**	0.982**					
Civilian Age (8)	2.659	0.897	0.054	0.006	0.023	0.016	-0.015	-0.062	-0.015				
PO Age (9)	2.746	0.317	-0.106	0.099	0.107	0.109	0.101	-0.057	-0.066	0.164			
Quality (10)	4.688	0.556	0.005	-0.428**	-0.372**	-0.429**	-0.409**	0.349**	0.369**	-0.082	-0.166		
Blur Civilian Race [^] (11)	0.390	0.153	0.103	-0.174	-0.079	-0.176	-0.189	0.104	0.167	0.146	0.071	0.005	
Time of Day (12)	0.414	0.450	0.196	0.078	0.145	0.064	0.089	0.188	0.203	-0.063	0.056	0.194	0.179

Note. PO = police officer; [^]ratings given on blurred videos

* $p < .05$, ** $p < .01$

civilians and 19 White civilians. In addition to eliminating potential confounds, equating the videos (e.g., on aggressiveness) allows for a clearer assessment of how individual differences impact evaluations of interracial police officer-civilian interactions for future research. Recall that for equating, analyses were performed as a function of videos. For each rating dimension, we (a) compared ratings based on civilian race (Black or White) and (b) calculated inter-rater reliability. Additionally, for aggressiveness and age ratings because individuals rated both police officers and civilians, we ran a 2 (Rating Target: police officer, civilian) \times 2 (Civilian Race: Black, White) linear mixed effects model using the function `lmer` in the R package `lme4` (Bates, Maechler, Bolker, & Walker, 2015). All *t*-test analyses were run in the R programming language using the stats package (R Core Team, 2018). We first present equating information for police officer and civilian characteristics (e.g., age, aggressiveness) and then present scene characteristics (e.g. time of day, duration). Lastly, we present interclass correlations of the ratings for the set of 38 equated videos.

POLICE OFFICER AND CIVILIAN CHARACTERISTICS

Videos did not differ in agreement of the racial categorization of the civilians ($M = 89.93\%$ rated the civilian as Black, $SD = 0.033$ and $M = 91.49\%$ rated the civilian as White, $SD = 0.034$; $t(37) = 1.429$, $p = .167$, $CI_{95\%} = [-0.006, 0.037]$). We also assessed whether age of the police officers or civilians varied, using a 2 (Rating Target: police officer, civilian) \times 2 (Civilian Race: Black, White) linear mixed effects model. We only observed a main effect of rating target, such that police officers ($M = 2.816$, $SD = 0.299$) were rated as older than civilians ($M = 2.487$, $SD = 0.781$) regardless of civilian race, $t(36) = 2.412$, $p = .021$, $CI_{95\%} = [0.031, 0.298]$.

Then, we wanted to ensure that the aggressiveness of the interaction between the police officer and the civilian, as well as aggressiveness of the police officers and civilians after blurring, were similar for videos of Black and White civilians. To equate for aggressiveness, we first ensured that the overall interaction was similarly rated in aggressiveness and arousal to the videos for Black and White civilians. Using ratings from the non-blurred videos, final Black and White civilian videos did not differ in overall aggressiveness of the interaction (Black civilian videos $M = 3.043$, $SD = 1.117$ and White civilian videos $M = 3.042$, $SD = 1.653$), $t(36) = 0.003$, $p = .998$, $CI_{95\%} = [-0.927, 0.930]$ (Figure 2a) or arousal to the videos (Black civilian videos $M = 2.718$, $SD = 0.781$ and White civilian videos $M = 2.656$, $SD = 1.162$), $t(36) = 0.193$, $p = .848$, $CI_{95\%} = [-0.590, 0.714]$ (Figure 2b).

Next, we wanted to ensure that aggressiveness ratings of the civilians or police officers in the videos were not driven by stereotypes; therefore, we examined aggressiveness ratings of the blurred videos. Participants were similarly below 50% accuracy in identifying civilian race from the blurred videos (accuracy = 34%) (Black civilian videos $M = 0.316$, $SD = 0.138$ and White civilian videos $M = 0.364$, $SD = 0.108$), $t(36) = -1.175$, $p = .248$, $CI_{95\%} = [-0.129, 0.034]$. Then, we assessed whether aggressiveness ratings from blurred videos varied as a function of ratings of the police officer and civilian using a 2 (Rating Target: police officer, civilian) \times 2

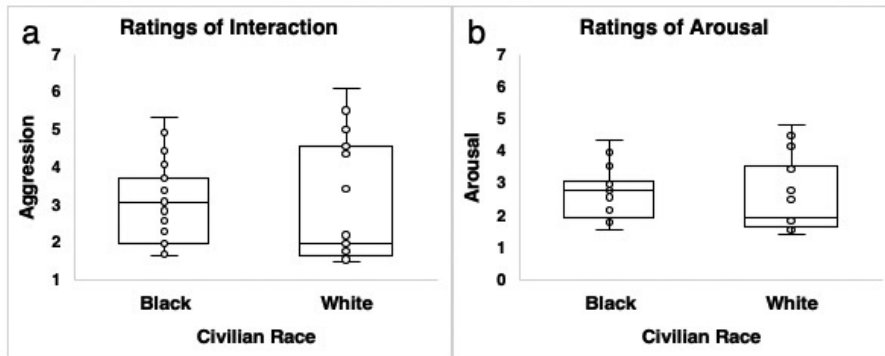


FIGURE 2. Box plots showing quartiles exclusive of the median for (a) ratings of the aggressiveness of the interaction (1 = “Not at all aggressive” to 7 = “Extremely aggressive”) and (b) ratings of arousal (1 = “Not at all arousing” to 7 = “Extremely arousing”) by civilian race. Data points represent the average rating by video. Both ratings were given on non-blurred videos. There were no significant differences as a function of civilian race in ratings of overall aggressiveness of the interaction or arousal.

(Civilian Race: Black, White) linear mixed effects model. We observed only a main effect of rating target such that police officers ($M = 2.826$, $SD = 1.505$) were rated as more aggressive than civilians ($M = 1.979$, $SD = 1.019$) regardless of civilian race, $t(36) = 6.338$, $p < .001$, $CI_{95\%} = [0.292, 0.554]$ (Figure 3). Note that although variance in ratings of aggressiveness for the blurred videos did differ for White civilian videos when comparing police officers versus civilians, $F(1,36) = 0.347$, $p = .030$, $CI_{95\%} = [0.134, 0.900]$, the variance in ratings of aggressiveness for the blurred videos did not differ for Black civilian videos when comparing police officers versus civilians, $F(1,36) = .666$, $p = .397$, $CI_{95\%} = [0.257, 1.730]$. Most importantly for equating as a function of race, the variance in ratings of aggressiveness for the blurred videos did not differ as a function of race for either civilian or police officer targets (civilian targets, $F(1,36) = 0.697$, $p = .276$, $CI_{95\%} = [0.362, 1.340]$; police officer targets, $F(1,36) = 0.538$, $p = .198$, $CI_{95\%} = [0.207, 1.397]$).

SCENE CHARACTERISTICS

Next, we wanted to ensure that aspects of the scene were similar for videos of Black and White civilians. Final stimuli of each condition were equally of average scene pleasantness (Black civilian videos $M = 3.438$, $SD = 0.604$ and White civilian videos $M = 3.443$, $SD = 0.481$), $t(36) = -0.024$, $p = .981$, $CI_{95\%} = [-0.363, 0.355]$; and safety (Black civilian videos $M = 3.491$, $SD = 0.592$ and White civilian videos $M = 3.586$, $SD = 0.451$), $t(36) = -0.553$, $p = .584$, $CI_{95\%} = [-0.441, 0.252]$. The video quality was average and did not differ by civilian race (Black civilian videos $M = 4.671$, $SD = 0.524$ and White civilian videos $M = 4.674$, $SD = 0.470$), $t(36) = -0.016$, $p = .987$, $CI_{95\%} = [-0.330, 0.325]$. The proportion of ratings for morning, early afternoon, late afternoon, and evening/night for the Black civilian videos were 0.063, 0.189,

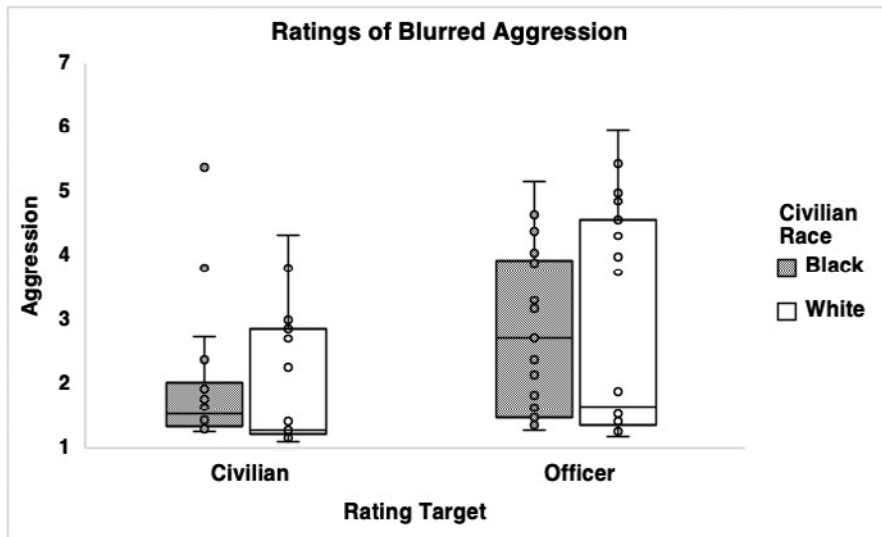


FIGURE 3. Box plots showing quartiles exclusive of the median for ratings of aggressiveness (1 = “Not at all aggressive” to 7 = “Extremely aggressive”) by civilian race. Data points represent the average rating by video. Police officers were rated as more aggressive than civilians regardless of civilian race. Aggressiveness ratings of the police officer and civilian are taken from blurred videos.

0.194, and 0.554, respectively, and for the White civilian videos were 0.075, 0.136, 0.248, and 0.541, respectively. For time of day, we created a composite proportion for each video of the number of participants who said the video took place during the day (i.e., either morning, early afternoon, or late afternoon) and compared that composite to the proportion of evening/night. Final stimuli in each condition were similarly a mixture of time of day (Black civilian videos $M = 0.446$, $SD = 0.458$ and White civilian videos $M = 0.459$, $SD = 0.470$), $t(36) = -0.083$, $p = .934$, $CI_{95\%} = [-0.318, 0.293]$. Video duration did not differ by civilian race (Black civilian videos $M = 15.790$ s, $SD = 5.308$, range: 7–27 s and White civilian videos $M = 14.263$ s, $SD = 6.026$, range: 4–23 s), $t(36) = 0.829$, $p = .413$, $CI_{95\%} = [-2.210, 5.263]$.

See Table 3 for a summary of ratings based on civilian race (Black or White) and Table 4 for correlations among ratings for the 38 equated videos.

INTERRATER RELIABILITY

We were also interested in evaluating how similar the videos were rated for the equated videos. In order to examine this, we calculated intraclass correlation coefficients (ICC) and 95% confidence intervals for the 38 equated videos using the function ICC in the R package psych (Revelle, 2018). We calculated two estimates of ICC: ICC2k and ICC3k (Koo & Li, 2016). ICC2k and ICC3k estimates revealed high consistency and agreement among video ratings across all dimensions. For example, there was high agreement for aggression ratings: interaction

TABLE 3. Video Equating for Each Dimension by Civilian Race

Dimension	Black Civilian Videos		White Civilian Videos		<i>t</i> (36)	<i>p</i>	<i>CI</i> _{95%}
	Mean	<i>SD</i>	Mean	<i>SD</i>			
Duration	15.790	5.308	14.263	6.026	0.829	0.413	[-2.210, 5.263]
Interaction Aggression	3.043	1.117	3.042	1.653	0.003	0.998	[-0.927, 0.930]
Civilian Aggression [^]	1.969	1.042	1.990	1.025			
PO Aggression [^]	2.822	1.276	2.830	1.740			
Arousing	2.718	0.781	2.656	1.162	0.193	0.848	[-0.590, 0.714]
Pleasant	3.438	0.604	3.443	0.481	-0.024	0.981	[-0.363, 0.355]
Safe	3.491	0.592	3.586	0.451	-0.553	0.584	[-0.441, 0.252]
Civilian Age	2.446	0.658	2.528	0.904			
PO Age	2.822	0.343	2.810	0.258			
Quality	4.671	0.524	4.674	0.470	-0.016	0.987	[-0.330, 0.325]
Blur Civilian Race [^]	0.316	0.138	0.364	0.108	-1.175	0.248	[-0.129, 0.034]
Time of Day	0.446	0.458	0.459	0.470	-0.083	0.934	[-0.318, 0.293]

Note. Equated videos do not significantly differ in any dimension depending on civilian race. PO = police officer; [^]Ratings given on blurred videos

aggression $ICC2k = .987$, $CI_{95\%} = [0.981, 0.993]$, $ICC3k = .993$, $CI_{95\%} = [0.989, 0.996]$; civilian aggression from the blurred videos $ICC2k = .987$, $CI_{95\%} = [0.981, 0.993]$, $ICC3k = .992$, $CI_{95\%} = [0.987, 0.995]$; and police officer aggression from the blurred videos $ICC2k = .993$, $CI_{95\%} = [0.989, 0.996]$, $ICC3k = .995$, $CI_{95\%} = [0.992, 0.997]$. See Table 5 for a summary of all ICC estimates collapsing across civilian race, as well as separately for Black civilian videos and White civilian videos. See Supplemental Material for detailed tables of all ICC estimates and 95% CIs.

DISCUSSION

We developed an IPOC database of videos depicting a third-person perspective, real-world interactions between a White, male police officer with either a Black or White, male civilian. These short, audio-less videos depict a range of aggressiveness and reactions from both police officers and civilians during arrests and represent some of the types of real-life videos the public and juries are exposed to. Our first aim was to collect ratings of dimensions that may vary between the 78 videos to allow future researchers to select videos based on their dimensions of interest. The videos were rated on police officer and civilian race and age, police officer and civilian aggressiveness, overall aggressiveness, arousal, scene characteristics (i.e., pleasantness, safety, and time of day), and video picture quality. Our second aim was to select a subset of Black and White videos that did not differ on the rated dimensions for use in future intergroup research. Our equated set of 38 videos consists of 19 of each civilian race (i.e., Black and White). The specific videos that were equated for Aim 2 are only one combination of videos that could possibly be equated across all dimensions. Researchers may select a different combination

TABLE 4. Correlations of equated video (n=38) ratings (Aim 2).

Dimension	Mean	SD	1	2	3	4	5	6	7	8	9	10	11
Duration (1)	15.026	5.654											
Interaction Aggression (2)	3.042	1.392	0.208										
Civilian Aggression [^] (3)	1.979	1.019	0.087	0.887**									
PO Aggression [^] (4)	2.826	1.505	0.210	0.984**	0.862**								
Arousing (5)	2.687	0.977	0.240	0.991**	0.870**	0.965**							
Pleasant (6)	3.440	0.539	0.069	-0.405*	-0.432**	-0.363*	-0.385*						
Safe (7)	3.538	0.522	0.086	-0.435**	-0.459**	-0.389*	-0.425**	0.982**					
Civilian Age (8)	2.487	0.781	0.170	0.244	0.255	0.258	0.230	-0.130	-0.080				
PO Age (9)	2.816	0.299	-0.084	0.134	0.235	0.126	0.127	0.132	0.105	0.020			
Quality (10)	4.672	0.491	0.114	-0.473**	-0.387*	-0.493**	-0.460**	0.100	0.122	0.026	-0.299		
Blur Civilian Race [^] (11)	0.340	0.125	0.060	0.022	0.110	0.033	0.014	-0.062	-0.012	-0.107	0.08	0.106	
Time of Day (12)	0.452	0.458	0.053	0.076	0.116	0.034	0.102	0.238	0.245	0.013	0.044	0.055	0.363*

PO=police officer, [^]Ratings given on blurred videos

*p<.05, **p<.01

TABLE 5. Overall and by Civilian Race Intraclass Correlation Coefficient (ICC) for Each Dimension for Equated Videos ($n = 38$)

Dimension	All Videos		Black Civilian Videos		White Civilian Videos	
	ICC2k	ICC3k	ICC2k	ICC3k	ICC2k	ICC3k
Interaction Aggression	0.987	0.993	0.979	0.989	0.992	0.995
Civilian Aggression [^]	0.987	0.992	0.988	0.993	0.987	0.991
PO Aggression [^]	0.993	0.995	0.989	0.992	0.995	0.996
Arousing	0.969	0.985	0.951	0.978	0.978	0.989
Pleasant	0.941	0.959	0.952	0.970	0.928	0.951
Safe	0.936	0.956	0.951	0.969	0.914	0.941
Civilian Age	0.955	0.975	0.934	0.969	0.968	0.980
PO Age	0.784	0.882	0.822	0.915	0.733	0.834
Quality	0.931	0.960	0.942	0.967	0.923	0.955
Blur Civilian Race [^]	0.803	0.845	0.679	0.739	0.851	0.885
Time	0.998	0.998	0.997	0.997	0.998	0.998

Note. PO = police officer; [^]Ratings given on blurred videos

from the entire set of 78 videos, focusing on the dimensions that are particularly central to their work.

It is our hope that the IPOC Database will contribute to future social cognition research by making it an easily accessible, publicly available resource. Specifically, the resulting database will enable researchers in future research to assess individual differences in response to interracial police officer-civilian interactions and/or to manipulate exposure to interracial police officer-civilian interactions.

VIDEO PERSPECTIVE

The IPOC Database represents a third-person perspective (e.g., filmed interaction). Video-based evaluations may differ depending on the viewing perspective, such as a first-person (e.g., body-worn camera or BWC) or a third-person (e.g., surveillance cameras) perspective. For example, studies have demonstrated that video evidence focusing exclusively on the civilians using BWC tend to lead to harsher judgments of the civilians and less severe judgments of the police officers than videos depicting the entire scene using surveillance cameras (Lassiter et al., 2007; Lassiter, Geers, Handley, Weiland, & Munhall, 2002). Because a video's perspective can impact evaluations even when people are experts or forewarned about possible biases, investigating how perceivers' evaluations are affected by various perspectives will ultimately be important for understanding the impact of the broad array of increasingly available video-recordings of police officer-civilian interactions. Specifically, given the growing prevalence and distribution of videos captured by bystanders on their cellphones, perceivers will be increasingly exposed to

these videos through traditional and social media. Therefore, understanding how aspects of these videos influence our evaluations is important. Although various types of first-person videos have been used in research, including police BWC, dash cam footage, and surveillance videos of a suspect that do not include police officers (Caruso, Burns, & Converse, 2016; Culhane et al., 2016; Hedberg, Katz, & Choate, 2017; Smykla, Crow, Crichlow, & Snyder, 2016), these databases can be difficult to obtain and may not include ratings across many dimensions that are necessary for some experiments. Therefore, in future iterations of the IPOC Database, we plan to include first-person perspective videos (e.g., surveillance, BWC, and dashboard videos).

LIMITATIONS

There are a few limitations of the current IPOC Database. For example, there is variability among the uniforms of the police officers and civilian clothes. Variation in police officer and civilian attire could impact evaluations (Simpson, 2017, 2018; Singer & Singer, 1985). Additionally, due to the limited availability of Black police officers and female police officers and civilians in the original source, the videos only contain White male police officers and either Black or White male civilians. In order to reflect actual population composition, it will be important to include future videos in the IPOC Database that include people of all races and genders.

We removed audio to reduce the influence of tone of voice or audio content on ratings. However, previous research has found that principles of procedural justice (e.g., giving the civilian an opportunity to express their views, explaining why the civilian was stopped) conveyed via speech leads to increased perceptions of police officer fairness and respect (Mazerolle, Bennett, Antrobus, & Eggins, 2012). Further, police officers are typically verbally more respectful towards White civilians (e.g., police officers apologizing; using last names) than Black civilians (e.g., telling civilians to put their hands on the wheel; using first names), even when controlling for severity and outcome of the interaction (Voigt et al., 2017). These findings suggest that future work could also include police officer-civilian videos with audio.

Another limitation with the current database is that the videos were shot with a camera crew present, perhaps impacting how the police officers and civilians behaved. For example, field research with police officers demonstrated that police officer use of force during BWC work shifts was reduced compared to work shifts without BWC (Ariel, Farrar, & Sutherland, 2015). Although beliefs are mixed regarding whether BWC will change policing (Jennings et al., 2014; Smykla et al., 2016), experimental data suggest that BWCs do change police officer behavior. Although the presence of the crew may have changed the police officers' and civilians' behavior, this was not a confound in the current database because all the videos were filmed with a crew present. Moreover, editing the videos to focus on the arrest without much additional contextual information minimized the camera crew's impact on ratings. Future iterations of the IPOC Database will include videos without camera crews present (e.g., surveillance footage).

An additional concern is that the original footage utilized from the Cops TV show was produced for entertainment and not experimental research. Nonetheless, the show depicts actual real-world interactions. Because the purpose of creating this database was to provide videos of police officer-civilian interactions for use in future research, the original TV format should not impact internal validity of the resulting equated videos depicting Black and White civilians. Additionally, the producers of the show purport that they endeavored to capture interactions as they are in real life as a documentary, while critics characterize the show as putting pressure on police officers to make arrests or even illegally targeting poor and minority communities (Taberski, 2019). In fact, some civilians may have preferred not to have their footage released. All of these factors impact the initial population of videos we used to create the database, and these factors are important avenues for future consideration of the external validity of these videos.

Although some of the videos in the current database portray aggressiveness and at times violence, the videos are still relatively less extreme than some publicly disseminated videos, for example those that portray shootings or beatings. However, given the number of exchanges between the public and police, these extreme encounters are far less representative. For example, the FBI estimates that in 2017 10.6 million arrests were made (FBI: Uniformed Crime Reporting, 2018). Of the 10.6 million arrests, *The Washington Post* reports that 987 civilians were shot and killed by police in 2017 (The Washington Post, 2017). In addition, in 2017 46 law enforcement officers died from injuries incurred in the line of duty during felonious incidents, and 60,211 officers were assaulted while performing their duties (FBI: Uniformed Crime Reporting, 2018). Therefore, extreme encounters are rare compared with the number of arrests. We believe it is important for researchers to consider a variety of interactions varying in aggressiveness to allow for more generalizability. However, with more police departments utilizing BWC and increases in publicly disseminated recordings, video recordings of the more extreme interactions that result in civilian or police officer deaths may become available in the future.

ACCESSING THE DATABASE

Both the 78 videos (Aim 1) and the subset of 38 videos equated by race (Aim 2) along with the ratings will be publicly available in order to facilitate future research on this important and timely topic. We will maintain this database and plan to accumulate more videos for future research. We provide unrestricted access for academic research to the IPOC Database. To access the database, interested users must submit a form (see Supplemental Material Figure S2) available on our lab website (<http://ifsnlab.org/stimuli>). We will not screen access to the videos; however, we will track the number of downloads and users' reported purposes for downloading the videos. Following form submission, users will be automatically redirected to a download link for a zipped file containing all of the IPOC videos, the ratings, and a variable guide.

REFERENCES

- Ariel, B., Farrar, W. A., & Sutherland, A. (2015). The effect of police body-worn cameras on use of force and citizens' complaints against the police: A randomized controlled trial. *Journal of Quantitative Criminology*, *31*, 509–535. <https://doi.org/10.1007/s10940-014-9236-3>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bosman, J., & Davey, M. (2019, January 17). 3 officers acquitted of covering up for colleague in Laquan McDonald killing. *New York Times*. Retrieved from <https://www.nytimes.com/2019/01/17/us/laquan-mcdonald-officers-acquitted.html>
- Caruso, E. M., Burns, Z. C., & Converse, B. A. (2016). Slow motion increases perceived intent. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 9250–9255. <https://doi.org/10.1073/pnas.1603865113>
- Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, *83*, 1314–1329. <https://doi.org/10.1037/0022-3514.83.6.1314>
- Culhane, S. E., Boman, J. H., & Schweitzer, K. (2016). Public perceptions of the justifiability of police shootings: The role of body cameras in a pre- and post-Ferguson experiment. *Police Quarterly*, *19*, 251–274. <https://doi.org/10.1177/1098611116651403>
- Devine, P. G., & Elliot, A. J. (1995). Are racial stereotypes really fading? The Princeton Trilogy revisited. *Personality and Social Psychology Bulletin*, *21*, 1139–1150. <https://doi.org/10.1177/01461672952111002>
- FBI: Uniformed Crime Reporting. (2018). 2017 law enforcement officers killed & assaulted. Retrieved April 12, 2019, from <https://ucr.fbi.gov/leoka/2017/home>
- Granot, Y., Balcetis, E., Feigenson, N., & Tyler, T. (2018). In the eyes of the law: Perception versus reality in appraisals of video evidence. *Psychology, Public Policy, and Law*, *24*, 93–104. <https://doi.org/10.1037/law0000137>
- Hedberg, E. C., Katz, C. M., & Choate, D. E. (2017). Body-worn cameras and citizen interactions with police officers: Estimating plausible effects given varying compliance levels. *Justice Quarterly*, *34*, 627–651. <https://doi.org/10.1080/07418825.2016.1198825>
- Jennings, W. G., Fridell, L. A., & Lynch, M. D. (2014). Cops and cameras: Officer perceptions of the use of body-worn cameras in law enforcement. *Journal of Criminal Justice*, *42*, 549–556. <https://doi.org/10.1016/j.jcrimjus.2014.09.008>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*, 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kubota, J. T., Banaji, M. R., & Phelps, E. A. (2012). The neuroscience of race. *Nature Neuroscience*, *15*, 940–948. <https://doi.org/10.1109/ICPR.2002.1048479>
- Lassiter, G. D., Diamond, S. S., Schmidt, H. C., & Elek, J. K. (2007). Evaluating videotaped confessions: Expertise provides no defense against the camera-perspective effect. *Psychological Science*, *18*, 224–226. <https://doi.org/10.1111/j.1467-9280.2007.01879.x>
- Lassiter, G. D., Geers, A. L., Handley, I. M., Weiland, P. E., & Munhall, P. J. (2002). Videotaped interrogations and confessions: A simple change in camera perspective alters verdicts in simulated trials. *Journal of Applied Psychology*, *87*, 867–874. <https://doi.org/10.1037/0021-9010.87.5.867>
- Lockhart, P. R. (2019, January 17). Chicago police officers found not guilty of covering up the Laquan McDonald shooting. *Vox*. Retrieved from <https://www.vox.com/2019/1/17/18187043/laquan-mcdonald-chicago-conspiracy-trial-not-guilty-verdict>
- Mazerolle, L., Bennett, S., Antrobus, E., & Eggins, E. (2012). Procedural justice, routine encounters and citizen perceptions of police: Main findings from the Queensland Community Engagement Trial (QCET). *Journal of Experimental*

- Criminology*, 8, 343–367. <https://doi.org/10.1007/s11292-012-9160-1>
- Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, 81, 181–192. <https://doi.org/10.1037/0022-3514.81.2.181>
- R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.r-project.org>
- Revelle, W. (2018). psych: Procedures for psychological, psychometric, and personality research. R CRAN Project. Retrieved from <https://cran.r-project.org/package=psych>
- Simpson, R. (2017). The Police Officer Perception Project (POPP): An experimental evaluation of factors that impact perceptions of the police. *Journal of Experimental Criminology*, 13, 393–415. <https://doi.org/10.1007/s11292-017-9292-4>
- Simpson, R. (2018). Officer appearance and perceptions of police: Accoutrements as signals of intent. *Policing: A Journal of Policy and Practice*. <https://doi.org/10.1093/police/pay015>
- Singer, M. S., & Singer, A. E. (1985). The effect of police uniform on interpersonal perception. *Journal of Psychology*, 119, 157–161. <https://doi.org/10.1080/00223980.1985.10542882>
- Smykla, J. O., Crow, M. S., Crichlow, V. J., & Snyder, J. A. (2016). Police body-worn cameras: Perceptions of law enforcement leadership. *American Journal of Criminal Justice*, 41, 424–443. <https://doi.org/10.1007/s12103-015-9316-4>
- Taberski, D. (Host). (2019, April 23). Running from COPS [audio podcast]. Retrieved from <https://www.topic.com/runningfromcops>
- The Washington Post. (2017). Fatal force. Retrieved April 12, 2019, from <https://www.washingtonpost.com/graphics/national/police-shootings-2017/>
- Voigt, R., Camp, N. P., Prabhakaran, V., Hamilton, W. L., Hetey, R. C., Griffiths, C. M., . . . Eberhardt, J. L. (2017). Language from police body camera footage shows racial disparities in officer respect. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 6521–6526. <https://doi.org/10.1073/pnas.1702413114>