

## BRIEF REPORT

# Psychometric properties of the Buss–Perry Aggression Questionnaire-short form among law enforcement officers

Michael Christopher<sup>1</sup>  | Marissa Ferry<sup>1</sup>  | Akeesha Simmons<sup>1</sup>  |  
Alicia Vasquez<sup>1</sup>  | Brooke Reynolds<sup>1</sup>  | Daniel Grupe<sup>2</sup> 

<sup>1</sup>School of Graduate Psychology, Pacific University, Hillsboro, Oregon, USA

<sup>2</sup>Center for Healthy Minds, University of Wisconsin-Madison, Madison, Wisconsin, USA

**Correspondence**

Michael Christopher, 190 SE 8th Ave, Hillsboro, OR 97123, USA.

Email: [mchristopher@pacificu.edu](mailto:mchristopher@pacificu.edu)

**Funding information**

National Institutes of Health, Grant/Award Numbers: R21AT008854, R01AT009841

**Abstract**

Repeatedly capturing national headlines, excessive law enforcement officer (LEO) use of force in critical incident encounters is one of the most divisive human rights issues in the United States. Valid and reliable measures of potential precursors to LEO excessive use of force, such as aggression, are needed. The Buss–Perry Aggression Questionnaire-short form (BPAQ-SF) is a validated measure of aggression across various populations; however, evaluation of this easily administered measure in high-stress, frontline populations such as LEOs is limited. The primary goal of this study was to evaluate the validity, reliability, and sensitivity to change the BPAQ-SF in a sample of LEOs. A confirmatory factor analysis suggested that the hierarchical solution provides a mixed fit to the data:  $SB\chi^2_{(25.84)} = 62.50$ ,  $p = .0001$ ; comparative fit index = .94, non-normed fit index = .92, root mean square error of approximation = .19 (90% confidence interval = .17–.21), standardized root mean squared residual = .08. The BPAQ-SF demonstrated good internal consistency ( $\alpha = .84$ ) and test–retest reliability ( $r = .86$ ), correlations in the expected direction with predictors of and buffers against aggression, and sensitivity to change among LEOs who participated in an intervention targeting aggression. Results support and extend previous findings suggesting that the BPAQ-SF is a valid and reliable measure of aggression among LEOs.

**KEYWORDS**

aggression, health, Law enforcement, mindfulness, psychometrics

## 1 | INTRODUCTION

Police officers' work entails extreme exposure to chronic and acute stressors, elevating risk for excessive use of force, such as unjustified shootings, severe beatings, and fatal chokings (Hine et al., 2018; Kurtz & Hughes, 2021; Verhage et al., 2018). Repeatedly capturing national headlines, excessive law enforcement officer (LEO) use of force in critical incident encounters is one of the most divisive human rights issues in the United States (U.S.) (Cobbina-Dungy &

Jones-Brown, 2023; McLean et al., 2022; Mourtgos & Adams, 2020). The Bureau of Justice Statistics estimates 53.8 million U.S. residents age 16 or older had one or more face-to-face contacts with police in 2020, the most recent year data were available (Tapp & Davis, 2022). Among them, an estimated 1.1 million experienced threat or use of force by police, including hitting or kicking, pepper spray, electroshock weapon, pointing a gun, or other force during the most recent contact, and 46% of those described it as excessive (Tapp & Davis, 2022).

Excessive aggression in policing can result in numerous negative outcomes, including physical injuries, trauma, and loss of life for both individuals involved and community members. Overly aggressive policing also erodes public trust and confidence in law enforcement, creating a widening divide between police and the citizens they serve (Haldipur, 2020; Nix et al., 2020), and hindering effective crime prevention efforts (Boxer et al., 2021; Hawkins, 2022). Victims of overly aggressive policing endorse elevated depression, anxiety, and trauma, particularly among members of racially minoritized groups (Alang et al., 2021; Bowleg et al., 2020; McLeod et al., 2020). In terms of economic cost, authors of a *Washington Post* (Alexander et al., 2022) article reported that the largest 25 police and sheriff's departments in the United States spent more than \$3.2 billion to resolve claims of police misconduct from 2012 to 2022. Reductions in overly aggressive policing have been linked to enhanced officer legitimacy and credibility, community cooperation, and collaborative crime prevention strategies (Wood et al., 2020).

Prior research highlights several key predictors of LEO aggression and potential mitigators of its impact. Repeated exposure to acute occupational stress increases the likelihood of excessive LEO aggression (Hine et al., 2018; Verhage et al., 2018), and LEOs who engage in hazardous drinking are four times more likely to report physical violence than officers who do not use alcohol (Oehme et al., 2012). High trauma-exposed LEOs report higher rates of job dissatisfaction and burnout (Juczyński & Ogińska-Bulik, 2022; Mumford et al., 2021) and are more likely to use overly aggressive policing tactics (Kurtz & Hughes, 2021). Burnout (Queirós et al., 2013) and sleep difficulties (Rajaratnam et al., 2011) exacerbate stress and trauma reactions and are associated with increased aggression among LEOs. Alternatively, contemplative practices such as mindfulness (Fix & Fix, 2013; Singh et al., 2014) and self-compassion (Morley, 2015) have been shown to reduce aggression in non-LEO populations, and there is preliminary evidence suggesting mindfulness may reduce aggression among LEOs (Khatib et al., 2022; Ribeiro et al., 2020).

The ability to accurately assess potentially overly aggressive LEOs in research and on-the-job contexts can be key to providing preventive strategies to mitigate excessive aggression. For example, screening for new police hiring often utilizes personality assessments to identify potentially problematic candidates, with the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; Butcher, 2010) being perhaps the most frequently used instrument (Cochrane et al., 2003; Loya et al., 2022). The MMPI-2 was designed to screen for possible psychopathology, not to predict aggressive behavior or other indicators of police performance. It does not contain a specific index of aggression, which has led researchers to create an aggression amalgam from various MMPI-2 items (Hargrave et al., 1988). A study of over 1000 police officers failed to identify a link between this measure and police acts of aggression (Davis et al., 2004), suggesting there may be practical utility for a brief, validated, theoretically derived, specific measure of aggression.

A commonly used self-report measure of aggression is the Buss-Perry Aggression Questionnaire (BPAQ; Buss & Perry, 1992). Aggression, as operationalized by the BPAQ, is a multidimensional

construct that assesses four behavioral subareas: physical aggression, verbal aggression, anger, and hostility. Physical aggression refers to engaging in aggressive behaviors that cause bodily harm or damage. Verbal aggression involves the use of harsh and hostile language to attack others verbally. Anger reflects the emotional arousal associated with the propensity to react aggressively in frustrating situations. Hostility represents a more enduring negative attitude and resentment toward others. Although the BPAQ has generally demonstrated adequate validity and reliability (Harris, 1997), Bryant and Smith (2001) developed a short-form version of the BPAQ to improve the psychometrics and reduce participant burden. The BPAQ-SF has consistently demonstrated excellent psychometric properties across various U.S. populations (Diamond & Magaletta, 2006; Webster et al., 2014).

Valid and reliable assessment of aggression in policing is crucial to advance understanding of its precursors and to assess the effectiveness of interventions to reduce its impact on the community. Despite the widespread use of the BPAQ-SF, only one published study to date has examined the validity and reliability of the BPAQ-SF among LEOs in the United States. Greenberg et al. (2003) found evidence supporting the construct validity of the BPAQ-SF, including a well-fitting hierarchical structure (with physical aggression, verbal aggression, anger, and hostility loading onto an overall aggression factor) and adequate reliability. Greenberg et al. (2003) found that the BPAQ-SF correlated with departmental measures of citizen complaints for LEO physical and verbal aggression, as well as the number of shooting incidents in which the officers were involved. Additionally, the study benefitted from a relatively large LEO sample ( $n = 250$ ). Despite these strengths, the Greenberg et al. study has yet to be replicated in an LEO sample. In addition, this study was conducted over 20 years ago, during which time the landscape of policing in the United States has seen dramatic changes, including increased militarization and a stream of high-profile police killings and excessive use-of-force incidents.

Additional research is needed to further examine the BPAQ-SF psychometrics among LEOs and, crucially, to examine the impact of interventions targeting excessive aggression in U.S. LEOs. Despite limitations to self-reported measures, including social desirability effects, the BPAQ-SF is a potentially useful measure in workplace and research contexts among LEOs given the high-stakes nature of overly aggressive policing, and the desire to identify interventions that may reduce it. Assessing BPAQ-SF psychometrics in randomized clinical trial research is an important step in optimizing the measurement of self-reported aggression, and enables future research investigating whether changes in self-reported aggression predict objective and behavioral outcomes of aggression. Moreover, establishing convergent validity of self-reported aggression in relation to factors known to be important for predicting aggressive behavior, or for mitigating its impact, can help overcome social desirability effects and other limitations of assessing self-reported aggression in isolation. Therefore, the two goals of the current study were to (1) replicate Greenberg et al.'s (2003) evaluation of the BPAQ-SF psychometrics in a sample of U.S. LEOs, and (2) extend the

field by providing novel evidence of BPAQ-SF test–retest reliability, convergent validity, and sensitivity to change in an intervention designed to reduce aggression.

## 2 | METHODS

### 2.1 | Participants

LEOs were recruited from urban police departments in the Pacific Northwest, the Southwest, and the Upper Midwest comprising a total of 15 law enforcement agencies. To be eligible, participants needed to be sworn LEOs. Data were combined from two separate randomized clinical trials assessing the impact of mindfulness-based resilience training (MBRT) on aggression and health outcomes among LEOs. Study 1 included  $n = 61$  participants at baseline (see Christopher et al., 2018 for study details) and Study 2 included  $n = 109$  participants at baseline (see Christopher et al. (2023) for study details) for a total baseline of  $n = 170$ . In Study 1, after baseline assessment participants were randomized to an 8-week MBRT ( $n = 31$ ) or no-intervention control (NIC,  $n = 30$ ) condition. Of the 31 MBRT participants, 24 provided post-intervention data. In Study 2, after baseline assessment a subset of participants were randomized to MBRT ( $n = 33$ ), active control ( $n = 24$ ), or NIC ( $n = 16$ ). Of the 33 MBRT participants, 31 provided post-intervention data. BPAQ-SF sensitivity to change pre-to-post MBRT was assessed in the pooled subsample of MBRT participants in Studies 1 ( $n = 24$ ) and 2 ( $n = 31$ ) for a total  $n = 55$ . BPAQ-SF test–retest reliability was assessed in the pooled subsample of NIC participants who provided baseline and post data in Studies 1 ( $n = 26$ ) and 2 ( $n = 14$ ) for a total  $n = 40$ .

At baseline, mean participant age was 41.17 years ( $SD = 8.54$ ) and mean years as an LEO was 14.01 ( $SD = 8.47$ ). Twenty-four percent ( $n = 40$ ) of participants identified as female and 76% as male ( $n = 130$ ). In terms of race, 84% ( $n = 143$ ) of participants identified as White American, 5% ( $n = 9$ ) as Black/African American, 4% ( $n = 7$ ) as Asian American, 2% ( $n = 3$ ) as Native American, and 5% ( $n = 8$ ) identified as other, and in terms of ethnicity 26% ( $n = 45$ ) identified as Hispanic or Latino and 74% ( $n = 125$ ) identified as Not Hispanic or Latino.

### 2.2 | Procedures and measures

The Pacific University IRB approved all procedures in Studies 1 and 2. After providing written informed consent, participants completed all self-report validated measures below online via Qualtrics. Unless otherwise noted, the following validated self-report measures were used in Studies 1 and 2.

#### 2.2.1 | Measures

The Buss–Perry Aggression Questionnaire-short form (BPAQ-SF; Bryant & Smith, 2001) is a 12-item scale of aggression derived

from the 29-item BPAQ (Buss & Perry, 1992). The BPAQ-SF was developed to assess four factors of aggression: physical aggression, verbal aggression, anger, and hostility. Items on the BPAQ-SF range from 1 *very unlike me* to 5 *very like me*, and are summed to create a total with higher scores indicating greater aggression (see Supporting Information File 1 for BPAQ-SF items). The BPAQ-SF has demonstrated good internal consistency and strong convergent and discriminant validity in non-LEO samples (Diamond & Magaletta, 2006).

Patient Reported Outcomes Measurement Information System (PROMIS® v1.0) short form versions were used to assess sleep disturbance (six items), alcohol use (seven items), and anger (six items). PROMIS measures have variable ranges (sleep disturbance [32–76], alcohol use [39–77], and anger [39–83]), with higher scores indicating a higher rate of the measured outcome. Scores were converted to standardized  $T$  scores ( $M = 50$ ;  $SD = 10$ ), centered on the general U.S. population mean. These short forms have demonstrated acceptable internal consistency and correlations with expected legacy measures (Pilkonis et al., 2011, 2016; Yu et al., 2012).

The PTSD Checklist for DSM-5 (PCL-5; Weathers et al., 2013) is a 20-item measure based on the DSM-5 symptoms of PTSD. Respondents report how much they were bothered by each symptom over the past month using a 5-point Likert scale ranging from 0 to 4. Items are summed to create a total score, with higher scores indicating greater PTSD symptoms. The PCL-5 has demonstrated good internal consistency and strong convergent and discriminant validity (Roberts et al., 2021). The PCL-5 was only administered to participants in Study 2.

The Perceived Stress Scale-10 (PSS-10; Cohen, 1988) is a 10-item measure used to assess the degree to which situations in life are perceived as stressful. Items are designed to capture how unpredictable, uncontrollable, and overloaded participants find their lives. Items are rated on a 5-point Likert-type scale ranging from 0 to 4. Items are summed to create a total score, with higher scores indicating greater stress. The PSS-10 has shown good internal consistency and demonstrated expected correlations with a variety of constructs (Taylor, 2015). The PSS-10 was only administered to participants in Study 2.

The Oldenburg Burnout Inventory (OLBI; Demerouti et al., 2003; Halbesleben & Demerouti, 2005) is a 16-item measure of burnout that assesses exhaustion and disengagement from work. The OLBI has acceptable internal consistency, factorial validity, and expected correlations with other constructs (Demerouti et al., 2010). Items are rated on a 4-point Likert-type scale ranging from 1 to 4. The OLBI ranges from 1 to 4, with higher scores indicating greater burnout.

The Five Facet Mindfulness Questionnaire-short form (FFMQ-SF; Bohlmeijer et al., 2011), a 24-item version of the FFMQ (Baer et al., 2006), assesses dispositional tendency to be mindful in daily life. The observe and describe facets of the scale have demonstrated weaker psychometric properties and issues with novice and non-meditating samples (Christopher et al., 2012; de Bruin et al., 2012). Thus, the current study used three of the five facets—acting with

awareness, nonjudging of experience, and nonreactivity to inner experience. Each facet has five items, resulting in a 15-item scale. Items are rated on a 5-point Likert-type scale and summed with higher scores indicating greater dispositional mindfulness. The 3-facet, 15-item FFMQ has demonstrated good internal consistency and expected correlations with other constructs (Christopher et al., 2018).

The Self-Compassion Scale-short form (SCS-SF; Raes et al., 2011) is a 12-item version of the 26-item SCS (Neff, 2003). It assesses kindness and understanding toward oneself in instances of pain or failure, perception of one's experiences as part of the larger human experience, and ability to hold painful thoughts and feelings in mindful awareness. Items are rated on a 5-point Likert-type scale and summed with higher scores indicating greater self-compassion. The SCS-SF demonstrated good internal consistency, factorial validity, and expected correlations with other constructs (Raes et al., 2011).

Three items from the Credibility/Expectancy Questionnaire (CEQ; Devilly & Borkovec, 2000; Hicks et al., 2016) were adapted to assess participant expectancy at baseline regarding how successful they believed MBRT would be in: (1) reducing job stress, (2) improving work performance, and (3) enhancing resilience. Each item is rated on a 7-point Likert-type scale and summed to create a total score, with higher scores indicating greater treatment expectancy. The CEQ was included as a measure of discriminant validity as there is no theoretical relationship between treatment expectancy and aggression.

### 2.3 | Statistical analysis

To assess factorial validity, an item-level BPAQ-SF hierarchical confirmatory factor analysis (CFA) model was tested using robust diagonal weighted least squares (RDWLS) estimation with LISREL 10.20 (Jöreskog & Sörbom, 2019) using baseline data from all 170 participants. RDWLS was used due to the ordinal scale nature of the BPAQ-SF (i.e., Likert-type items) and related issues of multivariate non-normality. A procedure described by Jöreskog (2004) was used in which polychoric correlations are estimated and then rescaled into a polychoric covariance matrix. To assess the fit of the data to the model, five fit indices were evaluated: the Satorra-Bentler adjusted  $\chi^2$  ( $SB\chi^2$ ; Satorra & Bentler, 2001), root mean square error of approximation (RMSEA; Marsh et al., 1996), comparative fit index (CFI; Bentler, 1990), non-normed fit index (NNFI; Bentler & Bonett, 1980), and the standardized root mean squared residual (SRMR; Hu & Bentler, 1998). RMSEA values of .06 or less are thought to indicate a close fit, .08 a fair fit, and .10 a marginal fit; CFI and NNFI values of .90 and greater, and SRMR values of approximately .08 or less, tend to indicate acceptable fit (Hu & Bentler, 1999). A  $SB\chi^2$  difference test (Satorra & Bentler, 2001) was performed to compare model fit between the hierarchical (four factors loading onto an overall aggression factor) and correlated four-factor models. A significant  $SB\chi^2$  difference test value in the comparison of two nested models indicates that the constraints specified in the more restrictive

model (i.e., hierarchical) degrade model fit. Practical evidence of invariance was also judged by evaluating the  $\Delta CFI$  when an equality constraint was imposed; values of  $\geq .01$  are indicative of difference in fit between models (Cheung & Rensvold, 2002). To compare to a baseline model, data from a single-factor model (all BPAQ-SF items loading onto an aggression factor) were also analyzed.

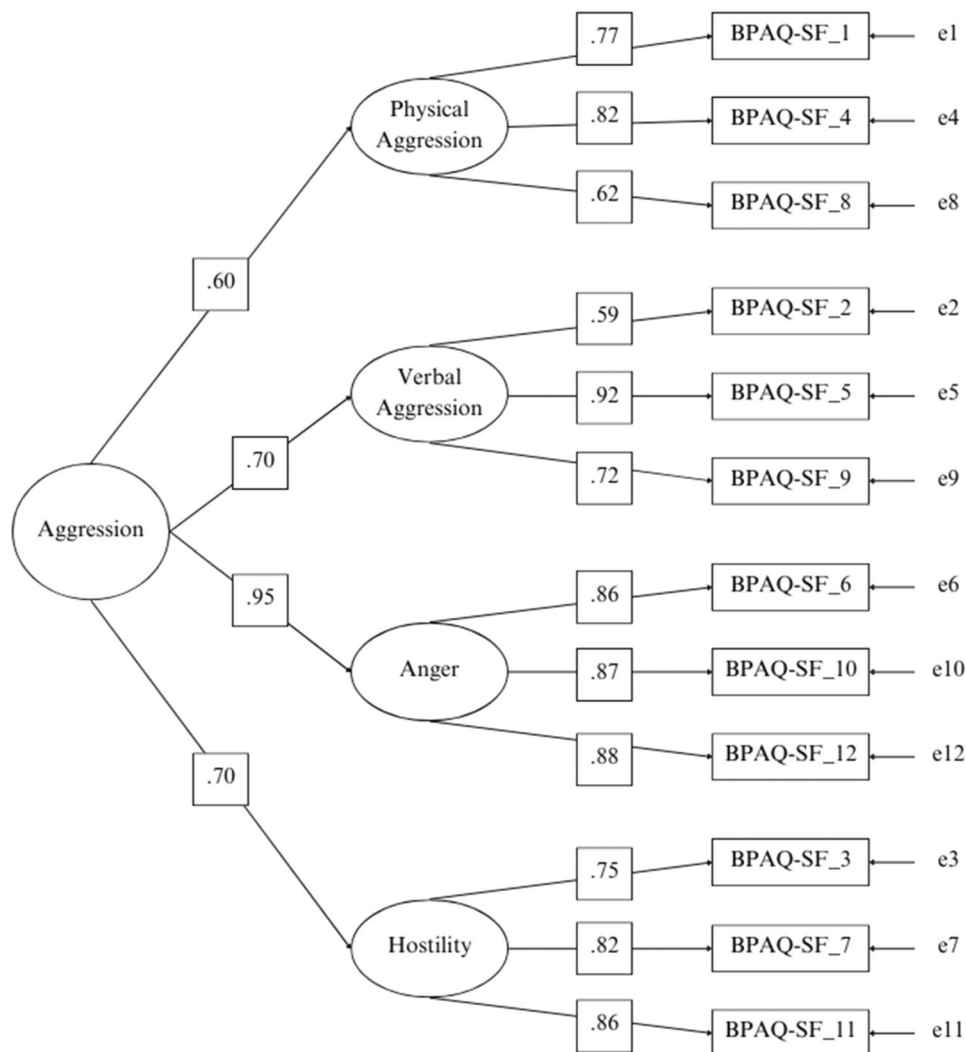
Cronbach's  $\alpha$  was used to assess internal consistency and Pearson's correlations were used to assess convergent and discriminant validity on baseline data ( $n = 170$ ). Only Study 1 participants completed the PROMIS-anger, and only Study 2 participants completed the PCL-5 and PSS-10, therefore, correlations between these and other variables had a smaller sample size ( $n = 61$  and  $n = 109$ , respectively). Similarly, participants who deny alcohol use on the initial item of the PROMIS alcohol use scale do not complete the measure, which resulted in a smaller sample size ( $n = 139$ ) for this measure. As noted above, test-retest reliability was assessed using a pre-to-post sample of NIC participants across Studies 1 and 2 ( $n = 40$ ; 8-week interval between assessments) and sensitivity to change was assessed using paired-samples  $t$  tests in a subsample of MBRT participants in Studies 1 and 2 ( $n = 55$ ).

We assessed differences in demographic and self-report variables at baseline between participants from Study 1 ( $n = 61$ ) and 2 ( $n = 109$ ) using independent  $t$ -tests and chi-square tests. Participants in Study 1 ( $M = 42.97$ ,  $SD = 6.05$ ) were older than Study 2 ( $M = 39.61$ ,  $SD = 9.31$ ,  $p = .001$ ), and a greater number of participants in Study 2 ( $n = 36$ , 33%) identified as Hispanic or Latino relative to Study 1 ( $n = 9$ , 15%,  $p = .007$ ). There were no other demographic differences and no statistical differences on any self-report measures.

## 3 | RESULTS

Fit indices for the hierarchical BPAQ-SF model were mixed, ranging from poor (RMSEA) to acceptable (SRMR) to good (CFI, NNFI):  $SB\chi^2_{(25.84)} = 62.50$ ,  $p = .0001$ ; CFI = .94, NNFI = .92, RMSEA = .19 (90% confidence interval = .17-.21), SRMR = .08. The 4-factor correlated model provided a comparable fit to the data:  $SB\chi^2_{(25.31)} = 62.08$ ,  $p = .0001$ ; CFI = .94, NNFI = .92, RMSEA = .19 (90% confidence interval = .17-.21), SRMR = .08. A  $SB\chi^2$  difference test indicated the two models did not differ in fit ( $\Delta SB\chi^2 = .42$ ,  $df = .53$ ,  $p = 1.00$ ), and practical evidence for similar model fit was supported by a  $\Delta CFI = .00$  when comparing models. All factor loadings were statistically significant in the hierarchical model (see Figure 1). Lastly, the single-factor model provided a poor fit to the data:  $SB\chi^2_{(23.99)} = 95.82$ ,  $p < .0001$ ; CFI = .87, NNFI = .83, RMSEA = .24 (90% confidence interval = .22-.26), SRMR = .13.

The BPAQ-SF and each factor demonstrated adequate to good internal consistency at baseline ( $\alpha$ 's: total = .84, anger = .85, hostility = .73, physical = .65, verbal = .74). As shown in Table 1, the BPAQ-SF and most factors were significantly correlated in the expected direction with demonstrated predictors of aggression in LEO samples. As expected, the CEQ did not correlate with the BPAQ-SF or three of four BPAQ-SF factors; however, it did correlate with



**FIGURE 1** Completely standardized Buss–Perry Aggression Questionnaire-short (BPAQ-SF) confirmatory factor analysis hierarchical solution. All factor loadings are statistically significant ( $p < .01$ ).

the BPAQ-SF anger factor (as well as PROMIS anger), providing partial evidence of discriminant validity. Test–retest reliability over an approximately 8-week period within waitlist control participants ( $n = 40$ ), was excellent for the total BPAQ-SF ( $r = .86$ ,  $p < .001$ ) and three of the four BPAQ-SF factors—anger ( $r = .86$ ,  $p < .001$ ), physical aggression ( $r = .89$ ,  $p < .001$ ), and verbal aggression ( $r = .83$ ,  $p < .001$ )—with questionable test–retest reliability for hostility ( $r = .60$ ,  $p < .001$ ). The BPAQ-SF also demonstrated sensitivity to change, as there was a statistically significant decrease pre-to-post MBRT in total aggression, anger, hostility, physical aggression, and verbal aggression (see Table 2).

## 4 | DISCUSSION

The goals of this study were to replicate the validity of the BPAQ-SF among LEOs, and to demonstrate novel evidence for convergent validity, test–retest reliability, and sensitivity to change in this

specific population. Consistent with a validation study of the BPAQ-SF in U.S. LEOs published over 20 years ago (Greenberg et al., 2003), results suggest the BPAQ-SF has evidence of factorial and convergent validity, and internal consistency among U.S. LEOs. The current results extend this prior research, additionally demonstrating the BPAQ-SF has excellent test–retest reliability and, crucially, is sensitive to change among LEOs.

Overall, the hierarchical BPAQ-SF CFA provided a mixed fit to the data. Although CFI, NNFI, and SRMR values indicated good to acceptable fit to the data, the RMSEA value (0.19) was indicative of poor fit. However, RMSEA is often artificially inflated in models with low degrees of freedom (e.g., CFAs), prompting some to argue that it should not be computed in these models (Kenny et al., 2015). Similarly, authors of a recent simulation study found CFA models with small sample sizes ( $N < 200$ ) and few observed variables ( $\leq 10$ ) resulted in dramatically inflated RMSEA values relative to other fit indices (Shi et al., 2019). These issues are particularly problematic in CFAs of short-form measures such as the BPAQ-SF, in which three

TABLE 1 Zero-order correlations between Buss-Perry Aggression Questionnaire-short form and conceptually related measures at baseline.

	BPAQ-SF M = 22.10 SD = 6.96	BPAQ-PA M = 4.37 SD = 1.96	BPAQ-VA M = 6.94 SD = 2.36	BPAQ-ANG M = 5.01 SD = 2.53	BPAQ-HOS M = 5.78 SD = 2.61	PROMIS- ANG M = 51.92 SD = 7.79	PROMIS-SD M = 52.40 SD = 7.39	PROMIS-AU M = 45.85 SD = 7.29	PCL-5 M = 34.18 SD = 11.90	PSS-10 M = 25.69 SD = 5.68	OLBI M = 2.38 SD = 0.40	FFMQ-SF M = 51.05 SD = 8.80	SCS-SF M = 39.99 SD = 7.40	CEQ M = 15.07 SD = 3.08
BPAQ-SF	.84													
BPAQ-PA	.62**	.65												
BPAQ-VA	.73**	.28**	.74											
BPAQ-ANG	.83**	.39**	.50**	.85										
BPAQ-HOS	.75**	.27**	.34**	.50**	.73									
PROMIS-ANG	.37**	.13	.15	.46**	.32**	.90								
PROMIS-SD	.34**	.24**	.13	.32**	.29**	.31*	.84							
PROMIS-AU	.23**	.12	.21*	.22**	.13	.35**	.20*	.92						
PCL-5	.56**	.30**	.30**	.54**	.48**	-	.51**	.15	.93					
PSS-10	.47**	.19*	.28**	.39**	.48**	-	.21*	.10	.46**	.78				
OLBI	.35**	.15	.11	.36**	.38**	.46**	.38**	.21*	.34**	.61**	.81			
FFMQ-SF	-.41**	-.022**	-.23**	-.35**	-.37**	-.32*	-.22**	-.11	-.41**	-.58**	-.47**	.85		
SCS-SF	-.42**	-.22**	-.23**	-.37**	-.38**	-.40**	-.27**	-.21*	-.37**	-.47**	-.39**	.62**	.82	
CEQ	.10	.06	-.10	.23*	.10	.39*	.01	.01	.03	-.15	-.07	-.12	.01	.92

Note: Cronbach's  $\alpha$ s are on the diagonal. The PROMIS-ANG was only administered in Study 1 ( $n = 61$ ), the PCL-5 and PSS-10 were only administered in Study 2 ( $n = 109$ ), and all other measures were administered in both studies ( $n = 170$ ).

Abbreviations: BPAQ-ANG, Buss-Perry Aggression Questionnaire-anger; BPAQ-HOS, Buss-Perry Aggression Questionnaire-hostility; BPAQ-PA, Buss-Perry Aggression Questionnaire-physical aggression; BPAQ-SF, Buss-Perry Aggression Questionnaire-short form; BPAQ-VA, Buss-Perry Aggression Questionnaire-verbal aggression; CEQ, Credibility/Expectancy Questionnaire; FFMQ-SF, Five Facet Mindfulness Questionnaire-short form; OLBI, Oldenburg Burnout Inventory; PCL-5, PTSD checklist for DSM-5; PROMIS-ANG, patient-reported outcomes measurement information system-anger; PROMIS-AU, patient-reported outcomes measurement information system-alcohol use; PROMIS-SD, patient-reported outcomes measurement information system-sleep disturbance; PSS, Perceived Stress Scale; SCS-SF, Self-Compassion Scale-short form.

\* $p < .05$ ; \*\* $p < .001$ .

**TABLE 2** Changes in Buss–Perry Aggression Questionnaire-short form scores from pre- to post-mindfulness-based resilience training.

	Pre-MBRT		Post-MBRT		<i>t</i>	<i>p</i>	<i>d</i>	95% CI of the effect size
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
BPAQ-SF total	23.22	7.93	19.31	5.69	4.94	<.0001	.67	.37, .96
BPAQ-SF anger	5.40	2.69	4.24	1.76	4.00	<.0001	.54	.25, .82
BPAQ-SF hostility	6.15	2.82	5.02	2.45	4.24	<.0001	.58	.29, .86
BPAQ-SF physical aggression	4.47	2.12	3.95	1.48	2.32	.02	.31	.04, .58
BPAQ-SF verbal aggression	7.20	2.46	6.11	2.40	3.83	<.0001	.52	.23, .80

Note: *n* = 55.

Abbreviations: BPAQ-SF, Buss–Perry Aggression Questionnaire-short form; CI, confidence interval; MBRT, mindfulness-based resilience training.

observed variables are assigned per factor. Moreover, based on simulation and real-world data, others have discouraged researchers from rejecting a model based on one fit index if the others are acceptable (Jackson et al., 2009; Lai & Green, 2016). Alongside evidence of acceptable fit from CFI, NNFI, and SRMR indices, and large and statistically significant factor loadings, these results suggest RMSEA may indeed be inflated and inaccurate. Despite the limited relevance to the current model, for full transparency, we opted to include RMSEA results. Overall, these results are very similar to Greenberg et al. (2003) who reported a “reasonable” data fit to the hierarchical BPAQ-SF CFA model with CFI = .91 and NNFI = .89 (RMSEA was not reported). Greenberg et al. (2003) also found “reasonable” fit to the 4-factor correlated model; however, they argue the hierarchical model is consistent with BPAQ conceptualization of aggression and is the preferred model.

Each item only loaded on the relevant latent construct, and each factor loaded onto an overall aggression factor, indicating that the BPAQ-SF has validity in assessing characteristics of anger, hostility, physical aggression, and verbal aggression in LEOs. Indices of internal consistency were good, with the possible exception of the physical aggression factor ( $\alpha = .65$ ). This appears to be due to the item “I have threatened people I know,” which had a smaller loading on the physical aggression factor than the other items, consistent with a prior validation study of the BPAQ-SF in U.S. LEOs (Greenberg et al., 2003). As the authors of that study suggested, this may reflect social desirability effects, given that LEOs are unlikely to report engaging in what they know to be illegal physically threatening behavior. Notably, the 12-item Brief Aggression Questionnaire (BAQ; Webster et al., 2014; Webster et al., 2015) replaced the “I have threatened people I know” item with “If I have to resort to violence to protect my rights, I will,” while keeping the other two physical aggression items from the BPAQ-SF. Given the BAQ initial validation samples consisted exclusively of undergraduate college students (Webster et al., 2014), its psychometrics should be further explored in an LEO sample.

The BPAQ-SF was positively correlated (all  $p$ 's < 0.05) with a number of demonstrated predictors of aggression among LEOs, including perceived stress, trauma exposure, alcohol misuse, burnout, and sleep disturbances (Hine et al., 2018; Juczyński & Ogińska-Bulik, 2022; Kurtz & Hughes, 2021; Mumford et al., 2021; Oehme et al., 2012; Queirós et al., 2013; Rajaratnam et al., 2011; Verhage

et al., 2018). The BPAQ-SF was also negatively correlated (all  $p$ 's < 0.05) with several buffers against aggression among LEOs, including self-reported mindfulness and self-compassion (Khatib et al., 2022; Ribeiro et al., 2020). This demonstration of convergent validity is important because it provides evidence that aggression correlates in the expected direction with factors that can precipitate or buffer against aggressive behavior among LEOs. Also supporting discriminant validity, the BPAQ-SF did not correlate with treatment expectancy. However, the BPAQ-anger factor and PROMIS anger both correlated with treatment expectancies. This unexpected finding suggests that LEOs with higher anger may perceive MBRT and perhaps other related interventions as helpful in remediating their anger.

Extending prior research on the factor structure and convergent validity of the BPAQ-SF in LEOs, our results provide novel evidence for excellent test–retest reliability and sensitivity to change in this population. Collapsing results across two intervention studies, total scores and each of the four BPAQ-SF factors showed significant reductions in 55 participants who completed mindfulness-based resilience training (MBRT), a police-specific training intended to improve LEO health and reduce aggression. In addition, we were able to utilize data from waitlist control participants to demonstrate excellent test–retest reliability ( $r = .86$ ) across an approximately 8-week period, and excellent test–retest reliability for three of four BPAQ-SF factors. Not only do these results highlight the potential for mindfulness training to reduce aggression in this population, as has previously been demonstrated in other groups (Fix & Fix, 2013; Tao et al., 2021), these psychometric properties of the BPAQ-SF make it an excellent candidate outcome measure for other intervention studies that seek to reduce LEO aggression. Relatedly, although the BPAQ-SF is often conceptualized as a measure of trait aggression, consistent with these results, others have found it sensitive to change in intervention research (e.g., Johnson, Sandel, et al., 2020; Johnson, Zisser, et al., 2020).

Findings from the current study support the use of the BPAQ-SF among LEOs in both research and applied settings. Higher levels of aggression are likely to be associated with excessive use of force (Greenberg et al., 2003) and other negative community outcomes. The reliability and validity of the BPAQ-SF support its inclusion as a brief measure in research settings to further understand individual

differences in aggression and their relationship with use of force and other real-world outcomes. Moreover, the excellent test-retest reliability and sensitivity to change following a mindfulness-based intervention support the expanded use of the BPAQ-SF in research on interventions intended to reduce aggression and its negative consequences. From an applied perspective, the validity and reliability of the BPAQ-SF support its use in clinical assessment settings to monitor treatment progress among LEOs, consistent with several previous studies which advocate the use of BPAQ-SF in clinical populations (e.g., Johnson, Zisser, et al., 2020; Torregrosa et al., 2020). The BPAQ-SF may also be used by law enforcement practitioners in nonclinical settings (e.g., by including the BPAQ-SF as a screening tool for early intervention or for departments that seek to assess aggression as part of regular department wellness check-ins). Updated guidelines from the International Chiefs of Police Psychological Services Section (2020) on "Preemployment Psychological Evaluation Guidelines" recommend the use of psychological assessments "with documented reliability, validity, and other empirical evidence supporting the use in the preemployment evaluation of public safety applicants" (guideline 7.1). As such, the current results are a valuable step toward advancing use of the BPAQ-SF as a potential screening tool that may offer benefits over existing screening tools such as the MMPI-2.

There are a number of limitations to this study. First, this was a largely homogeneous sample (primarily white males), which precludes generalizability of findings. Future studies with larger and more diverse LEO samples are needed. Second, it would have been ideal to have administered the full BPAQ compared to the short form, particularly because BPAQ physical aggression items excluded from the BPAQ-SF assess additional elements that are relevant to LEOs (e.g., "if I have to resort to violence to protect my rights, I will."). Third, there is an inherent limitation when utilizing self-report measures to assess aggression, as these reports are subject to social desirability effects, particularly for a highly scrutinized population asked to report on socially undesirable and even illegal behavior. This underscores the importance of anchoring self-reported aggression in objective or third-party measures of aggression, such as administrative use-of-force data or reports from community members.

Overall, the results from this study support our hypothesis that the BPAQ-SF is a psychometrically sound measure of self-reported aggression among LEOs. These results provide clinicians, law enforcement practitioners, and researchers a useful tool to investigate individual differences in LEO aggression, its downstream consequences, and its reduction through evidence-based interventions.

## ACKNOWLEDGMENTS

The authors express their sincere gratitude to the law enforcement officers who participated in the study and to Amit Bernstein, Ph.D. for his assistance with data interpretation. Research reported in this publication was supported by the National Center for Complementary & Integrative Health of the National Institutes of Health under Awards R21AT008854 and R01AT009841. The content is solely the

responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

## DATA AVAILABILITY STATEMENT

The data set in this study is publicly available at <https://osf.io/bskn9/>.

## ETHICS STATEMENT

The Pacific University IRB approved all procedures (IRB#s 089-15, 090-18). Informed consent was obtained from all individual participants included in the study.

## ORCID

Michael Christopher  <http://orcid.org/0000-0003-4657-4571>

Marissa Ferry  <http://orcid.org/0000-0001-6417-0568>

Akeesha Simmons  <http://orcid.org/0000-0002-2494-3332>

Alicia Vasquez  <http://orcid.org/0000-0001-6820-0373>

Brooke Reynolds  <http://orcid.org/0000-0003-2965-3964>

Daniel Grupe  <http://orcid.org/0000-0003-4927-7572>

## REFERENCES

- Alang, S., McAlpine, D., & McClain, M. (2021). Police encounters as stressors: associations with depression and anxiety across race. *Socius: Sociological Research for a Dynamic World*, 7, 2378023 12199812. <https://doi.org/10.1177/2378023121998128>
- Alexander, K. L., Rich, S., & Thacker, S. (2022). What you need to know about the cost of police misconduct. *The Washington Post*, March 10. <https://www.washingtonpost.com/investigations/2022/03/10/what-you-need-know-about-cost-police-misconduct/>
- Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment*, 13(1), 27–45. <https://doi.org/10.1177/1073191105283504>
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588–606. <https://doi.org/10.1037/0033-2909.88.3.588>
- Bohlmeijer, E., ten Klooster, P. M., Fledderus, M., Veehof, M., & Baer, R. (2011). Psychometric properties of the five facet mindfulness questionnaire in depressed adults and development of a short form. *Assessment*, 18(3), 308–320. <https://doi.org/10.1177/1073191111408231>
- Bowleg, L., Maria del río-González, A., Mbaba, M., Boone, C. A., & Holt, S. L. (2020). Negative police encounters and police avoidance as pathways to depressive symptoms among US black men, 2015–2016. *American Journal of Public Health*, 110(S1), S160–S166. <https://doi.org/10.2105/ajph.2019.305460>
- Boxer, P., Brunson, R. K., Gaylord-Harden, N., Kahn, K., Patton, D. U., Richardson, J., Rivera, L. M., Smith Lee, J. R., Staller, M. S., Krahé, B., Dubow, E. F., Parrott, D., & Algrim, K. (2021). Addressing the inappropriate use of force by police in the United States and beyond: A behavioral and social science perspective. *Aggressive Behavior*, 47(5), 502–512. <https://doi.org/10.1002/ab.21970>
- de Bruin, E. I., Topper, M., Muskens, J. G. A. M., Bögels, S. M., & Kamphuis, J. H. (2012). Psychometric properties of the five facets mindfulness questionnaire (FFMQ) in a meditating and a



- non-meditating sample. *Assessment*, 19(2), 187–197. <https://doi.org/10.1177/1073191112446654>
- Bryant, F. B., & Smith, B. D. (2001). Refining the architecture of aggression: A measurement model for the Buss–Perry Aggression Questionnaire. *Journal of Research in Personality*, 35(2), 138–167. <https://doi.org/10.1006/jrpe.2000.2302>
- Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, 63(3), 452–459. <https://doi.org/10.1037/0022-3514.63.3.452>
- Butcher, J. N. (2010). Minnesota multiphasic personality inventory. *The Corsini encyclopedia of psychology* (1–3). Wiley. <https://doi.org/10.1002/9780470479216.corpsy0573>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233–255. [https://doi.org/10.1207/S15328007SEM0902\\_5](https://doi.org/10.1207/S15328007SEM0902_5)
- Christopher, M., Bowen, S., Witkiewitz, K., Grupe, D., Goerling, R., Hunsinger, M., Oken, B., Korecki, T., & Rosenbaum, N. (2023). A multisite feasibility randomized clinical trial of mindfulness-based resilience training for aggression, stress, and health in law enforcement officers. *PsyArXiv*, <https://doi.org/10.31234/osf.io/wm6ch>
- Christopher, M. S., Hunsinger, M., Goerling, L. R. J., Bowen, S., Rogers, B. S., Gross, C. R., Dapollonia, E., & Pruessner, J. C. (2018). Mindfulness-based resilience training to reduce health risk, stress reactivity, and aggression among law enforcement officers: A feasibility and preliminary efficacy trial. *Psychiatry Research*, 264, 104–115. <https://doi.org/10.1016/j.psychres.2018.03.059>
- Christopher, M. S., Neuser, N. J., Michael, P. G., & Baitmangalkar, A. (2012). Exploring the psychometric properties of the five facet mindfulness questionnaire. *Mindfulness*, 3, 124–131.
- Cobbina-Dungy, J. E., & Jones-Brown, D. (2023). Too much policing: Why calls are made to defund the police. *Punishment & Society*, 25(1), 3–20. <https://doi.org/10.1177/14624745211045652>
- Cochrane, R. E., Tett, R. P., & Vandecreek, L. (2003). Psychological testing and the selection of police Officers: A national survey. *Criminal Justice and Behavior*, 30(5), 511–537. <https://doi.org/10.1177/0093854803257241>
- Cohen, S. (1988). Perceived stress in a probability sample of the United States. In S. Spacapan, & S. Oskamp (Eds.), *The social psychology of health*. Sage.
- Davis, R. D., Rostow, C. D., Pinkston, J. B., Combs, D. R., & Dixon, D. R. (2004). A re-examination of the MMPI-2 aggressiveness and immaturity indices in law enforcement screening. *Journal of Police and Criminal Psychology*, 19(1), 17–26. <https://doi.org/10.1007/bf02802571>
- Demerouti, E., Bakker, A. B., Vardakou, I., & Kantas, A. (2003). The convergent validity of two burnout instruments. *European Journal of Psychological Assessment*, 19(1), 12–23. <https://doi.org/10.1027/1015-5759.19.1.12>
- Demerouti, E., Mostert, K., & Bakker, A. B. (2010). Burnout and work engagement: A thorough investigation of the independency of both constructs. *Journal of Occupational Health Psychology*, 15(3), 209–222. <https://doi.org/10.1037/a0019408>
- Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the Credibility/Expectancy Questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry*, 31(2), 73–86.
- Diamond, P. M., & Magaletta, P. R. (2006). The short-form Buss–Perry Aggression Questionnaire (BPAQ-SF): A validation study with federal offenders. *Assessment*, 13(3), 227–240. <https://doi.org/10.1177/1073191106287666>
- Fix, R. L., & Fix, S. T. (2013). The effects of mindfulness-based treatments for aggression: A critical review. *Aggression and Violent Behavior*, 18(2), 219–227. <https://doi.org/10.1016/j.avb.2012.11.009>
- Greenberg, B. E., Riggs, M., Bryant, F. B., & Smith, B. D. (2003). Validation of a short aggression inventory for law enforcement. *Journal of Police and Criminal Psychology*, 18(2), 12–19. <https://doi.org/10.1007/BF02807243>
- Halbesleben, J. R. B., & Demerouti, E. (2005). The construct validity of an alternative measure of burnout: Investigating the English translation of the Oldenburg Burnout Inventory. *Work & Stress*, 19(3), 208–220. <https://doi.org/10.1080/02678370500340728>
- Haldipur, J. (2020). No place on the corner: The costs of aggressive policing. *Social Forces*, 99(2), e1–e3. <https://doi.org/10.1093/sf/soaa059>
- Hargrave, G. E., Hiatt, D., & Gaffney, T. W. (1988). F + 4 + 9 + Cn: An MMPI measure of aggression in law enforcement officers and applicants. *Journal of Police Science & Administration*, 16(4), 268–273.
- Harris, J. A. (1997). A further evaluation of The aggression questionnaire: Issues of validity and reliability. *Behaviour Research and Therapy*, 35(11), 1047–1053. [https://doi.org/10.1016/S0005-7967\(97\)00064-8](https://doi.org/10.1016/S0005-7967(97)00064-8)
- Hawkins, D. S. (2022). "After philando, I had to take a sick day to recover": Psychological distress, trauma and police brutality in the black community. *Health Communication*, 37(9), 1113–1122. <https://doi.org/10.1080/10410236.2021.1913838>
- Hicks, M., Hanes, D., & Wahbeh, H. (2016). Expectancy effect in three mind-body clinical trials. *Journal of Evidence-Based Complementary & Alternative Medicine*, 21(4), NP103–NP109. <https://doi.org/10.1177/2156587216652572>
- Hine, K. A., Porter, L. E., Westera, N. J., Alpert, G. P., & Allen, A. (2018). Exploring police use of force decision-making processes and impairments using a naturalistic decision-making approach. *Criminal Justice and Behavior*, 45(11), 1782–1801. <https://doi.org/10.1177/0093854818789726>
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to under-parameterized model misspecification. *Psychological Methods*, 3, 424–453. <https://doi.org/10.1037/1082-989x.3.4.424>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Jackson, D. L., Gillaspay, Jr., J. A., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods*, 14(1), 6–23. <https://doi.org/10.1037/a0014694>
- Johnson, S. L., Sandel, D. B., Zisser, M., Pearlstein, J. G., Swerdlow, B. A., Sanchez, A. H., Fernandez, E., & Carver, C. S. (2020). A brief online intervention to address aggression in the context of emotion-related impulsivity for those treated for bipolar disorder: Feasibility, acceptability and pilot outcome data. *Journal of Behavioral and Cognitive Therapy*, 30(1), 65–74. <https://doi.org/10.1016/j.jbct.2020.03.005>
- Johnson, S. L., Zisser, M. R., Sandel, D. B., Swerdlow, B. A., Carver, C. S., Sanchez, A. H., & Fernandez, E. (2020). Development of a brief online intervention to address aggression in the context of emotion-related impulsivity: Evidence from a wait-list controlled trial. *Behaviour Research and Therapy*, 134, 103708. <https://doi.org/10.1016/j.brat.2020.103708>
- Jöreskog, K. G. (2004). Structural equation modeling with ordinal variables using LISREL. <http://www.ssicentral.com/lisrel/techdocs/ordinal.pdf>
- Jöreskog, K. G., & Sörbom, D. (2019). LISREL 10.20. Scientific Software International.
- Juczyński, Z., & Ogińska-Bulik, N. (2022). Ruminations and occupational stress as predictors of post-traumatic stress disorder and burnout among police officers. *International Journal of Occupational Safety and Ergonomics*, 28(2), 743–750. <https://doi.org/10.1080/10803548.2021.1907986>
- Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, 44(3), 486–507.

- Khatib, L., Glaser-Reich, J., Mosbey, D., Oliva, V., Riegner, G., Dean, J. G., Harth, N. M., & Zeidan, F. (2022). Mindfulness meditation training reduces aggression and improves well-being in highly stressed law enforcement officers. *Journal of Police and Criminal Psychology*, 37(4), 972–983. <https://doi.org/10.1007/s11896-022-09554-2>
- Kurtz, D. L., & Hughes, V. (2021). Using structural equation modeling to explore the role of early strain on later stress responses and aggression by police officers. *Police Practice and Research*, 22(1), 355–369. <https://doi.org/10.1080/15614263.2020.1749623>
- Lai, K., & Green, S. B. (2016). The problem with having two watches: Assessment of fit when RMSEA and CFI disagree. *Multivariate Behavioral Research*, 51(2–3), 220–239. <https://doi.org/10.1080/00273171.2015.1134306>
- Loya, K. M., Peskosky, E. C., & Jacquin, K. M. (2022). Predictive validity of MMPI-2 immaturity index for early termination of police officers. *Journal of Police and Criminal Psychology*, 37(1), 222–228. <https://doi.org/10.1007/s11896-022-09499-6>
- Marsh, H. W., Balla, J. R., & Hau, K.-T. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical properties. *Advanced Structural Equation Modeling: Issues and Techniques*, 315–353.
- McLean, K., Stoughton, S. W., & Alpert, G. P. (2022). Police uses of force in the USA: A wealth of theories and a lack of evidence. *Cambridge Journal of Evidence-Based Policing*, 6(3–4), 87–108.
- McLeod, M. N., Heller, D., Manze, M. G., & Echeverria, S. E. (2020). Police interactions and the mental health of black Americans: A systematic review. *Journal of Racial and Ethnic Health Disparities*, 7(1), 10–27. <https://doi.org/10.1007/s40615-019-00629-1>
- Morley, R. H. (2015). Violent criminality and self-compassion. *Aggression and Violent Behavior*, 24, 226–240. <https://doi.org/10.1016/j.avb.2015.05.017>
- Mourtgos, S. M., & Adams, I. T. (2020). Assessing public perceptions of police use-of-force: Legal reasonableness and community standards. *Justice Quarterly*, 37(5), 869–899. <https://doi.org/10.1080/07418825.2019.1679864>
- Mumford, E. A., Liu, W., & Taylor, B. G. (2021). Profiles of U.S. law enforcement officers' physical, psychological, and behavioral health: Results from a nationally representative survey of officers. *Police Quarterly*, 24(3), 357–381. <https://doi.org/10.1177/10986111211991111>
- Neff, K. D. (2003). The development and validation of a scale to measure self-compassion. *Self and Identity*, 2(3), 223–250.
- Nix, J., Pickett, J. T., & Wolfe, S. E. (2020). Testing a theoretical model of perceived audience legitimacy: The neglected linkage in the dialogic model of police-community relations. *Journal of Research in Crime and Delinquency*, 57(2), 217–259. <https://doi.org/10.1177/0022427819873957>
- Oehme, K., Donnelly, E. A., & Martin, A. (2012). Alcohol abuse, PTSD, and officer-committed domestic violence. *Policing*, 6(4), 418–430. <https://doi.org/10.1093/policing/pas023>
- Piilkonis, P. A., Choi, S. W., Reise, S. P., Stover, A. M., Riley, W. T., & Cella, D. (2011). Item banks for measuring emotional distress from the patient-reported outcomes measurement information system (PROMIS®): Depression, anxiety, and anger. *Assessment*, 18(3), 263–283. <https://doi.org/10.1177/1073191111411667>
- Piilkonis, P. A., Yu, L., Dodds, N. E., Johnston, K. L., Lawrence, S. M., & Daley, D. C. (2016). Validation of the alcohol use item banks from the patient-reported outcomes measurement information system (PROMIS®). *Drug and Alcohol Dependence*, 161, 316–322. <https://doi.org/10.1016/j.drugalcdep.2016.02.014>
- Queirós, C., Kaiseler, M., & Leitão da Silva, A. (2013). Burnout as predictor of aggressivity among police officers. *European Journal of Policing Studies*, 1(2), 110–134.
- Raes, F., Pommier, E., Neff, K. D., & Van Gucht, D. (2011). Construction and factorial validation of a short form of the Self-Compassion Scale. *Clinical psychology & psychotherapy*, 18(3), 250–255. <https://doi.org/10.1002/cpp.702>
- Rajaratnam, S. M. W., Barger, L. K., Lockley, S. W., Shea, S. A., Wang, W., Landrigan, C. P., O'Brien, C. S., Qadri, S., Sullivan, J. P., Cade, B. E., Epstein, L. J., White, D. P., Czeisler, C. A., & Harvard Work Hours, Health and Safety Group (2011). Sleep disorders, health, and safety in police officers. *Journal of the American Medical Association*, 306(23), 2567–2578. <https://doi.org/10.1001/jama.2011.1851>
- Ribeiro, L., Colgan, D. D., Hoke, C. K., Hunsinger, M., Bowen, S., Oken, B. S., & Christopher, M. S. (2020). Differential impact of mindfulness practices on aggression among law enforcement officers. *Mindfulness*, 11(3), 734–745. <https://doi.org/10.1007/s12671-019-01289-2>
- Roberts, N. P., Kitchiner, N. J., Lewis, C. E., Downes, A. J., & Bisson, J. I. (2021). Psychometric properties of the PTSD checklist for DSM-5 in a sample of trauma exposed mental health service users. *European Journal of Psychotraumatology*, 12(1), 1863578. <https://doi.org/10.1080/20008198.2020.1863578>
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, 66(4), 507–514. <https://doi.org/10.1007/bf02296192>
- Shi, D., Lee, T., & Maydeu-Olivares, A. (2019). Understanding the model size effect on SEM fit indices. *Educational and Psychological Measurement*, 79(2), 310–334. <https://doi.org/10.1177/0013164418783530>
- Singh, N. N., Lancioni, G. E., Karazsia, B. T., Winton, A. S. W., Singh, J., & Wahler, R. G. (2014). Shenpa and compassionate abiding: Mindfulness-based practices for anger and aggression by individuals with schizophrenia. *International Journal of Mental Health and Addiction*, 12, 138–152.
- Tao, S., Li, J., Zhang, M., Zheng, P., Lau, E. Y. H., Sun, J., & Zhu, Y. (2021). The effects of mindfulness-based interventions on child and adolescent aggression: A systematic review and meta-analysis. *Mindfulness*, 12(6), 1301–1315. <https://doi.org/10.1007/s12671-020-01570-9>
- Tapp, S. N., & Davis, E. (2022). Contacts between police and the public, 2020 (NCJ 304527). Bureau of Justice Statistics. Research and Development Series, November.
- Taylor, J. M. (2015). Psychometric analysis of the ten-item perceived stress scale. *Psychological Assessment*, 27(1), 90–101. <https://doi.org/10.1037/a0038100>
- Torregrosa, M. S., Gómez-Núñez, M. I., Inglés, C. J., Ruiz-Esteban, C., Sanmartín, R., & García-Fernández, J. M. (2020). Buss and Perry aggression questionnaire-short form in Spanish children. *Journal of Psychopathology and Behavioral Assessment*, 42(4), 677–692. <https://doi.org/10.1007/s10862-020-09809-4>
- Verhage, A., Noppe, J., Feys, Y., & Ledegen, E. (2018). Force, stress, and decision-making within the Belgian police: The impact of stressful situations on police decision-making. *Journal of Police and Criminal Psychology*, 33(4), 345–357. <https://doi.org/10.1007/s11896-018-9262-4>
- Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). The PTSD checklist for DSM-5 (PCL-5). Scale available from the National Center for PTSD at [www.ptsd.va.gov](http://www.ptsd.va.gov).
- Webster, G. D., DeWall, C. N., Pond, Jr., R. S., Deckman, T., Jonason, P. K., Le, B. M., Nichols, A. L., Schember, T. O., Crysel, L. C., Crosier, B. S., Smith, C. V., Paddock, E. L., Nezek, J. B., Kirkpatrick, L. A., Bryan, A. D., & Bator, R. J. (2014). The brief aggression questionnaire: Psychometric and behavioral evidence for an efficient measure of trait aggression. *Aggressive Behavior*, 40(2), 120–139. <https://doi.org/10.1002/ab.21507>
- Webster, G. D., DeWall, C. N., Pond, Jr., R. S., Deckman, T., Jonason, P. K., Le, B. M., Nichols, A. L., Schember, T. O., Crysel, L. C., Crosier, B. S.,

- Smith, C. V., Paddock, E. L., Nezelek, J. B., Kirkpatrick, L. A., Bryan, A. D., & Bator, R. J. (2015). The brief aggression questionnaire: Structure, validity, reliability, and generalizability. *Journal of Personality Assessment*, *97*(6), 638–649. <https://doi.org/10.1080/00223891.2015.1044093>
- Wood, G., Tyler, T. R., & Papachristos, A. V. (2020). Procedural justice training reduces police use of force and complaints against officers. *Proceedings of the National Academy of Sciences*, *117*(18), 9815–9821.
- Yu, L., Buysse, D. J., Germain, A., Moul, D. E., Stover, A., Dodds, N. E., Johnston, K. L., & Pilkonis, P. A. (2012). Development of short forms from the PROMIS™ sleep disturbance and sleep-related impairment item banks. *Behavioral Sleep Medicine*, *10*(1), 6–24. <https://doi.org/10.1080/15402002.2012.636266>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Christopher, M., Ferry, M., Simmons, A., Vasquez, A., Reynolds, B., & Grupe, D. (2024). Psychometric properties of the Buss–Perry Aggression Questionnaire-short form among law enforcement officers. *Aggressive Behavior*, *50*, e22145. <https://doi.org/10.1002/ab.22145>