Downward Spiral: Police-Threat Associations and Perceptions of Aggression During Arrests are Mutually Reinforcing

Vincenzo J. Olivett[1], Madeliene Stults[2], and David S. March[1]

[1]Florida State University, Department of Psychology
[2]University of Denver, Graduate School of Social Work

Corresponding Author:
Vincenzo J. Olivett
olivett@psy.fsu.edu

This manuscript is a preprint accepted for publication in *Journal of Personality and Social Psychology*. Please cite as:

Olivett, V. J., Stults, M. S., & March, D. S. (in press). Downward Spiral: Police-Threat Associations and Perceptions of Aggression During Arrests are Mutually Reinforcing. *Journal of Personality and Social Psychology.*

Abstract

In the United States, encounters among police officers and civilians are laden with the potential for dangerous outcomes. At the same time, the ubiquity of digital and social media has made observing violent police-civilian encounters easier than ever. Perhaps consequently, recent evidence suggests that Americans automatically associate the police with and behaviorally respond to officers as a source of physical threat. Yet little is known about the interplay between observations of violent police encounters and automatic police-threat associations. Four studies (N = 857) reveal a mutually reinforcing dynamic in which (a) automatic police-threat associations shape perceptions of aggression during arrests, (b) perceptions of aggression during arrests influence automatic police-threat associations, and (c) changes in automatic police-threat associations influence downstream perceptions of aggression. That is, people perceive aggression during arrest encounters through the lens of their existing police-threat associations, and these perceptions in turn reinforce those associations.

Statement of Limitations

The current work is primarily limited in generalizability. Our samples consist of U.S. online convenience samples and U.S. undergraduate students. Although diverse in age, race, and gender, our samples are limited in their generalizability to populations outside of the U.S. Additionally, we measured outcomes in an artificial context using experimental measures. While these practices were chosen to strengthen internal validity, they limit ecological validity.

In 2023, more than one-thousand Americans were fatally wounded by the police (Washington Post, 2024). The ubiquity of body worn cameras and digital media have made it easier than ever for people to witness such encounters (Graeff et al., 2014). Perhaps consequently, Americans view police officers as sources of danger (Pickett et al., 2022), hold automatic associations linking the police with physical threat (Olivett & March, 2021; Sargent & Newman, 2020, 2023), and respond defensively to the police (Olivett & March, 2024)[1]. Despite this bourgeoning work and the negative social consequences of witnessing police violence (Campbell & Valera, 2020; Graziano, 2019; Weitzer, 2002; Weitzer & Tuch, 2005), the relationship between observing police violence during arrests and police-threat associations is empirically untested. The current work focuses on three non-mutually exclusive ways in which automatic police-threat associations and witnessing officer use-of-force may causally interact: associations influence perception, perception influences associations, and associations and perceptions are mutually reinforcing.

Associations, Perceptions, and Their Mutually Reinforcing Dynamic

People interpret the world though their lens of pre-existing beliefs and expectations, and social group-level associations are no exception (Bruner, 1957). In other words, existing associations shape the ways in which people perceive and interpret subsequent information.

[1]We use the term "association" to refer to a mental representation linking a concept and an attribute in memory (Greenwald, et al., 1998). We use the term "attitude" to refer to specific associations between a concept and an evaluative valence (i.e., negativity/positivity; Fazio, 2001). We use the term "evaluation" to refer to the activation and expression of associations or attitudes on a behavioral task (Eagly & Chaiken, 1993). In the current work, we assess *associations* linking *police officers* to *physical threat* using indirect behavioral tasks. We describe such behavioral tasks as "indirect", because they infer the strength of an individual's association in ways other than prompting that person to introspect and verbally report (Corneille & Hütter, 2020). Because indirect measures entail relatively less opportunity for respondents to engage in controlled processing, we classify the police-threat association as relatively more *automatic.* Based on features of the indirect measures used to assess the police-threat association in the current (the Misattribution Procedure) and past research (postural sway, startle eyeblink, approach-avoidance task), we argue that the police-threat association is automatic in that it may become active unintentionally under certain conditions.

Indeed, group-based associations shape trait impressions of individual members of those groups and perceptions of their behaviors (Duncan, 1976; Fazio, 1990; Olson & Fazio, 2004). For instance, Germans who held more negative automatic associations of Turkish people more negatively rated ambiguous behaviors (e.g., "approaching a woman on the dance floor and asking for her telephone number") when performed by Turkish compared to German men (Gawronski et al., 2003). Meaning, when witnessing police violence, people's automatic associations towards the police may shape how they perceive the behaviors of officers' and civilians' during arrests. Specifically, we expect that individuals with stronger police-threat associations will perceive officer's use of force as more aggressive and civilian's behavior as less aggressive.

Just as preexisting associations shape perceptions, perceptions may likewise shape associations. Indeed, witnessing others experiencing physical harm can foster or reinforce (i.e., strengthen) an association between the agent of harm and threat (i.e., "social fear learning"; Olsson & Phelps, 2007). And moreover, highly negative information can rapidly produce short-term shifts in automatic associations (Cone & Ferguson, 2015; Kurdi & Charlesworth, 2023). Meaning, perceptions of violent police-civilian encounters may shape automatic police-threat associations, at least in the short-term. We therefore expect that perceiving more (vs. less) aggression in officer use-of-force will commensurately strengthen one's preexisting police-threat association.

And critically, the temporal dynamic between associations and perception is not fixed, but iterative. Meaning, associations can shape perceptions, perceptions can influence associations, and changes in either inform the next iteration (Cunningham & Zelado, 2007). Thus, the above-described effects may operate simultaneously in a mutually reinforcing cycle.

That is, in an iterative dynamic playing out continuously over time, police-threat associations and perceptions of aggression during witnessed encounters may be mutually reinforcing. We therefore expect that perceiving more versus less aggression in officer use-of-force may strengthen one's police-threat association, and these strengthened associations may subsequently influence perceptions of use-of-force aggression in future encounters.

This mutually reinforcing dynamic is consistent with theoretical proposals suggesting a causal pathway from automatic associations to downstream behaviors (e.g., Fazio, 1990; Olson & Fazio, 2004). However, evidence that experimentally induced changes in automatic associations have "causal force" in shaping downstream judgments and behaviors is modest and inconsistent. A recent meta-analysis (Forscher et al., 2019) found that although automatic associations (as assessed by indirect measures) can indeed be modified experimentally, these changes typically produce weak, unreliable effects on explicit attitudes and downstream behavior. The meta-analysis further indicated that implicit measure changes usually do not strongly mediate subsequent changes in explicit attitudes or behaviors. In other words, the lasting effects of experimentally induced automatic association change has been difficult to evidence empirically. The current set of studies provides multiple within-person tests of automatic association change *and* its effects on downstream judgments. Therefore, beyond clarifying the dynamics between perceptions of police violence and associations towards the police, the current research offers empirical tests of classic theories of automatic associations that have previously seen limited empirical evidence.

Current Work

The current work tests the idea of a mutually reinforcing dynamic between perceiving aggression and automatic police-threat associations across four studies. Study 1 examines the

basic relationship between existing police-threat associations and perceptions of aggression during witnessed arrests. We expect that stronger baseline police-threat associations will lead to heightened perceptions of officers relative to civilians as aggressive. Study 2 tests the causal impact of perceived police versus civilian aggression on post-encounter police-threat associations. We test causality in two ways: by manipulating perceived aggression cross-sectionally, and by examining the time-lagged relationship between perceptions of aggression and subsequent police-threat associations. We predict that witnessing higher versus lower hostility arrests will heighten perceptions of police relative to civilian aggression, which will strengthen subsequent police-threat associations. Study 3 examines the iterative and mutually reinforcing nature of these processes by testing whether the perception-driven changes in police-threat associations established in Study 2 shape subsequent perceptions of aggression. Specifically, we expect that strengthened police-threat associations will lead to greater perceived officer versus civilian aggression in subsequently viewed arrests. Finally, to explore the persistence of this effect, a preregistered Study 4 (1) replicates Studies 2 and 3 and (2) extends Study 3 by introducing an approximately 1-day delay before assessing the critical dependent measure (i.e., perceptions following induced changes in associations). Two plausible outcomes exist, along with considerable intermediate possibilities. If the change in automatic police-threat associations persist over this delay, shifts in police-threat associations should influence perceptions of aggression during arrests viewed one day later. Conversely, if such automatic associations are transient or stable for an unknown time before decaying at an unknown rate, we would not expect them to affect perceptions after this delay.

Transparency and Openness

Studies 1-3 are not preregistered. Study 4 is preregistered. Study 2 replicates and extends Study 1. Study 3 replicates and extends Study 2. Study 4 replicates and extends Studies 2 and 3. We report how we determined sample sizes, all data exclusions (if any), all manipulations, and all measures in each study. Supplemental analyses mentioned in the text are available in the supplemental materials. Primary data, code, and materials are available at https://osf.io/rfn9h/?view_only=126462cde54940d2886527019ffb98d0. The Study 4 preregistration is available at https://osf.io/5auje/?view_only=1c969c72176c471db845ce3da84b635e

Study 1

Study 1 tests whether the strength of preexisting automatic police-threat associations predicts perceptions of police aggression during observed arrests. Participants complete a measure of automatic police-threat associations prior to viewing several videos of police officers arresting civilians. Participants then rate each of these videos on their perceptions of officer and civilian aggression.

Methods

Data were collected from one-hundred fifty-three Prolific workers. Twenty-three participants were excluded from analysis for failing to complete the study. One participant was excluded for indicating the same response on all trials of the measure of police-threat association (Kurdi et al., 2024). Statistical conclusions do not differ when including this participant. This resulted in a final sample of 129 individuals (47 women, 81 men, 1 nonbinary; 59 White, 58 Black, 5 Asian, 3 Native American, 4 other; 12 Hispanic, 117 Nonhispanic; M_{age} = 35.37, SD_{age} = 7.91). Sample size was determined by the maximum number of participants who could be

compensated within budget. A post hoc sensitivity analysis of the critical effect revealed an observed effect size greater than the predicted minimal detectable effect size assuming 80% power to detect that effect (Faul et al., 2007).

Misattribution Procedure. We utilized a misattribution procedure to measure automatic police-threat associations (Imhoff et al., 2011; Payne et al., 2005). Misattribution procedures (MPs) are sequential priming tasks that indirectly assesses the strength of automatic associations between a prime category (e.g., police officers) and target attributes (e.g., dangerous). Participants were told that on each trial, a prime (e.g., an image of a police officer) would rapidly flash on the screen before being replaced by a momentarily appearing ambiguous target (i.e., a Chinese symbol). Participants were instructed to ignore the prime and instead indicate whether they feel the Chinese character symbolizes something physically "Safe" or "Dangerous." To the extent that a prime category is associated with danger threat versus safety, participants should more frequently judge Chinese characters following primes of that category as symbolizing something physically "Dangerous" versus "Safe." Participants completed an MP assessing their strength of the association between police and threat.

Randomly ordered trials consisted of 30 police officer primes and 30 nonpolice primes (e.g., civilians and uniformed nonpolice professionals; Olivett & March, 2021, 2024). Faces on prime images were blurred to minimize effects of idiosyncratic features such as attractiveness and masculinity (see Supplemental Materials for all stimuli). Each trial began with a 1,000 ms fixation cross, followed by a 100 ms prime image, a 125 ms blank screen, a 125 ms Chinese character, and then ended with a visual noise mask that lasted until participants indicated "Safe" or "Dangerous" using a designated key response. After completing 8 practice trials with neutral image primes (e.g., desk lamps, filing cabinets), participants completed the full block.

Perceptions of Aggression During Arrests. To assess perceptions of police officer and civilian aggression during police-civilian encounters, participants viewed and responded to 38 videos of police officers arresting civilians (Dang et al., 2020, 2022; a database constructed of clips from the television show *Cops*). Videos range from 7s to 22s in duration and depict real life footage of an officer detaining a civilian with handcuffs.[2] Immediately after viewing each video, participants rated perceptions of officer and civilian aggression. Specifically, they reported agreement that officers and arrested persons were "behaving with aggression," respectively. Since officer and civilian aggression does not occur in a vacuum, but is often consequent of the other (i.e., civilian aggression may lead to officer aggression; Logan, 2016), we account for both officer and civilian aggression within the context of each video. To this end, we created difference scores in perceptions of police versus civilian aggressiveness, which we refer to as aggression *bias* scores. Higher scores indicate more police relative to civilian aggression, which serves to directionally isolate police aggression distinct from civilian aggression. Bias scores above 0 indicate stronger perceptions of officer relative to civilian aggression.

Results

MP trials with slow reaction times (trials with reaction times greater than the third quantile plus three times the interquartile range; Tukey, 1977) were dropped from analyses (2.94% of trials). We operationalized the police-threat association as the proportion of MP trials in which targets were evaluated as "Dangerous" compared to "Safe" following police primes. "Dangerous" evaluations were coded as a 1 and "Safe" evaluations were coded as a 0 (Olivett & March, 2021). Evaluations on the 30 police prime trials were averaged into a single score such that values above 0.5 indicate an increasingly stronger police-threat than safe association and

[2] Half of the videos portrayed a White civilian and half a Black civilian. All officers were White. Videos containing White versus Black civilians were pre-rated as aggregately equal on officer and civilian aggression.

values below 0.5 indicate an increasingly stronger police-safe than threat association. We also averaged responses on trials where targets were preceded by nonpolice primes to derive a nonpolice-threat association score.

On average, police were evaluated as equally dangerous and safe (M = 0.53, CI_{95}[0.48, 0.57]), $t(128) = 1.28$, $p = .203$. Nonpolice were evaluated as more safe than dangerous, (M = 0.40, CI_{95}[0.37, 0.43]), $t(128) = -6.98$, $p < .001$. Police were evaluated as more dangerous than safe than were nonpolice, $t(128) = 4.74$, $p < .001$.

The critical question of Study 1 is whether baseline police-threat associations are associated with perceptions of officer and civilian aggression during witnessed encounters. A regression analysis revealed that heightened baseline police-threat associations were associated with stronger perceptions of police relative to civilian aggression, $\beta = .24$, $b = 0.75$, $t(127) = 2.74$, $p = .007$, $\eta_p^2 = .06$.

Discussion

Study 1 established the basic relationship between automatic police-threat associations and perceptions of police aggression during later viewed arrests. Stronger baseline automatic police-threat associations corresponded with later perceptions of officers relative to civilians as more aggressive. These findings imply that how people view (i.e., interpret, perceive) aggressive policing is related to their preexisting associations towards the police as threatening.

Yet, Study 1 cannot speak to a causal relationship between automatic associations and perceptions. People may perceive police aggression through the lens of their automatic police-threat associations, and those associations may be shaped by their perceptions. Studies 2-4 test these possibilities. Study 2 assesses whether perceiving more versus less police-aggression

causally shapes later measured automatic police-threat associations regardless of one's preexisting police-threat association.

Study 2

Study 2 tests if perceiving more police aggression strengthens later measured automatic police-threat associations. Like Study 1, participants complete an initial MP indexing baseline automatic police-threat associations. Participants then view and rate videos of arrests that were pre-rated to be either relatively high or low on officer hostility. Last, they complete a post-exposure MP assessing police-threat associations. We hypothesized that witnessing higher versus lower hostility arrests would result in stronger post-exposure automatic police-threat associations. And that this effect would be explained by (i.e., mediated by) increased perceptions of officers relative to civilians as aggressive in the high compared to low hostility videos.

Methods

Three-hundred forty-seven undergraduates at a large southeastern U.S. university participated for partial course credit. Participants were excluded from analyses if they failed an attention check (n = 12) or indicated the same response on all trials of either time 1 or time 2 MP (n = 14; Kurdi et al., 2024). Statistical conclusions remain when including these participants. Exclusions resulted in a final sample of 321 individuals (254 women, 60 men, and 7 nonbinary or other; 251 White, 31 Black, 23 Asian, 16 mixed race or other; 89 Hispanic, 232 Nonhispanic; M_{age} = 19.66, SD_{age} = 1.39). Sample size was determined by the maximum number of participants who could be recruited within a single semester. A post hoc sensitivity analysis of the critical effect revealed an observed effect size greater than the predicted minimal detectable effect size assuming 80% power to detect that effect (Faul et al., 2007).

Participants first completed a time 1 baseline MP indexing automatic police-threat evaluations (T1-MP). This MP was identical to that used in Study 1. Immediately following the T1-MP, participants were shown a one-minute relaxation video to buffer against lingering effects from the T1-MP. Participants were then randomly assigned to view 19 videos depicting relatively high (pre-rated mean police hostility = 4.16; n = 161) or relatively low police hostility (pre-rated mean police hostility = 1.49; n = 160) arrests from the same set of 38 videos used in Study 1 (Dang et al., 2020). As in Study 1, for each video, participants rated their perceptions of officer and civilian aggression, respectively. After viewing and rating videos, all participants completed time 2 post-exposure MP (T2-MP) identical to the T1-MP.

Results

MP trials with slow reaction times (trials greater than the third quantile plus three times the interquartile range; Tukey, 1977) were dropped from analyses (2.19% of trials on the T1-MP; 3.04% of trials on the T2-MP). Police-threat versus non-police threat evaluations were indexed as in Study 1. See Table 1 for descriptive statistics among all measures.

Table 1.

Raw means and standard deviations for Study 2 variables

	High Hostility Arrests		Low Hostility Arrests	
	M	***SD***	***M***	***SD***
Baseline Police-Threat Association	0.52	0.20	0.49	0.20
Post-Exposure Police-Threat Association	0.56	0.21	0.49	0.23
Police Aggression	3.92	0.54	2.12	0.75
Civilian Aggression	2.36	0.58	1.87	0.58
Aggression Bias	1.56	0.94	0.26	0.56

We first tested if viewing high versus low hostility arrests manifests as relatively higher automatic police-threat associations. We submitted T2-MP scores to a between subjects

ANCOVA testing for the effect of arrest condition when controlling for T1-MP scores[3]. Police were evaluated as significantly more dangerous relative to safe among participants who viewed high (M_{ls} = 0.55, $CI_{95\%}$[0.53, 0.58]) compared to low hostility arrests (M_{ls} = 0.50, $CI_{95\%}$[0.47, 0.52]), $F(1, 318) = 8.45$, $p = .004$, $CI_{95\%}$[0.02, 0.09], $\eta_p^2 = .03$. Evaluations of nonpolice were unaffected by aggression condition, $F(1, 318) = 0.28$, $p = .598$, $CI_{95\%}$[-0.04, 0.02], $\eta_p^2 = .001$. In other words, viewing more versus less hostile arrests manifested in stronger police-threat associations.

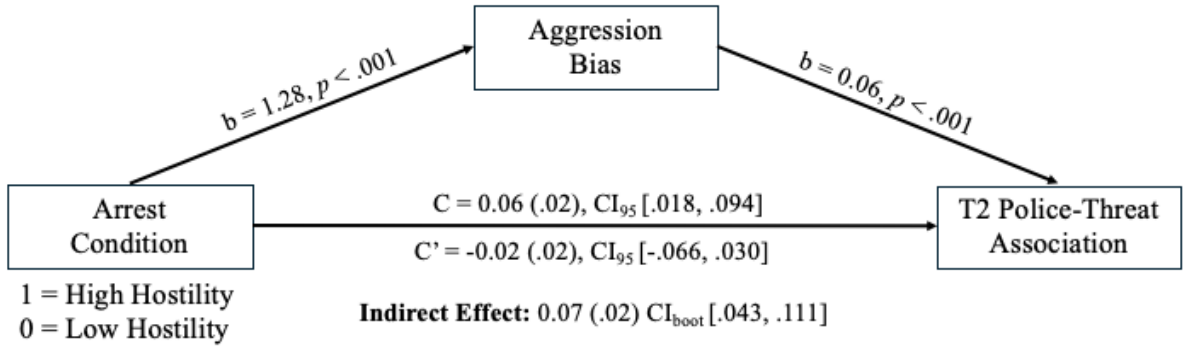
Next, we used a residualized change model to test if perceptions of police relative to civilian aggression promotes stronger automatic police-threat associations. We regressed T2-MP scores onto aggression bias scores, controlling for T1-MP scores. Higher aggression bias scores predicted stronger automatic police-threat associations at time 2, $\beta = .23$, $b = 0.05$, $t(318) = 5.57$, $p < .001$, $\eta_p^2 = .09$. This effect was unmoderated by arrest condition, $p = .489$. That is, regardless of arrest condition, perceiving police versus civilians as more aggressive strengthened the automatic association between police and threat.

Finally, we further assessed whether perceptions of police relative to civilian aggression causally strengthens automatic police threat associations by assessing whether the conditional effect of viewing high versus low hostility arrests on police-threat associations was explained by the effect of aggression bias scores on police-threat associations. Using PROCESS for SAS (Model 4; Hayes, 2022), arrest condition (0 = low hostility, 1 = high hostility) predicted T2-MP scores through aggression bias scores while controlling for T1-MP scores. Confirming the primary and causal role of perceptions of aggression on automatic threat associations, aggression

[3] The same patterns and directions of effects were found when testing the conditional effect using T2-MP minus T1-MP difference scores controlling for T1-MP scores. We report this analysis in the Supplemental Materials.

bias scores fully mediated the effect of arrest condition on T2-MP scores. See Figure 1 for all pathways and direct and indirect effects.

Figure 1. *Study 2 mediation model whereby viewing aggressive arrests strengthens the police threat association through promoting stronger perceptions of police relative to civilian aggression controlling for Time 1 police-threat associations.*



Discussion

Viewing high compared to low hostility arrests manifested as stronger automatic police-threat associations and, across conditions, perceptions of police relative to civilian aggression predicted stronger post-encounter police-threat associations (controlling for pre-encounter associations). Further, the effect of arrest condition on automatic police-threat associations was fully mediated by stronger perceptions of officers relative to civilians as aggressive among those who viewed high versus low hostility arrests. That is, perceiving higher officer relative to civilian aggression during arrest strengthens automatic police-threat associations.

Together, the findings of Studies 1 and 2 support our idea that there exists an iterative cycle where people's associations shape their perceptions, their perceptions then shape their associations, and so on. This may take the form of, for example, people's associations affecting how they perceive aggressive arrests or how they interpret media coverage of violent police, which then reinforces and strengthens the original association, which is then more likely to affect

future perceptions. In other words, just as the association is shaped by an intermediate perception, the change in association may likewise causally shape later perceptions of witnessed arrests. Study 3 tests this possibility by first replicating Study 2's change in the strength of automatic police-threat associations and then assessing whether such changes affect later perceptions of aggressive policing.

Study 3

Study 3 tests the fully mutually reinforcing dynamic between automatic police-threat associations and perceptions of aggression during witnessed arrests. Specifically, Study 3 tests whether the *perception-driven change* in police-threat associations found in Study 2 amplifies perceptions of officer aggression in subsequently viewed arrests. Participants first complete the same procedure as in Study 2. In addition, Study 3 participants then watch a single set of arrests depicting relatively moderate hostility. We expected that induced increases in police-threat associations would shape perceptions of police relative to civilian aggression during subsequently viewed arrests.

Methods

Data were collected from two-hundred seventy-four Prolific workers. Participants were excluded from analyses if they recommended the exclusion of their data (n = 5), indicated that they spoke or read Chinese (n = 5) or indicated the same response on all trials of either time 1 or time 2 MP (n = 21; Kurdi et al., 2024). Statistical conclusions do not differ when including these participants (see Supplemental Materials). Exclusions resulted in a final sample of 243 individuals (113 women, 124 men, and 6 nonbinary or other; 157 White, 64 Black, 14 Asian, 8 mixed race or other; 10 Hispanic, 233 Nonhispanic; M_{age} = 40.74, SD_{age} = 14.44). Sample size was determined by the maximum number of participants who could be compensated within

budget. A post hoc sensitivity analysis of the critical effect revealed an observed effect size greater than the predicted minimal detectable effect size assuming 80% power to detect that effect (Faul et al., 2007).

The initial three phases of the Study 3 were the same as those of Study 2. Participants first completed a time 1 baseline measure of automatic police-threat associations (T1-MP). Immediately following the T1-MP, participants were shown a one-minute relaxation video to buffer effects of completing the T1-MP on subsequent measures. Then, participants were randomly assigned to view 8 videos depicting high (pre-rated mean police aggression = 4.94; n = 116) or low police hostility (pre-rated mean police hostility = 1.32; n = 127) arrests from the same set of videos used in Studies 1 and 2. Participants rated officers and civilians, respectively, within each video on aggression using the same items as in Studies 1 and 2. Participants next completed a time 2 post-exposure police-threat MP (T2-MP). In the fourth and final phase, all participants viewed and rated the same set of 8 moderate hostility arrest videos (pre-rated mean police hostility = 2.72) on officer and civilian aggression.

Results

MP trials with slow reaction times (trials greater than the third quantile plus three times the interquartile range; Tukey, 1977) were dropped from analyses (2.90% of trials on the T1-MP; 2.68% of trials on the T2-MP). Police-threat and non-police threat evaluations were indexed in the same way as in Studies 1 and 2. See Table 2 for descriptive for all measures in the high and low aggression arrest conditions, respectively.

We first sought to replicate Study 2's main findings that perceptions of police relative to civilian aggression manifested in stronger automatic police-threat associations. T2-MP scores were submitted to a between subjects ANCOVA testing for the effect of arrest video condition

while controlling for T1-MP scores. Once again, police were evaluated as more dangerous relative to safe among participants who viewed high (M_{ls} = 0.57, $CI_{95\%}$[0.53, 0.61]) compared to low hostility arrests (M_{ls} = 0.45, $CI_{95\%}$[0.42, 0.49]), $F(1, 240) = 16.76$, $p < .001$, $CI_{95\%}$[0.06, 0.17], $\eta_p^2 = .065$.[4] Unlike in Study 2, evaluations of *non*police were affected by arrest condition. Here, nonpolice were evaluated as more safe relative to dangerous among participants who viewed high, (M_{ls} = 0.34, $CI_{95\%}$[0.32, 0.37]), compared to low hostility arrests, (M_{ls} = 0.39, $CI_{95\%}$[0.37, 0.42]), $F(1, 240) = 7.00$, $p = .009$, $CI_{95\%}$[-0.09, -0.01], $\eta_p^2 = .03$.

Table 2.

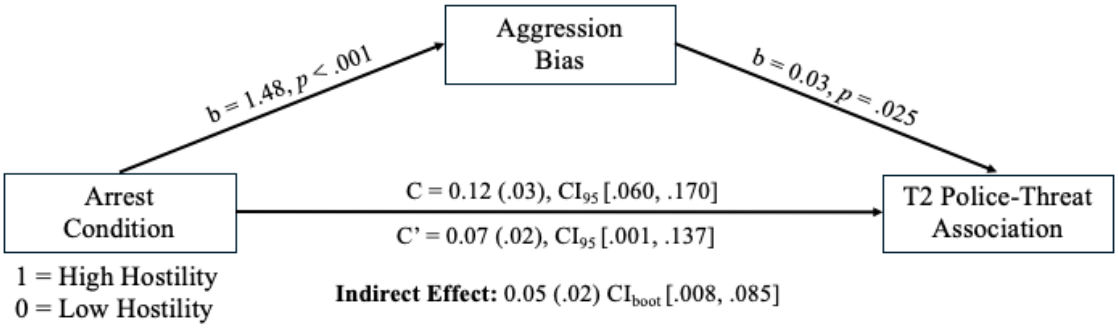
Means and standard deviations for Study 3 variables

	High Hostility Arrests		Low Hostility Arrests	
	M	***SD***	***M***	***SD***
Baseline Police-Threat Association	0.49	0.27	0.50	0.25
Post-Exposure Police-Threat Association	0.57	0.31	0.46	0.27
Police Aggression Time 1	4.18	0.72	1.61	0.82
Civilian Aggression Time 1	2.44	0.90	1.37	0.51
Aggression Bias Time 1	1.73	1.30	0.24	0.72
Police Aggression Time 2	2.83	0.97	3.39	0.92
Civilian Aggression Time 2	1.63	0.69	1.89	0.68
Aggression Bias Time 2	1.21	1.06	1.50	1.12

Also replicating Study 2, perceived aggression during arrest videos predicted T2-MP scores controlling for T1-MP scores, $\beta = .21$, $b = 0.05$, $t(242) = 4.23$, $p < .001$, $\eta_p^2 = .07$. And again, this effect was unmoderated by arrest condition, $p = .232$. As in Study 2, aggression bias scores mediated the effect of Condition on T2-MP scores. See Figure 2 for all pathways and indirect effects. All critical effects from Study 2 replicated in Study 3.

[4] The same patterns and directions of effects were found when testing the conditional effect using T2-MP minus T1-MP difference scores controlling for T1-MP scores. We report this analysis in the Supplemental Materials.

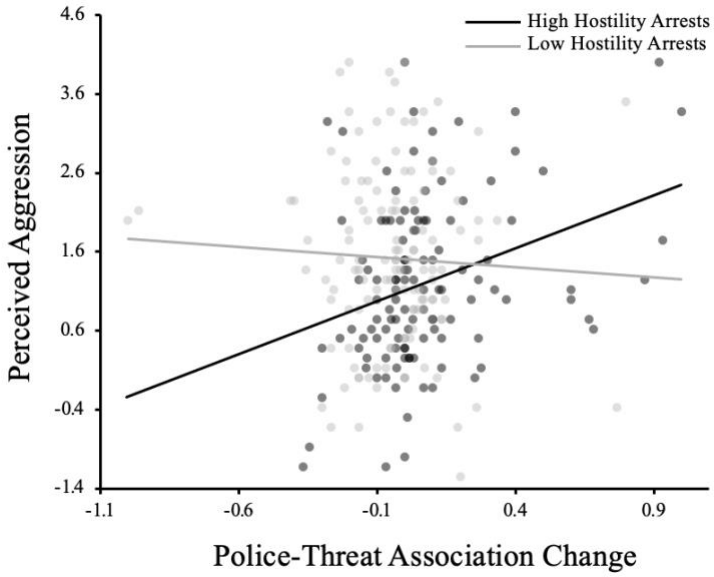
Figure 2. *Study 3 mediation model whereby viewing aggressive arrests strengthens the police threat association through promoting stronger perceptions of police relative to civilian aggression controlling for Time 1 police-threat associations.*



We then tested for the presence of a mutually reinforcing dynamic between perceptions of aggression and police-threat associations. Specifically, we examined whether perception-driven change in automatic police-danger associations affects subsequent perceptions of aggression. We regressed aggression bias scores from the follow-up (moderate) arrest videos onto the change in police-threat association (the difference of T2-MP score minus T1-MP score), arrest condition (high versus low hostility), and their interaction term. This revealed a significant interaction between condition and association change, $F(1,239) = 7.06$, $p = .008$, such that among participants who initially viewed high hostility arrests, increases in the automatic police-threat association led to heightened perceptions of officers relative to civilians as aggressive in later viewed arrests, $\beta = .29$, $b = 1.34$, $t(239) = 3.35$, $p < .001$, $CI_{95\%}[0.55, 2.13]$. Among those who viewed low aggression arrests, changes in the automatic police-threat association did not affect subsequent perceptions of aggression in later viewed arrests, $\beta = -.06$, $b = -0.26$, $t(239) = -0.57$, $p = .589$, $CI_{95\%}[-1.14, 0.63]$. See Figure 3. In other words, changes in (or strengthening of)

the police-threat association due to perceptions of aggression in previously viewed arrests shaped stronger perceptions of aggression in later viewed arrests.

Figure 3. *The relationships between change in police-threat* associations *and later perceptions of officer relative to civilian aggression in Study 3.*



Discussion

Replicating Study 2, viewing high compared to low hostility arrests manifested in stronger automatic associations between police and threat and, across conditions, perceptions of police relative to civilian aggression predicted stronger post-encounter police-threat associations (controlling for pre-encounter associations). This effect was mediated by heightened perceptions of police versus nonpolice as behaving aggressively. Furthermore, these perceived aggression-driven increases in automatic police-threat associations influenced how people interpreted subsequent instances of police aggression. Specifically, increased perceptions of officer versus civilian aggression strengthened later measured police-threat associations, which in turn led to heightened perceptions of police versus civilians as acting aggressively in later viewed arrests.

Together, findings from Studies 1-3 demonstrate that people perceive police aggression through the lens of their existing police-threat associations, and those perceptions subsequently and mutually reinforce those associations. This implies that real-world associations towards the police and perceptions of police aggression exist in an iterative cycle and are mutual reinforcing. People may view the police as more threatening upon viewing aggressive arrests. Consequently, they may perceive later observed arrests as more aggressive in nature, in turn further reinforcing their police-threat associations, and so on.

In terms of downstream implications (i.e., causal force), Study 3 demonstrates that within-person changes in automatic police-threat associations influence judgments measured shortly thereafter. However, Study 3 addresses only short-term causal effects within a single experimental session. Study 4 extends this investigation to explore the durability of these causal effects, specifically testing whether experimentally induced changes in automatic police-threat associations persist and influence judgments beyond the immediate experimental context (i.e., after approximately 24 hours).

Study 4

Study 4 has two aims. The first aim is to provide a preregistered replication of the time-lagged relationship between perceiving aggression during arrests and subsequent automatic police-threat associations. Second, it tests the long-term causal impact (i.e., causal force) of changes in automatic associations on downstream perception. To these aims, participants in an initial experimental session first complete a baseline measure of police-threat associations, then view and rate hostile arrests on police and civilian aggression, and last complete a post-exposure measure of associations. This replicates the entirety of Study 2 and all of Study 3 prior to the second perception task. Then, approximately one day later, participants complete a follow-up

session where they view and rate moderately violent arrests on police and civilian aggression. We expected to replicate the findings of Studies 2 and 3 that percieving aggression while viewing arrests produces a time-lagged change in automatic police-threat associations. As stated in the Introduction, if the change in automatic police-threat associations persist over this delay, shifts in police-threat associations should influence perceptions of aggression during arrests viewed one day later. Conversely, if such automatic associations are transient or stable for an unknown time before decaying at an unknown rate, we would not expect them to affect perceptions after this delay.

Methods

An *a priori* power analyses using the observed effect size from Study 3 (i.e., the relationship between police-threat association change and later aggression perceptions among participants who initially viewed high hostility arrests; $\eta_p^2 = .10$) indicated that a sample of 99 participants would provide 95% power to replicate this effect. We aimed to over-recruit because we expected (a) attrition between session 1 and session 2 and (b) the effect size of association change on perception to decline given the increased time between association change induction and measurement of perception. Accordingly, we preregistered to collect a sample of approximately 175 participants via Prolific. One hundred eighty-three completed all measures in the first session. A total of 8 participants were excluded for failing to meet inclusion criteria (i.e., if they recommended the exclusion of their data ($n = 2$) and/or indicated that they spoke or read Chinese ($n = 6$). Twelve participants were excluded for pressing the same response on all trials of either time 1 or time 2 MP (Kurdi et al., 2024). As in all prior studies, statistical conclusions do not differ when including these participants (see Supplemental Materials). Exclusions resulted in a final sample of 164 individuals who completed session 1 (77 women, 83 men, and 4

nonbinary or other; 116 White, 28 Black, 7 Asian, 13 mixed race or other; 21 Hispanic, 143 Nonhispanic; M_{age} = 40.43, SD_{age} = 12.99).

Study 4 was nearly identical to Study 3, but with a few notable differences. As in Study 3, in session 1, participants completed a time 1 baseline measure of automatic police-threat associations (T1-MP). Immediately following the T1-MP, participants were shown a one-minute relaxation video to buffer effects of completing the T1-MP on subsequent measures. New to Study 4, following the relaxation video, participants completed a brief self-report scale of trust in the police as well as single item assessing belief that police officers are heroes. We return to the relevance of these measures in the General Discussion. Following these measures, all participants viewed 8 videos depicting high hostility arrests. These were the same videos as in the Study 3 high hostility condition. We did not assign participants to a low hostility condition in Study 4 because Study 3 did not evidence a relationship between change in associations and perception of aggression in later viewed arrests among people who initially viewed low hostility arrests (and therefore did not exhibit changes in automatic associations). Participants rated officers and civilians, respectively, within each video on aggression using the same items as in Studies 1-3. Participants next completed a time 2 post-exposure police-threat MP (T2-MP) and then reported demographic information. Instead of immediately viewing and rating the 8 moderate hostility arrest videos (used in Study 3) on officer and civilian aggression, participants were invited back approximately 18-hours later to complete this measure within a second session. The second session remained open to participants to complete between approximately 18- and 36-hours following session 1 (M = 20.77 hours). Among the 164 participants who met inclusion criteria for session 1, 141 participants completed session 2.

Results

MP trials with slow reaction times (trials greater than the third quantile plus three times the interquartile range; Tukey, 1977) were dropped from analyses (3.40% of trials on the T1-MP; 3.36% of trials on the T2-MP). Police-threat versus non-police threat evaluations were indexed in the same way as in Studies 1-3. See Table 3 for descriptive statistics among all measures.

Table 3.

Means and standard deviations for Study 4 variables

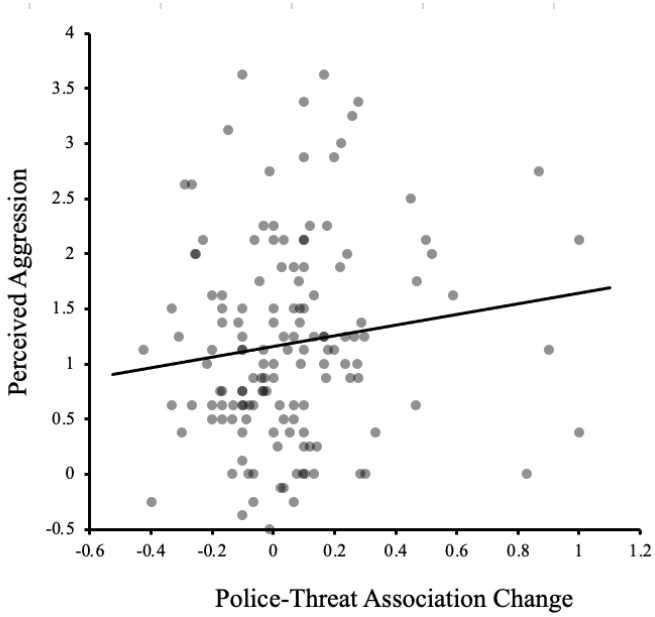
	M	***SD***
Baseline Police-Threat Association	0.51	0.25
Post-Exposure Police-Threat Association	0.57	0.29
Police Aggression Time 1	4.23	0.60
Civilian Aggression Time 1	2.70	0.86
Aggression Bias Time 1	1.53	1.13
Police Aggression Time 2	2.89	0.87
Civilian Aggression Time 2	1.70	0.68
Aggression Bias Time 2	1.18	0.90

Demonstrating that participants experienced a change in automatic associations from pre- to post-exposure, T2-MP scores (M = 0.57, $CI_{95\%}$[0.52, 0.61]) were greater than T1-MP scores (M = 0.51, $CI_{95\%}$[0.48, 0.55]), $t(163) = 2.81$, $p = .006$. Replicating Studies 2 and 3, perceptions of officer relative to civilian aggression during arrest videos were associated with heightened T2-MP scores controlling for T1-MP scores, $\beta = .21$, $b = 0.05$, $t(161) = 3.42$, $p < .001$, $\eta_p^2 = .07$.

To assess whether the association between perceived aggression induced changes in automatic associations and later perceptions of aggression persists across a longer (~1-day) delay, we regressed aggression bias scores from the session 2 (moderate) arrest videos onto the change in police-threat association (the difference of T2-MP score minus T1-MP score) from session 1. This revealed a nonsignificant association in the hypothesized direction, $\beta = .13$, $b =$

0.48, $t(141) = 1.59$, $p = .113$, $\eta_p^2 = .02$, see Figure 4. This effect was unmoderated by time elapsed between sessions ($p = .578$).

Figure 4. *The relationship between change in police-threat association and later perception (~1-day) of officer relative to civilian aggression in Study 4.*



Discussion

Viewing hostile arrests strengthens automatic police-threat associations, and this pattern is explained by perceptions of aggression. Study 4 provides a preregistered replication of the effects initially established in Study 2 and confirmed in Study 3. Additionally, the findings of Study 4 offer insights into the nature of the causal force of automatic association change over time. Specifically, we found that the causal influence of automatic associations on aggression perceptions diminishes to statistical nonsignificance when arrests are viewed after an approximately 1-day delay. Meaning, the mutually reinforcing dynamic between police-threat associations and perceptions of police aggression holds primarily in short-term iterative contexts. Despite this, we caution against concluding that social perception induced automatic association

changes cannot operate beyond the immediate experimental session. As we describe in further detail in the General Discussion, long-term changes in automatic police-threat associations may occur via several repeated short-term iterations of the mutually reinforcing patterns laid out in our single session Study 3.

Moreover, although we observed a statistically nonsignificant relationship between changes in automatic associations and delayed perceptions of aggression, this effect was in the hypothesized direction. A Monte Carlo power simulation (reviewed at length in the General Discussion) indicated that our current design was likely underpowered to reliably detect this effect. Indeed, it is likely that we underestimated the decay rate of change in automatic associations. If, for example, the follow up session had been conducted at shorter intervals, say 4, 8, or 12 hours post manipulation (or, as we discuss later, had we collected data from 480 participants), we may have found a statistically significant effect. This limitation highlights the potentially transient nature of experimentally induced automatic association changes and the many unknowns that likely govern the stability and decay of changes in automatic association, including, for example, the nature of the association, magnitude of initial change, the time elapsed between association induction and its measurement, and variable decay rates of the change over time. Consistent with Forscher and colleagues (2019), our findings underscore the modest and often transient nature of experimentally induced changes in automatic association, which emphasizes the need for adequately powered future research to systematically explore how long association-perception dynamics meaningfully persist. We expand on these considerations in the General Discussion.

General Discussion

Encounters among police officers and civilians are laden with potential violence. Both access to visual accounts of hostile policing and evidence that people automatically associate the police with threat are growing. The current work revealed an iterative and (short-term) mutually reinforcing dynamic through which these outcomes are related. Study 1 revealed that existing police-threat associations are related to perceptions of officers' and civilians' behaviors during later viewed arrests. Those who more strongly associate the police with threat more readily perceive officers relative to civilians as behaving aggressively. Studies 2-4 demonstrated that these perceptions in turn reinforce the police-threat association. Viewing more versus less hostile arrests led to greater perceptions of officer relative to civilian aggression and perceptions of aggression strengthened the police-threat association. Study 3 found that increases in the police-threat association strengthen peoples' subsequent perceptions of officer relative to civilian aggression in later viewed arrests. Study 4 replicated the findings of Studies 2 and 3 while demonstrating that the implications of change in police-threat associations for later perceptions of arrest diminishes across a 1-day delay. In the remainder of the General Discussion, we situate our findings within the literatures of automatic social cognition and police-community relations.

Implications for Automatic Social Cognition

Our findings align with existing evidence that highly trait diagnostic information can produce rapid, though potentially short term, shifts in automatic associations at the individual level (Cone & Ferguson, 2015; Kurdi & Charlesworth, 2023; Shen et al., 2020). Verbal descriptions of diagnostic behavior (e.g., highly negative actions) induce quick negative shifts in automatic impressions toward novel individual targets. Extending these findings, we here show that visual depictions of diagnostic behavior (i.e., physically aggressive actions) can likewise

rapidly produce shifts in automatic associations. Critically, our results differ from prior work by demonstrating rapid shifts in automatic associations towards established groups for whom participants already possess prior attitudes. Whereas previous research primarily examined automatic first impressions toward unfamiliar individuals (Cone & Ferguson, 2015; Shen et al., 2020), our studies uniquely illustrate that automatic associations toward familiar groups (i.e., here, police officers) are similarly susceptible to rapid shifts based on diagnostic behavioral information.

Another strength of our work is the demonstration of automatic association shifts via within-person change. We did so by combining experimental manipulations with pre- and post-manipulation measures of automatic associations, rather than relying solely on post-manipulation measures compared to a baseline control group. We established these effects in Study 2 and replicated them in Study 3. Study 4 provided further additional confirmatory evidence of perception driven within-person association change. Together, our findings provide insights into a social cognitive mechanism underlying within-person automatic association shifts. Here, perceptions of officers (relative to civilians) as behaving aggressively explained shifts in automatic police-threat associations following exposure to arrest videos.

Specifically, across three studies, we consistently found that experiences occurring between pre-post measures of automatic associations produced shifts in those associations linearly proportionate with participant's subjective perceptions of the intervening experience. The robustness and replicability of these findings reinforce contemporary theories of automatic attitude change (e.g., Forscher et al., 2019; Kurdi & Charlesworth, 2023), and likewise inform ongoing discussions on the reliability, sensitivity, and utility of implicit measurement methods (Corneille & Gawronski, 2025; Moreau, 2025), particularly misattribution procedures (Payne et

al., 2005; Payne & Lundberg, 2014). This facet of our work exemplifies the systematic and reliable nature of within-person association change driven by brief, socially meaningful experiences.

As alluded to in the Introduction, empirical evidence for the causal force of automatic attitude and attitude change on subsequent judgments and decision remains mixed (Forscher et al., 2019). This inconsistency likely stems from numerous factors that vary nonsystematically across studies. The current findings demonstrate a short-term perceptual consequence of shifts in automatic associations: aggression induced changes in police threat-associations shaped subsequent perceptions of police relative to civilians as more aggressive, at least in the short-term (i.e., within a single experimental session). In other words, we provide experimental evidence that automatic attitude shifts can influence downstream judgements in the short term, which speaks directly to theories of attitude-motivated perception and automatic social cognition (Dunning & Balcetis, 2013; Fazio, 1990; Olson & Fazio, 2004).

However, Study 4 suggests that the effect of automatic association change on downstream perception does not remain reliable after a roughly 1-day delay. While this null finding might appear to challenge the stability of automatic association change (and its causal force), we caution against overinterpreting its theoretical or practical significance. Although our effect was not statistically significant, and we will therefore not oversell its significance, the trend was in the expected direction. Our sample size (targeted at 175) was conservative relative to the power needed for the effect observed in Study 3 ($N = 99$), but clearly underestimated the extent to which the association shift would decay. To refine this point, we conducted a Monte Carlo power simulation to estimate the power to detect the observed effect of police-threat association change on perceptions of aggression in arrests viewed approximately 1-day later.

Using Study 4's observed parameter estimates, specifically the regression intercept and slope, residual variance, and the standard deviation of the predictor, we generated 1,000 synthetic datasets by sampling from a normal distribution with matched parameters. For each simulated dataset, we fit the same linear model (i.e., perceptions of aggression regressed onto change in automatic associations) and recorded whether the slope reached statistical significance at $\alpha = .05$, with power defined as the proportion of simulations yielding a significant effect. This indicated that our detected effect with $N = 164$ had only 34% power and would require approximately 480 participants to achieve 80% power (assuming all parameters remained proportional). Meaning, although we were underpowered to detect the small(er than expected) effect, that effect may have been detected had we been able to gather a larger sample. Thus, while our test was underpowered to detect a smaller-than-expected effect, this does not mean the effect was absent—only that it may have decayed more than anticipated.

Notably, we found that the post-exposure (time 2) police-threat associations predicted post-delay perception of aggression ($r = .22$, $p = .006$), whereas the baseline (time 1) associations did not ($r = .11$, $p = .169$). And the relationship between post-exposure (time 2) police-threat associations and post-delay perception remained significant when controlling for baseline (time 1) associations ($\beta = .21$, $b = 0.66$, $t(141) = 2.07$, $p = .040$). This suggests that the association was still meaningfully updated and remained predictive approximately 24 hours later, even if the magnitude of change had diminished. Had we measured the causal force of association change on perception at shorter intervals, say an hour, four hours, or some other number less than twenty-four hours later, we might have observed that the shift in automatic associations predicted perception. Discussed below, this highlights a critical avenue for future research in systematically mapping the time course and decay rate of automatic association change.

This nuance has clear implications for both theoretical models of automatic social cognition and real-world applications. Decisions and judgments made shortly after an association shift are more likely to be influenced by that shift. This supports dynamic models of social cognition in which attitudes are not static but are shaped and reshaped by salient, meaningful social information (Lai et al, 2016; Powell & Fazio, 1984). Durable changes in automatic attitudes, perhaps particularly toward groups like the police for which people likely have preexisting stable attitudes, may not result from a single exposure but from repeated encounters that accumulate over time. Delay between exposure and response, as well as the frequency and intensity of exposure, are factors known to moderate long-term attitude change (Bornstein, 1989). Put differently, the long-term causal force of automatic attitude change may emerge through repeated iterations of the mutually reinforcing processes demonstrated here: perception influences associations, and those associations, in turn, shape future perception.

Consider how this may play out in a real-world setting. American adults now spend on average nearly two and a half hours a day on social media and increasingly rely on social media sites for news (Kumar, 2024; McClain, 2024). Content depicting police violence is widely posted and circulated, with nearly 90% of Americans reporting that they have viewed videos of police use of force on social media (Horowitz et al., 2023). While Study 4 implemented a 1-day delay after a single exposure, real-world exposure patterns suggest that when encountering a single instance of violent police media, people are more likely to then encounter violent police media more frequently and across several contexts. In such conditions, each exposure may incrementally shift automatic police-threat associations and, over time, produce stable and consequential changes in perception and judgment.

In this light, future work could explore long-term automatic attitude change as a function of multiple (short-term) mutually reinforcing iterations over time. We observed a mutually reinforcing pattern between automatic associations and social perceptions across two short-term (i.e., a single session) iterations. Indeed, in Study 3, automatic associations shifted in response to perceptions of aggression and those updated associations then iteratively shaped subsequent perceptions. A reasonable next step would be to extend this design across both shorter and longer time spans and multiple exposures, simulating real-world media consumption patterns. For example, researchers could expose participants to a series of aggression-laden arrest videos spaced across several hours and/or days, measuring association strength and perception at each timepoint. Such work could clarify whether repeated exposure produces cumulative, lasting changes in automatic associations and whether those changes meaningfully alter how individuals interpret ambiguous social encounters over time. Of course, the strengthening of the police-threat association is unlikely to continue indefinitely. At some point, repeated exposure may yield only diminishing returns, resulting in incremental increases or merely maintaining an already heightened association. Future research can test the boundary conditions and potential ceiling of the iterative dynamic outlined in the current work.

Implications for Police-Community Relations

Our findings contribute to a growing body of working demonstrating that media portrayals of police aggression can sow public discontent with policing institutions. Viewing such videos has been linked to more negative evaluations of police (Campbell & Valera, 2020; Graziano, 2019), and favorability ratings of police departments reliably decline following publicized incidents of aggressive arrests (Weitzer, 2002; Weitzer & Tuch, 2005). Our work reflects these patterns while extending them by demonstrating that increasing discontent with the

police (e.g., as indexed by strengthened police-threat association) may not only result from perceptions of police aggression but also feed back into them. That is, discontent may promote heightened perceptions of police aggression, which in turn may reinforce negative associations in a mutually reinforcing cycle.

Our demonstration of this process at the level of automatic associations underscores its potential relevance for real-world behavioral consequences of witnessing police aggression. One implication is that people with stronger police-threat associations may be more likely to exhibit automatic defensive behaviors to police officers (Olivett & March, 2024). Strengthening the police-threat association via repeated media exposure to aggressive policing could therefore strengthen these associations over time, increasing the likelihood of defensive behaviors among civilians when encountering officers. Such behaviors may then be interpreted by officers as noncompliant or aggressive and thus provoke reciprocal aggression among officers. Future work should test whether perceptions of police aggression increase the likelihood of defensive responses to officers.

Strengths and Limitations

Strengths and limitations regarding internal and statistical validity as well as generalizability are systematically outlined in Table 3. Notably, our studies focused on policing in an U.S. context. Our samples exclusively consisted of American adults. Therefore, our findings cannot necessarily be generalized to police-community relations in non-U.S. contexts.

An incisive reader (or reviewer) may question whether responses on our MPs genuinely reflect “police-threat” associations, as we claim. We argue that police-threat associations reflect a view of officers as sources of threat to the self. Yet, one might argue that the MPs capture a police-with-danger association, as officers are often seen as heroic agents who frequently

encounter danger, rather than police-as-dangerous. While this alternative interpretation is intuitively plausible, both the current findings and previous work utilizing the same MPs provide strong support the "police-as-threat" conceptualization. First, previous studies have shown that that stronger police-threat associations indexed with the same MP predicted lower self-reported trust in the police (Olivett & March, 2021). Moreover, participants in those studies consistently associated police with danger more strongly than they did firefighters, despite firefighters' profession inherently involving frequent exposure to dangerous situations. If participants simply associated danger with occupational risk, firefighters should have elicited stronger danger associations, yet they did not. Second, the current work shows that perceptions of aggression by *officers, not civilians,* is the key driver of increases in police-threat associations. If MPs were capturing a belief that officers are agents who frequently encounter danger, then perceiving civilian aggression (i.e., danger to officers) should have increased MP scores. Instead, we find the opposite pattern, suggesting that participants interpret aggressive officers as themselves being *sources* of threat. Finally, data from Study 4 further support this interpretation. Recall that we included both a self-report measure of trust in police and a single item assessing view of police officers as heroes. Baseline police-threat associations correlated strongly with both diminished trust in police (r = -.50, p < .001) and diminished view of officers as heroic (r -.43, p < .001). These associations are inconsistent with the idea that our MP indexes police as brave protectors who face danger. Rather, they align with our view that the MP indexes police as entities posing danger.

Conclusions

Three studies converge to find that viewing aggressive policing and automatic associations towards the police are iteratively related and mutually reinforcing in the short-term.

When viewing arrests, people perceive officer and civilian aggression through the lens of their preexisting police-threat associations. In turn, these perceptions reinforce the police-threat association.

Table 3.
Assessment of Strengths and Limitations

Dimension	Assessment
Internal Validity	
Is the phenomenon assessed with experimental methods?	Yes (Studies 2-3)
Is the phenomenon assessed with longitudinal methods?	Yes (Studies 2-4)
Were manipulations validated?	Yes, we used previously piloted videos and gathered our own manipulation checks.
Was mechanism examined?	Yes (Studies 2-3)
Statistical Validity	
Was statistical power adequate?	Post hoc sensitivity analyses revealed observed effect sizes greater than the predicted minimal detectable effect sizes assuming 80% power for the critical effect Study 1, the conditional effects in Studies 2 and 3, and the interaction in Study 3. Power considerations for the indirect mediation effects in Studies 2 and 3 are unaddressed.
Was dependent measure reliability established within this work or elsewhere?	Yes, previous work speaks to the reliability of MPs (Payne & Lundberg, 2014)
Were variables assessed for normal distribution?	Visual inspection indicated that all DVs were normally distributed.
Generalizability	
Were different experimental manipulations used?	No, we relied on the same experimental methods in Studies 2-3.
Were different dependent measures used?	No, we used MPs to measure association and self-reports to measure perceptions. We did not use a direct measure of associations. We did not measure behavior.
Were the methods assessed in a field setting	No, our research was conducted via online data collection platforms
Were the methods artificial?	Yes, we measured automatic associations which necessitates the use of indirect measures.
Are the results generalizable across populations?	Our samples are diverse in age, race, and gender, but not nationality
Are the results generalizable across time, historic period?	Not tested. We suspect that these results are contingent on social context, and therefore may not generalize across time.

References

Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64*(2), 123—152.

Campbell, F., & Valera, P. (2020). "The only thing new is the cameras": A study of U.S. college Students' perceptions of police violence on social media. *Journal of Black Studies, 51*, 654–670.

Cone, J. & Ferguson M. J. (2015). He did what? The role of diagnosticity in revising implicit evaluations. *Journal of Personality and Social Psychology, 108*(1), 37—57.

Corneille, O. & Gawronski, B. (2024). Self-reports are better measurement instruments than implicit measures. *Nature Reviews Psychology, 3,* 835—846.

Corneille, O. & Hütter, M. (2020). Implicit? What do you mean? A comprehensive review of the delusive implicitness construct in attitude research. *Personality and Social Psychology Review, 24*(3), 212—232.

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.

Correll, J., Park, B., Judd, C. M., Wittenbrink, B., Sadler, M. S., & Kreese, T. (2007). Across the thin blue line: Police officers and racial bias in the decision to shoot. *Journal of Personality and Social Psychology, 92,* 1006–1023.

Correll, J., Wittenbrink, B., Crawford, M. T., & Sadler, M. S. (2015). Stereotypic vision: How stereotypes disambiguate visual stimuli. *Journal of Personality and Social Psychology, 108,* 219–233.

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: A social cognitive neuroscience perspective. *Trends in cognitive sciences*, *11*(3), 97-104.

Dunning, D., & Balcetis, E. (2013). Wishful seeing: How preferences shape visual perception. *Current Directions in Psychological Science, 22*(1), 33–37. https://doi.org/10.1177/0963721412463693

Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of Blacks. *Journal of Personality and Social Psychology, 34*(4), 590—598.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes.* Harcourt Brace Jovanovich College Publishers.

Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology, 23,* 75—109. *of Social Psychology, 163*(4), 439—458.

Fazio, R. H. (1995). Attitudes as object evaluations. In R. Petty & J. A. Krosnick (Eds.), Attitude strength: Antecedents and consequences (pp. 247-285). Mahwah, NJ: Erlbaum.

Fazio, R. H. (2001). On the automatic activation of associated evaluations: An overview. *Cognition & Emotion, 15*(2), 115—141.

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., & Devine, P. G. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology: Attitudes and Social Cognition, 117*(3), 522—559.

Gawronski, B., Geschke, D., & Banse, R. (2003). Implicit bias in impression formation:

Associations influence the construal of individuating information. *European Journal of Social Psychology, 33,* 573—589.

Graeff, E., Stempeck, M. & Ethan Zuckerman, E. (2014). The battle for Trayvon Martin: Mapping a media controversy online and offline. *First Monday, 19.*

Graziano, L. M. (2019). News media and perceptions of the police. *Policing: An International Journal, 42*(2), 209–225.

Greenwald, A.G., McGhee, D.E., & Schwartz, J.L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, *74*(6):1464—1480.

Horowitz, J. M. (2023, June 14). *2. The impact of videos of police violence against Black people.* Pew Research Center. https://www.pewresearch.org/social-trends/2023/06/14/the-impact-of-videos-of-police-violence-against-black-people/

Imhoff, R., Schmidt, A. F., Bernhardt, J., Dierksmeier, A., & Banse, R. (2011). An inkblot for sexual preference: A semantic variant of the affect misattribution procedure. *Cognition and Emotion*, 25, 676–690.

Kumar, N. (2024). *Average time spent on social media per day.* DemandSage. https://www.demandsage.com/average-time-spent-on-social-media/

Kurdi, B., & Charlesworth, T. E. S. (2023). A 3D framework of implicit attitude change. *Trends in Cognitive Sciences, 27*(8), 745–758. https://doi.org/10.1016/j.tics.2023.05.009

Kurdi, B., Melnikoff, D. E., Hannay, J. W., Korkmaz, A., Lee, K. M., Ritchie, E., Surdel, N., Vuletich, H. A., Yang, X., Payne, B. K., & Ferguson, M. J. (2024). Testing the automaticity features of the Affect Misattribution Procedure: The roles of awareness and

intentionality. Behavior Research Methods, 56(4), 3161–3194. https://doi.org/10.3758/s13428-023-02291-2

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., Marshburn, C. K., Simon, S., Blanchar, J. C., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhaas, F. M. H., Burns, M., . . . Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General, 145*(8), 1001–1016. https://doi.org/10.1037/xge0000179

Logan, M. H. (2016). Police violence: a two-way street. *Violence and Gender*, *3*, 3-6.

March, D. S., Gaertner, L., & Olson, M. A. (2017). In harm's way: On preferential response to threatening stimuli. *Personality and Social Psychology Bulletin, 43*(11), 1519–1529.

March, D. S., Gaertner, L., & Olson, M. A. (2018a). On the prioritized processing of threat in a dual implicit process model of evaluation. Psychological Inquiry, 29, 1–13.

March, D. S., Gaertner, L., & Olson, M. A. (2022). On the automatic nature of threat: Physiological and evaluative reactions to survival-threats outside conscious perception. *Affective Science, 3*(1), 135–144.

McClain, C. (2024, August 20). *About half of TikTok users under 30 say they use it to keep up with politics, news.* Pew Research Center. https://www.pewresearch.org/short-reads/2024/08/20/about-half-of-tiktok-users-under-30-say-they-use-it-to-keep-up-with-politics-news/

Mobbs, D., Headley, D. B., Ding, W., & Dayan, P. (2020). Space, time, and fear: Survival computations along defensive circuits. *Trends in Cognitive Sciences, 24*(3), 228–241.

Moreau, D. (2025). Temporal stability in measurements matters only for stable constructs.

Nature Reviews Psychology, 4, 361.

Olivett, V. (2025, May 14). Police-Threat Associations and Perceptions of Aggressive Policing. Retrieved from osf.io/rfn9h

Olivett, V. J. & March, D. S. (2021). White civilians' implicit danger evaluation of police officers underlies explicit perception of the police. *Cognitive Research: Principles and Implications, 6*(1), Article 81.

Olivett, V. J. & March, D. S. (2024). The civilian's dilemma: Civilians exhibit automatic defensive responses to the police. *Journal of Personality and Social Psychology, 126*(5), 841—851.

Olson, M. A. & Fazio, R. H. (2004). Trait inferences as a function of automatically activated racial attitudes and motivation to control prejudiced reactions. *Basic and Applied Social Psychology, 26*(1), 1—11.

Olsson, A., Nearing, K. I., & Phelps, E. A. (2007). Learning fears by observing others: the neural systems of social fear transmission. Social cognitive and affective neuroscience, 2(1), 3-11.

Olsson, A. & Phelps, E. A. (2007). Social learning of fear. *Nature Neuroscience*, 10(9), 1095—1102.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*(3), 277–293.

Payne, B. K. & Correll, J. (2020). Chapter one: Race, weapons, and the perception of threat. *Advances in Experimental Social Psychology, 62,* 1—50.

Payne, B. K. & Lundberg, K. (2014). The affect misattribution procedure: Ten years of evidence

on reliability, validity, and mechanisms. *Social and Personality Psychology Compass, 8*(12), 672—686.

Powell, M. C., & Fazio, R. H. (1984). Attitude accessibility as a function of repeated attitudinal expression. *Personality and Social Psychology Bulletin, 10*(1), 139–148. https://doi.org/10.1177/0146167284101016

Sargent, R. H. & Newman, L. S. (2020). An investigation of civilian implicit attitudes toward police officers. *Journal of Police and Criminal Psychology, 35,* 360—376.

Sargent, R. H. & Newman, L. S. (2023). An investigation of emotional and evaluative implicit associations with police using four versions of the implicit association test. *The Journal of Social Psychology, 163*(4), 439—458.

Shen, X., Mann .T. C., & Ferguson, M. J. (2020). Beware a dishonest face?: Updating face-based implicit impressions using diagnostic behavioral information. *Journal of Experimental Social Psychology, 86,* 103888.

Washington Post (2024, October 10). Police shootings database. Washington Post. Retrieved October 22, 2024 from https://www.washingtonpost.com/graphics/investigations/police-shootings-database/

Weitzer, R. (2002). Incidents of police misconduct and public opinion. *Journal of Criminal Justice, 30*(5), 397—408.

Weitzer, R. & Tuch, S. A. (2005). Racial biased policing: Determinants of citizen perceptions. *Social Forces, 83*(3), 1009—1030.