A Self-Report Approach to Screening Police Candidates’ Aggressive Tendencies

Frank Joseph Gallo

University of Rhode Island
A SELF-REPORT APPROACH TO SCREENING POLICE CANDIDATES' AGGRESSIVE TENDENCIES

BY

FRANK JOSEPH GALLO

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN PSYCHOLOGY

UNIVERSITY OF RHODE ISLAND 2005
Abstract

This paper introduces the concepts of labeling (detection of aggression severity) and provocation (response to directed aggression) as meaningful dimensions for evaluating police candidates' patterns of aggressive tendencies. The evaluation uses candidates' judgments on instances of suspect behaviors during hypothetical arrest situations. Findings showed that candidates agreed on an ordered continua of behavior severity, alpha = 0.99. One was able to predict very well candidates' provocation tendencies from knowing their labeling tendencies and vice versa, $r^2 = 0.76$. Labeling and provocation tendencies were related to other established measures of aggression (e.g., Buss-Perry Aggression Questionnaire, 1992). I discuss potential applications of candidates' labeling and provocation tendencies for use in police training sessions and employment evaluations.
Acknowledgements

Although this research bears my name, its beginning and completion would not be possible without the assistance of many other people.

To the police agencies and training academies who participated, their administrators and personnel who supported my request to survey police recruits, I thank them for their assistance and trust.

Members of my dissertation committee, Dr. Jerry L. Cohen, Dr. Patricia L. Gallagher, and Dennis C. Hilliard helped give substance and direction to this research. Dr. Patricia L. Gallagher was instrumental in sharing information on psychological testing in police employment settings.

My major professor, Dr. Charles E. Collyer, has been supportive and has offered valuable perceptions on measuring police candidates' aggressive tendencies. His thoughts on concepts of labeling and provocation were important to this study.

Finally, to my wife Kim and sons Frank and Nico, without their unending source of support this work would not be possible, thank you.
| Appendices                      | Page |
|--------------------------------|------|
| Appendix A: Labeling Task      | 49   |
| Appendix B: Items Constituting the Labeling and Provocation Tasks | 50   |
| Appendix C: Provocation Task   | 51   |
| Appendix D: Response to Provocation Short-Form | 53   |
| Appendix E: Interview Protocol | 54   |
| Appendix F: Buss-Perry Short-Form Aggression Questionnaire | 55   |
| Appendix G: Buss-Perry Physical Aggression Subscale | 56   |
| Appendix H: Aggressiveness Subscale Verbal Aggressiveness Scale | 57   |
| Appendix I: Anger Arousal Subscale Multidimensional Anger Inventory | 58   |
| Appendix J: Cook-Medley Hostility Scale | 59   |
| Appendix K: Marlowe-Crowne Short-Form Social Desirability Scale | 61   |
| Appendix L: Behavioral Experience Questionnaire | 62   |
| Appendix M: Average Ratings and Rankings of Suspect Behaviors | 65   |
| Appendix N: Labeling Task Behavior Items and Varimax Pattern | 66   |
| Appendix O: Frequency Count Behavioral Experience Questionnaire | 67   |

Bibliography                          | 68   |
List of Tables

Table 1. Descriptives for Labeling and Provocation Tasks by Recruit Classes........ 27
Table 2. Shapiro-Wilk Tests Labeling and Provocation Tasks by Recruit Classes .... 28
Table 3. Correlations of Labeling and Provocation Ratings with other Measures ..... 38
List of Figures

Figure 1. Severity ratings of suspect behaviors by corresponding severity rank........ 32
Figure 2. Severity ratings of suspect behaviors against One-for-One model.............. 33
Figure 3. Average responses to provocation ........................................................... 34
Figure 4. Recruits’ provocation ratings plotted by their labeling ratings............... 37
Police have the responsibility to safeguard the well-being of the community, prevent crime, and enforce laws (Adams, 1999). Constitutional provisions recognize that such work demands carry with them the need to use some degree of coercion to effect them (Terry v. Ohio, 1968). As long as some citizens continue to demonstrate their willingness to break the law, the use of force will remain an unavoidable activity of police work (Skolnick & Fyfe, 1993). The importance of aggression in police work demands selection procedures that screen in police candidates who are willing to be forceful, yet screen out those candidates who demonstrate a lack of restraint and self-control. Psychological screening has long been one component of evaluating a candidate’s ability to balance aggression (Benner, 1986). Yet, highly publicized abuses of force by the police remind contemporary thinkers that psychological screening efforts might not always identify the right balance. The tasks of this study are to (a) look at screening approaches used to decide which police candidates to select and which to weed out, (b) review psychological tests used for police selection and their links to predicting subsequent abuses of aggression, and (c) propose an aggression assessment to identify which candidates might experience difficulties with using aggression in a law enforcement capacity.

Approaches to Psychological Screening

Most police agencies recognize psychological evaluations as a required component of the police selection process (Detrick, Chibnall, & Rosso, 2001). The courts have looked at failure to provide psychological screening as negligence (Bonsignore v. City of New York, 1981; Conte v. Horcher, 1977; McKenna v. Fargo, 1
1978). Today, more than 80% of U.S. police agencies require the administration of psychological tests to screen police candidates (Craig, 2005). Traditionally, determining the psychological suitability of police candidates might involve two selection events: a “screen-out” decision or a “select-in” decision (Benner, 1986). Selecting in desirable police candidates involves choosing the most qualified who demonstrate positive qualities necessary to be successful in the work field. A work task analysis of successful officers is one strategy usually adopted for gaining select-in information that identifies important police functions and the necessary police characteristics to perform them (Craig, 2005; Inwald, Knatz, & Shusman, 1983). Effective policing and its performance require a mixture of tasks. The evolving nature of policing might cause difficulties in identifying certain qualities necessary to perform such tasks (Cohen & Chaiken, 1973). Consequently, the select-in decision might focus on particular police characteristics that no longer reflect existing police practices (Brengelman, 1982; as cited in Grant & Grant, 1995). For example, in the 1990s law enforcement saw a shift from car-based policing to intra community-based patrols that focused on problem-oriented policing. Police agencies recognized a new need for selecting officers who have certain problem solving skills.

Although psychologists might use select-in criteria to arrive at an accept decision, there is a lack of consensus among police stakeholders on the qualities needed to be successful in the law enforcement profession (Benner, 1986). Benner recognized that there is more agreement on the unwanted or negative qualities of police candidates and that selection usually involves a screen-out decision. This kind of selection event calls for eliminating those candidates who demonstrate undesirable
police characteristics. Psychological stability is the major concern. Psychologists hypothesize that a psychologically unstable officer is more likely to perform poorly in the work field than is a "normal" officer. However, some reviews on the screen-out approach suggest a lack of consistent evidence on predicting which candidates are more likely to experience on-the-job difficulties (e.g., Daley, 1982; Varela, Boccaccini, Scogin, Stump, & Caputo, 2004).

The select-in and screen-out decisions might include evaluations of mental health, which psychologists must carry out in accordance with the Americans with Disabilities Act (ADA), and only after a conditional offer of employment to the police candidate (Hibler & Kurke, 1995). For pre-conditional offers, psychologists use personality tests and other police screening methods that do not include evaluations of mental health (Vetter, 1999). Both conditional and pre-conditional offer psychological evaluations focus on screening for suitable candidates. The President's Commission on Law Enforcement and Administration of Justice (1967) recommends psychological research and development of valid tests for screen-out and select-in procedures.

In practice, pre-employment psychological evaluations should focus on the suitability of police candidates to perform essential job tasks, while screening for characteristics that may adversely affect their job performance (International Association of Chiefs of Police, 2004). Although some psychological screeners might lean toward a select-in or screen-out approach, many screeners favor a psychological test battery that satisfies both selection events: screen-out psychopathology (clinical) and select-in ideal police attitudes, traits, and background.
Tests for Psychological Screening

The International Association of Chiefs of Police (IACP) suggests twenty-two recommendations for the pre-employment psychological evaluation of police candidates (IACP, 2004). Among the criteria, the IACP recommends the use of objective and validated tests that specify what job-related functions they intend to measure. The Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1943), the California Psychological Inventory (CPI; Gough, 1975), and the Inwald Personality Inventory (IPI; Inwald et al., 1983) are some of the more commonly used psychological tests for police screening (Wrightsman, 2005).

The MMPI and the CPI are general personality inventories used to assess the relatively stable and enduring characteristics of test takers. They tap a number of dimensions (or factors) thought to make up the respondent’s personality, which might affect subsequent uses of job-work aggression. The MMPI is a clinical instrument designed to measure dimensions of psychopathology. Whereas the CPI is a non-clinical instrument designed to measure normal personality traits important for social living and interaction. Authors of the MMPI and the CPI did not initially design the instruments to screen police candidates. There are, however, police and public safety reports available for both the MMPI and CPI.

In contrast to the MMPI and CPI, Inwald et al. (1983) developed the IPI to predict normal as well as deviant job-performance patterns of police candidates. Four general content areas of the IPI measure job-related criteria: guardedness, acting out behaviors, internalized conflict, and interpersonal conflict. Psychological tests such as the MMPI, CPI, and IPI capture an objective measure of a sample of the candidate’s
behaviors (Anastasi & Urbina, 1997). The diagnostic value of these tests is to forecast what the candidate might say or do under work conditions. Therefore, “forming the connection between applicants’ test responses and eventual job performance is crucial in the evaluation of a test’s general usefulness” (Inwald & Shusman, 1984, p. 1). Do personality tests do well at predicting which police candidates will have difficulties with on-the-job uses of aggression?

Hargrave, Hiatt, and Gaffney (1988) found that elevations on an “aggression index” composed of MMPI scales F (infrequency), 4 (psychopathic deviate), and 9 (hypomania), combined with elevated Cn (control in psychological adjustment) scale scores correctly classified aggressive incumbent officers who received disciplinary actions for aggressive misconduct against offenders, inmates, co-workers, or family members. Costello, Schneider, and Schoenfeld (1996) observed that elevations on the F+4+9 aggression index predicted suspensions of officers after three years of service. Although MMPI scale scores have generated some discussion on their usefulness for predicting job-related aggression, there appears an absence of research on the utility of the MMPI's current version, the MMPI-2 (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989), for predicting police difficulties with uses of force.

There is some psychological literature showing validity of the CPI for predicting abuses of aggression. Hargrave and Hiatt (1989) reported an association between low CPI scale scores on socialization, self-control, and well-being and disciplinary actions against incumbent police officers. Job difficulties that led to disciplinary actions included unnecessary uses of force. Fitzgerald (1987) found that officers with low CPI Re (responsibility) scores tended to receive citizen complaints,
which included uses of unnecessary force. Sarchione, Cutler, Muchinsky, Nelson-Gray (1998) reported that low CPI scale scores on Re, So (socialization), and Sc (self-control) discriminated officers who received disciplinary action from those officers who did not. Reported job dysfunctions that led to discipline included using excessive force and inappropriate verbal conduct toward the public.

Scogin, Schumacher, Gardner, and Chaplin (1995) found that IPI scales Absence Abuse, Anxiety, Substance Abuse, Rigid Type, Critical Items, Undue Suspiciousness, Unusual Experience, and Sexual Concerns best predicted which officers would receive citizen complaints. The authors did not report on the types of citizen complaints filed against officers.

Much of the literature on personality tests used for predicting subsequent police job performance has linked tests scores with objective criteria such as disciplinary actions, absenteeism, and citizen complaints, or with subjective criteria such as supervisory performance ratings (Varela et al., 2004). Where disciplinary actions and citizen complaints are job performance criteria, they are often composite measures that might include uses of abusive aggression that authors report or sometimes fail to report. Generally, there is a lack of evidence that ties down specific psychological constructs of personality tests to certain measures of job work aggression (Grant & Grant, 1995).

Which police candidates are prone to unreasonable uses of aggression? The Independent Commission on the Los Angeles Police Department (ICLAPD; 1991) found that officers who had high rates of excessive force complaints also received superior supervisory performance ratings and psychologists rated them as suitable for
police work. If personality tests are contributing some knowledge about aggressive tendencies, then being prone to abuses of aggression might be more than a matter of measuring personality traits (Grant & Grant, 1995). Toch (1995) recognized that not all psychologically healthy officers are free from abuses of aggression. Situational factors might contribute to aggressive overreactions (Benner, 1986; Mills & Stratton, 1982). Abuses of aggression might be an artifact of attitudes and belief systems that develop after selection (ICLAPD, 1991) or be independent of them. Police experience and effects of the occupational culture might lead to job-related problems not predicted by candidates' psychological profiles. Megargee (1969) suggests that instigation, inhibition, and situational factors interact to determine some acts of aggression. In short, test responses, in part, help identify aggressive tendencies. The test data when coupled with the personal history (e.g., legal difficulties, physical altercations, other antisocial or unconventional tendencies) improve identification. Adding additional valid approaches may further enhance this process.

Thinking about Aggression Assessment

In this research, I propose rendering an aggression assessment of police candidates to determine their present aggressive tendencies in the management of potential job-related conflict. Consider that psychologists can reproduce important aggressive behavior patterns if they present hypothetical force situations to candidates and ask them to predict their performance. Candidates must use their experience to answer such questions when they have little direct knowledge on how to employ force in the work field. In creating hypothetical force situations, it is important to include conditions in which candidates need to use some force to manage job-related
problems. Most police use-of-force incidents result in arrests (Croft & Austin, 1987; as cited in Adams, 1995). Using arrests as a source of data about use of force allows psychological screeners to approximate work conditions in which police might use varying degrees of force against suspects. While suspects might respond to arrests by using firearms, knifes, kicks, punches, or profanity against the police, the police have a range of possible forceful responses that might include the use of police equipment such as firearms, batons, chemical agents, stun guns or the use of weaponless tactics such as arm bars, pressure points, or verbal commands. Police trainers usually represent these types of police and suspect behaviors along a continuum of force (Garner, Buchanan, Schade, & Hepburn, 1996; Garner & Maxwell, 1999; McLaughlin, 1992). The continuum categorizes behaviors and orders the categories on their relative severity. These gradations capture important variations in the types of force police and suspects might use in a given encounter. The continuum of force illustrates how police agencies conceptualize the measurement of aggression.

Designing an inventory of force situations that include a continuum of force (a) provides a set of behaviors that reveal levels of aggression, (b) gives psychological screeners guidelines for defining levels of force candidates might use in response to levels of suspect behaviors, and (c) allows screeners to define “excessive force” in terms of police training practices.

Of course, not all candidates are equal in the way they might use force to solve hypothetical arrest situations. Collyer, Gallo, and Boney-McCoy (2004) suggest that individuals may differ in two ways. First, there can be differences in the threshold adopted by two candidates for the use of a given response tactic. For example, if a
suspect begins to use profanity, one candidate might begin to employ physical responses; another candidate might require a physical action by the suspect before responding physically him/herself. Note that in this example, both candidates can be in agreement about the underlying ordered continua of force that apply to their own and the suspect's behavior; their disagreement is over where to "draw the line" (or threshold) with respect to physical tactics. Most candidates can probably be trained to have a common conception of the continuum, even if individual differences in thresholds remain.

A more significant kind of difference arises when candidates have different conceptions of the continuum of force itself – that is, when their ordering of actions by severity are not the same. For example, one candidate may rate a threat of physical force, such as shaking a stick, as more provocative than profanity. Another candidate (perhaps for personal or cultural reasons) may regard the profanity as more provocative, and even as the triggering event for a forceful response. In this case, there is a need for psychological screeners and police trainers to recognize that the two candidates have not internalized the same continuum of force.

I have been discussing what could be called provocation – the use of some level of force in response to a suspect's behavior. Another measurable aspect of a candidate's approach to force is labeling. Labeling refers to the way a police candidate assesses the severity of aggression of an action, not necessarily in a threatening situation, but generally. Collyer et al. (2004) found that some people regard profanity as a moderately violent behavior, while others rated profanity as "not violent at all." Again, there are two ways police candidates can differ in their labeling judgments: (1)
they may agree on the underlying order of behaviors with respect to severity of
ger
taggression, but differ in where they draw the line for applying labels such as
"aggressive" and "nonaggressive"; and (2) they may differ in their actual ordering of
behaviors by severity.

The data of Collyer et al. so far suggest that individuals have a common
conception of the relative severity of aggressive actions, and that this common
ordering by severity underlies both labeling and provocation judgments. They argue
that threshold differences predominate over differences in ordering, because the
correlations among individuals are high ($r=+0.85$); on the other hand, if it were the
case that differences in ordering predominated, these correlations would be low by
definition. One may speculate, however, that a shared understanding of how to order
behaviors may arise from common cultural experience in a homogeneous group, and
that differences in ordering may become more frequent when individuals come from
different backgrounds.

Are labeling and provocation judgments related? Collyer et al. (2004) observed
a relationship between college students' labeling and provocation ratings on instances
of violent actions, $r=+0.36$, $p<.01$. In a scatterplot, the paired values (individuals’
average ratings) using median splits on low and high labeling and provocation ratings
define four aggression types: individuals with high labeling and high provocation
ratings, high labeling and low provocation, low labeling and low provocation, and low
labeling and high provocation ratings. In this research, police candidates with low and
high labeling and provocation ratings will be sorted into the four aggression types:
1. HH – high labeling raters and high provocation raters. Candidates see many types of suspect force as aggression, and are easily provoked into being aggressive. Such candidates are more likely to use force when needed. Extreme HH types are more likely to use force in excessive ways.

2. HL – high labeling raters and low provocation raters. Candidates see many types of suspect force as aggression, but are difficult to provoke into being aggressive. Such candidates are more likely to be overly cautious in employing force.

3. LL – low labeling raters and low provocation raters. Candidates discount many types of suspect force as aggression, and are difficult to provoke into being aggressive. Such candidates are more likely to be permissive and have difficulty making decisions to use force.

4. LH – low labeling raters and high provocation raters. Candidates discount many types of suspect force as aggression, but are easily provoked into being aggressive. Such candidates are more likely to be unpredictable in their uses of force.

Distinguishing aggression types consists of using candidates’ labeling and provocation judgments on instances of suspect behaviors during hypothetical arrest situations. Those judgments might reflect four dispositional subtraits of aggression: “Physical and verbal aggression, which involve hurting or harming others, represent the instrumental or motor component of behavior. Anger, which involves physiological arousal and preparation for aggression, represents the emotional or affective component of behavior. Hostility, which consists of feelings of ill will and injustice, represents the cognitive component of behavior” (Buss and Perry, 1992, p. 457). Self-report measures such as questions on candidates’ past behavioral
expressions of these subclasses of aggression can alert police screeners to the ways in which candidates' experiences influence their labeling and provocation judgments.

Identifying a connection between candidates' behavioral experiences and their labeling and provocation judgments would be beneficial because: (a) candidates can reflect on their acquired attitudes and beliefs, and understand themselves better; (b) candidates can uncover prior behavior patterns that might lead to reasonable or unreasonable uses of aggression; (c) psychologists can use labeling and provocation performance data to cross-validate other personality test data sources; and (d) psychologists can give candidates, training officers, and hiring agencies formative feedback on relevant aggression issues and problem solving skills that should be worked on in training sessions.

In summary, this paper investigates police candidates' labeling and provocation responses to a set of constructed hypothetical force situations. It is specifically concerned with studying (a) candidates' underlying conception of a continuum of force; (b) the relationship between candidates' labeling and provocation judgments; (c) the theoretical relationship of candidates' labeling and provocation judgments with established measures of physical and verbal aggression, anger and hostility; and (d) the link between candidates' life events that might involve use of aggression and their labeling and provocation judgments. I discuss the potential applications of candidates' labeling and provocation judgments for use in police training sessions and both conditional and pre-conditional employment evaluations.
Chapter 2: Method

Participants and Samples

Participants were state and municipal police candidates who entered their agency-affiliated recruit training academies. State police candidates enter a six-month training academy for state police trainees only. The academy gives recruit training following a hiring process. Recruit classes might occur several years apart. The municipal police training academy provides fifteen weeks of recruit training for most city and town police candidates. They conduct recruit classes several times per year, generally fifteen weeks apart. The numbers of city and town police candidates that compose the classes can vary. Police agencies usually request class seats based on personnel attrition.

At the time of academy entrance, both state and municipal police candidates had passed general requirements (i.e., minimum education level of a high school degree or its equivalent, written exam, physical agility test, and an oral board interview) of their hiring agencies. All candidates had been through a psychological screening process; psychologists who conducted the evaluations judged them to be suitable for police work. Demographic data collected on police candidates included gender, racial group membership, age, and education achievement. The study consisted of four police recruit classes.

Recruit class 1. The first class consisted of 42 municipal police candidates. Thirty-nine (or 92.9%) recruits were male and three (or 7.1%) were female. The racial composition of the class was 38 (or 90.5%) White recruits, 2 (or 4.8%) Black, 1 (or
2.4%) Asian, and 1 (or 2.4%) Hispanic. On average, recruits were roughly 26 years of age and they had earned around 72 college credits.

**Recruit class 2.** The second class, which was state police candidates, consisted of 20 (or 83.3%) male and 4 (or 16.7%) female recruits. Sixteen (or 90.5%) recruits were White, 4 (or 4.8%) were Black, and 3 (or 2.4%) were Hispanic. The average age of recruits was around twenty-seven. Fourteen recruits (or 58.3%) had earned 120 or more college credits and 4 (or 22.2%) recruits had earned at least 60 but less than 120 college credits. Information on the academic achievement of six recruits was unavailable.

**Recruit class 3.** The third class consisted of 30 municipal police candidates who were all male. Most recruits were White (28 or 93.3%) and the remainder was Hispanic (2 or 6.7%). On average, recruits were roughly 27 years of age and they had earned around 86 college credits.

**Recruit class 4.** The fourth class, which was also municipal police candidates, was composed of 40 (or 83.3%) males and 8 (or 16.7%) females. Most recruits were White (45 or 93.8%), followed by Black (1 or 2.1%), Asian (1 or 2.1%), and Hispanic (1 or 2.1%). The average age of recruits was almost twenty-seven. Their average college credits earned was about ninety.

**Development of a Reaction Inventory-Force**

The Reaction Inventory – Force (RIF) was designed to measure the degree to which stimulus situations (or arrest encounters involving force) may reveal labeling and provocation tendencies. A careful review of the police literature did not uncover any uses of hypothetical arrest situations that might have provided a starting point, and
so the RIF is a new first-stage instrument. The method of inventory development involved generating suspect behaviors and police responses categorized along a continuum of force. A review of the items by police candidates and a panel of police experts served to assess the adequacy and meaningfulness of the inventory.

Item generation. A review of use of force continua from northeast U.S. police agencies, and a review of the literature on force by the police guided the selection of six categories of force that suspects might use against police during an arrest: (1) nonverbal resistance — the suspect’s intentional use of nonverbal behaviors that indicate his or her attitude, appearance, or physical readiness to resist the officer; (2) verbal resistance — the suspect’s intentional use of verbal responses that indicate his or her unwillingness to cooperate with the officer; (3) passive resistance — the suspect’s intentional use of physical actions not directed against the officer, with no intent to prevent the officer’s attempt to take control; (4) defensive resistance — the suspect’s intentional use of physical actions to escape, with no intent to cause harm to the officer; (5) assaultive resistance — the suspect’s intentional use of physical actions against the officer, with intent to cause harm to the officer; and (6) deadly force resistance — the suspect’s intentional use of physical actions or weapons against the officer, with intent to cause serious bodily harm or death to the officer. The categories provided a framework to organize different types of forceful behaviors by their relative degree of severity. Initial categorization was important because there is no one accepted configuration by police agencies (Garner et al., 1996). Category labels and definitions were sufficiently broad to include most items police professionals might suggest.
Initial item generation of behaviors within each category involved a content analysis of literature on force by and against the police and my police experience. A tentative item pool of 37 forceful behaviors was comprehensive enough to represent the six categories of suspect force (6 nonverbal resistance behaviors, 5 verbal resistance, 6 passive resistance, 6 defensive resistance, 7 assaultive resistance, and 7 deadly force resistance behaviors).

A five-person focus group of police trainers convened to help examine the properties of the proposed 37 behavior items: category assignment and relevance (i.e., high, moderate, or low), and vocabulary clarity and conciseness (Fowler, 1995). The group had a total of 86 years of police experience ($M = 17.2$). Members of the group discussed whether the range of behaviors and categories adequately mapped gradations of force by suspects. Results of the focus group's review showed difficulties with the wording of some behaviors. I used those results to revise the set of 37 behavior items.

A second five-person focus group of different police trainers reviewed the revised inventory of items. This group had a total of 72 years of police experience ($M = 14.4$). After suggestions for revisions, the final pool of 37 behavior items for administration was judged to be representative of the proposed six categories and their conceptual definitions.

Labeling task. I arranged the generated 37 suspect behaviors as part of a paper-and-pencil survey called “labeling task” (see Appendix A). Appendix B shows the constituent behavior items and their assigned force categories. The order of suspect behaviors consisted of non-repeated gradations of force. The labeling task asked
police candidates how much force they would associate with each of the 37 suspect behaviors used during an arrest. Candidates used a six-point Likert type response format, which ranged from no force to maximum force (1=no force, 2=low force, 3=moderate force, 4=intermediate force, 5=high force, 6=maximum force). A number and label assignment to each scale option can improve the reliability of the response task by providing a basis for discriminating between options (Converse & Presser, 1986; Fowler, 1995).

**Provocation task.** The task was similar to the labeling task, except that I framed the suspect behaviors as directed against candidates; the instructions asked candidates to give their opinions aloud on how they would respond to the behaviors; the response task involved timed-conditions; and a computer program (Macromedia Flash Movie) presented the behaviors (i.e., text items) on a projection screen. Appendix C gives the provocation task.

Using a timed-condition response procedure artificially increases states of arousal usually experienced in force situations. Under timed-conditions, I believe that candidates are more likely to operate on “automatic,” put little effort into being guarded or defensive with their responses, and use their experience to choose a response predictive of their response under real conditions.

**Field interview.** The goal of the field interview was to find out how potential respondents would understand the behavior items and respond to the hypothetical force situations. Available incumbent police officers (2 males) and police candidates (3 males) participated in a presurvey evaluation. They were members of the target population who might complete the actual labeling and provocation tasks.
constructed five pencil-and-paper short-forms of the provocation task in which the responses required no time constraints. Each short-form consisted of six different behaviors from all categories of suspect force. Short-forms are less demanding and allow respondents to recall and elaborate on how they arrive at their responses (Sudman & Bradburn, 1982). Appendix D gives an example of one short-form.

I chose the provocation task for the field interview because there is a level of threat in asking about candidates' aggressive behavior that may reveal some undesirable characteristics. Clearly, it is more difficult for inexperienced police candidates to answer questions about using aggression than to answer nonthreatening questions about labeling severity.

Procedures for the presurvey evaluation consisted of having each respondent go twice through a different short-form: respondents first answered the behavior items in the usual way, and then discussed the process they used for answering each item (Forsyth & Lessler, 1991). The question-by-question review followed a standard interview protocol (Fowler, 1995). Appendix E provides the protocol used.

Individual interview sessions lasted roughly 45 minutes. Officers and police candidates expressed that both the instructions and the response scale format were clear and succinct. They agreed that what force meant was not limited to physical actions or the use of weapons, but included officer presence and verbal techniques. Officers and candidates pointed out that they were able to discriminate well between forceful response categories. Their examples of behaviors linked to response categories confirmed clear gradations along the severity dimension. They were using the response categories correctly.
In discussing each behavior item with the incumbent police officers, they found that suspect behaviors were obvious. They went about predicting their responses by using prior experiences with suspects. Officers expressed being very confident when giving their responses. Similarly, police candidates found the suspect behaviors clear. They spoke about visualizing suspect behaviors during arrests, and constructing some strategy from experience (e.g., personal or media) to predict their responses. For example, one candidate mentioned having been punched and used this experience to help him decide on an appropriate response. Police candidates said they were very confident when giving their responses. As expected, for both incumbent officers and police candidates, there was a convergence between their experience and response choices when forecasting their performance in hypothetical force situations. In light of the interview sessions and focus group discussions, I administered the final forms of the labeling and provocation tasks to samples of police recruits.

**Concurrent Validation**

How do labeling and provocation tendencies behave relative to previously established measures of physical aggression (PA), verbal aggression (VA), anger (ANG), and hostility (HO)? Police recruits responded to a set of criterion measures for use in evaluating the concurrent validity of labeling and provocation tendencies.

*Recruit classes 1, 2, and 3.* For criterion measures of PA, VA, ANG, and HO, I used the Buss-Perry short-form aggression questionnaire (AQ short-form; Bryant & Smith, 2001). It is a psychometrically refined 12-item subset of Buss and Perry's (1992) original questionnaire. Subscales of PA, VA, ANG, and HO compose the AQ short-form. Cross-validation procedures have linked these subscales with measures of
aggression by incumbent law enforcement officers (Greenberg, Riggs, & Bryant, 2003). Appendix F gives the AQ short-form.

Recruit class 4. Police recruits completed the AQ short-form. They also filled out an additional set of criterion measures. The Physical Aggression subscale (see Appendix G) from Buss and Perry's original Aggression Questionnaire (AQ long-form; 1992), the Aggressiveness subscale (see Appendix H) from the Verbal Aggression Scale (VAS; Infante & Wigley, 1986), the Anger Arousal subscale (see Appendix I) of the Multidimensional Anger Inventory (MAI; Siegel, 1986), and the Cook-Medley Hostility Scale (Cook & Medley, 1954; see Appendix J) served as standards for evaluating PA, VA, ANG, and HO respectively.

The Buss-Perry Aggression Questionnaire (1992) is a widely used self-report measure of different dimensions of aggression having correlations with other aggression measures (Suris, Lind, Emmett, Borman, Kashner, & Barrat, 2004). Infante and Wigley's (1986) Verbal Aggression Scale is a frequently used valid measure of trait verbal aggression (Beatty, Rudd, & Valencic, 1999; Levine, Beatty, Limon, Hamilton, Buck, & Chory-Assad, 2004). Recent research suggests that the scale is bi-dimensional having subscales of aggressiveness and benevolence (Beatty et al., 1999; Levine et al., 2004; Suzuki & Rancer, 1994). Levine et al. (2004) recommend using the Aggressiveness subscale, which comprises 10 aggressively worded items, as a conceptually valid measure of verbal aggression.

The definition of the Anger Arousal subscale from the MAI involves physiological arousal, which fits Buss and Perry's (1992) definition of anger (Bryant & Smith, 2001). Cook-Medley's Hostility Scale is a frequently used self-report
measure of hostility, which has shown to be a valid predictor of a person's physiological and interpersonal functioning (Conrada & Jussim, 1992). Bryant and Smith (2001) reported significant correlations of Anger Arousal scores from the MAI and Cook-Medley Hostility Scale scores with AQ short-form subscale scores on ANG and HO respectively.

Survey Schedule

Recruit classes 1, 2, and 3. There were preliminary decisions made regarding the collection of data. First, recruits should be relatively naive to police endorsed training practices. When recruits have little knowledge on how to use force, they must use prior experiences; their answers will better reflect what they might say or do in force situations (Fowler, 1995; Poland, 1978; Smith & Klein, 1984). Second, a time interval between survey sessions should be long enough to eliminate or reduce sources of extraneous variability particularly carryover effects. Survey procedures for recruit classes 1, 2, and 3 followed a three-day data collection schedule, which began before recruits received extensive training on uses of force.

1. Day 1: In a group session, recruits completed the AQ short-form. It took them roughly 5 minutes to complete.

2. Day 2 (four days from day 1): In a group session, recruits completed the labeling task. It took roughly 15 minutes to complete.

3. Day 3 (four days from day 2): In single-participant sessions, recruits completed the provocation task. It took each recruit roughly 6 minutes to complete. Following the task, recruits completed the Marlowe-Crowne 13-item short form social desirability questionnaire (Reynolds, 1982) because they might lean toward making
favorable impressions in an attempt to appear well suited for police work. The questionnaire is psychometrically superior (Zook & Sipps, 1985; Reynolds; 1982; Silverstein, 1983) to most other short-form alternatives to the original Marlowe-Crowne scale (Crowne & Marlowe, 1960). Appendix K gives the Marlowe-Crowne scale. Scoring the scale entailed summing incorrect scores (correct response = 0, incorrect response = 1). At the end of each test session, I asked the recruit not to discuss the survey with other recruits.

Recruit class 4. The survey schedule closely matched the schedule for recruit classes 1, 2, and 3 except that on Day 1 recruits filled out the above set of additional criterion measures including the behavioral experience questionnaire (BEQ; see Appendix L). The questionnaire revealed self-report information about different kinds of life events that might involve aggressive behavior. Questions were drawn from clinical and life history questionnaires and were used to obtain recruits' background information as a source for describing labeling and provocation tendencies.

Analyses

An exploratory data analyses examined whether I could combine data sets collected from recruit classes into a single data set to obtain improved variability estimates. Researchers must be careful to avoid combining data sets that resemble "apples and oranges." There was no expectation that recruit classes were coming from dissimilar pools of police applicants that would require separate analyses.

Suspect behavior as unit of analysis. The correlation of average behavior ratings for the labeling task with those for the provocation task evaluated the extent of recruits' agreement on an ordered continuum of suspect behaviors. I expected a high
correlation between the average behavior ratings on the two tasks, which would signify a stable underlying stimulus dimension of severity.

How many categories are necessary to map an ordered dimension of severity? A principle factor analysis (PFA) with Varimax rotation was performed on the sample inter-item correlation matrix. The data, 6-point scaled items, has a quasi-continuous quality necessary for using analytic techniques such as factor analysis (Floyd & Widaman, 1995). Because PFA accounts for only covariation among variables, one can test certain hypotheses about whether recruits have an underlying conception of ordered continua (Fabrigar, MacCallum, Wegener, & Strahan, 1999). I preferred PFA for uncovering the nature and number of factors needed to understand the pattern of relationships in the data. The scree test, parallel analysis procedure, and substantive subject matter sense helped determine the number of factors to retain. To interpret the meaning of factors by the variables that correlate with them, I chose a minimum variable loading of .40.

To check whether recruits' average behavior ratings differed from a test model of conservative responses, I constructed an ordered continuum of force for comparison. The philosophy underlying the generated values was a "One-for-One" concept in which recruits would select levels of force that paralleled levels of resistance. Values for items constituting the categories nonverbal and verbal resistance, passive resistance, defensive resistance, assaultive resistance, and deadly force resistance were 2, 3, 4, 5, and 6 respectively. I used a one-sample t-test procedure. The numeric test value (or average rating) was 3.81 (see Appendix M).
Individual recruit as unit of analysis. The correlation of recruits' average ratings for the labeling task with those for the provocation task assessed whether there was a strong relationship between them. I expected that knowing recruits' labeling ratings would provide some information about their provocation ratings and vice versa.

The correlation of recruits' average labeling and provocation ratings with those average ratings for selected criterion measures of PA, VA, ANG, and HO measured how recruits' labeling and provocation ratings behaved relative to other established measures of aggression. The scoring strategy for the Cook-Medley Hostility Scale used recruits' summated scores. The behavioral experience questionnaire served as a count index of some actual life events that might involve recruits' use of aggression. I explored whether there was a convergence between recruits' life history and their labeling and provocation ratings.
Exploratory Data Analysis

Recruit responses to the labeling and provocation tasks were initially screened for missing value patterns. There were no items on the labeling task with 5% or more missing values. On the provocation task, there was one item with 5% or more missing: item 2, “suspect runs out of their house away from me” \((n = 7)\). Twenty-four recruits (17.1%) did not respond (within the 3 second interval) to a range of suspect behaviors \((n = 38)\) at arrest. Eighty-four percent (or 32) of those behaviors involved suspects using defensive resistance (16), passive resistance (7), nonverbal resistance (7), or verbal resistance (2) against recruits. Some recruits had difficulty responding to nonphysical directed acts of resistance. For missing values, cases were excluded analysis-by-analysis (pairwise deletion).

On substantial subject matter grounds, the way police agencies selected candidates for academy training was comparable: candidates had a minimum education level of a high school degree or its equivalent and had passed a written exam, physical agility test, oral board interview, and a battery of psychological tests. The onset of data collection from police recruit classes took place at different times in the course of the research: day 1, recruit class 1; day 4, recruit class 2; day 105, recruit class 3; and day 252, recruit class 4. No apparent historical events as sources of extraneous variability were associated with any of the recruit classes. On statistical grounds, diagnostics to determine if the data sets could be combined as a single set involved obtaining descriptive statistics by recruit classes, testing the normality of the
data, testing the equality of variances in recruit classes, and conducting ANOVAs on
the labeling and provocation tasks.

Table 1 provides descriptive statistics for the labeling and provocation tasks by
recruit classes as well as for the pooled data set. The table values appear to signal that
the samples of recruit classes are coming from similar distributions. Testing the
normality of the data for both labeling and provocation tasks by recruit classes
involved checking the Normal Q-Q plots of both the raw data and standardized
residuals and conducting Shapiro-Wilk (SW) tests of non-normality. Visual inspection
of the Q-Q plots showed only small discrepancies between the observed data and
expected normal values for both tasks. Table 2 provides SW test values, which
indicated no significant departures from normality.
### Table 1

*Descriptive Statistics for Labeling and Provocation Tasks by Recruit Classes*

| Recruit Class | $n$ | $M$ | $Mdn$ | $SD$ | $Min$ | $Max$ | $Q1$ | $Q3$ |
|---------------|-----|-----|-------|------|-------|-------|------|------|
| **Labeling Task** |     |     |       |      |       |       |      |      |
| 1 Municipal   | 42  | 3.82| 3.74  | 0.67 | 2.49  | 5.35  | 3.38 | 4.21 |
| 2 State       | 23  | 3.84| 4.14  | 0.70 | 2.54  | 5.22  | 3.19 | 4.30 |
| 3 Municipal   | 30  | 3.53| 3.43  | 0.58 | 2.32  | 4.81  | 3.05 | 3.97 |
| 4 Municipal   | 48  | 3.55| 3.54  | 0.41 | 2.84  | 4.54  | 3.26 | 3.84 |
| Pooled Classes| 143 | 3.67| 3.59  | 0.59 | 2.32  | 5.35  | 3.22 | 4.11 |
| **Provocation Task** |     |     |       |      |       |       |      |      |
| 1 Municipal   | 41  | 3.60| 3.62  | 0.67 | 2.43  | 5.65  | 3.08 | 4.01 |
| 2 State       | 23  | 3.69| 3.89  | 0.74 | 2.44  | 5.05  | 3.05 | 4.19 |
| 3 Municipal   | 28  | 3.48| 3.44  | 0.57 | 2.54  | 4.97  | 3.12 | 3.76 |
| 4 Municipal   | 48  | 3.37| 3.34  | 0.40 | 2.76  | 4.38  | 3.08 | 3.61 |
| Pooled Classes| 140 | 3.51| 3.46  | 0.59 | 2.43  | 5.65  | 3.08 | 3.89 |
Table 2

*Shapiro-Wilk Tests for Labeling and Provocation Tasks by Recruit Classes*

| Recruit Class | Statistic | df  | p    |
|---------------|-----------|-----|------|
| 1 Municipal   | 0.98      | 42  | 0.65 |
| 2 State       | 0.96      | 23  | 0.44 |
| 3 Municipal   | 0.98      | 30  | 0.70 |
| 4 Municipal   | 0.98      | 48  | 0.53 |
| Pooled Classes| 0.99      | 143 | 0.25 |

| Recruit Class | Statistic | df  | p    |
|---------------|-----------|-----|------|
| 1 Municipal   | 0.97      | 41  | 0.29 |
| 2 State       | 0.96      | 23  | 0.52 |
| 3 Municipal   | 0.96      | 28  | 0.28 |
| 4 Municipal   | 0.96      | 48  | 0.06 |
| Pooled Classes| 0.99      | 140 | 0.14 |
Testing the equality of variances in the labeling and provocation data sets from recruit classes entailed checking the side-by-side boxplots of both the raw data and standardized residuals and calculating Hartley's $F_{\text{MAX}}$ test for homogeneity of variances. Box and Whisker plots of the data suggested that there is an overlap of variances and few outliers. Values for $F_{\text{MAX}}$ (labeling, 2.88; provocation, 3.43) indicated relatively small degrees of differences between recruit class data sets. Within the limits of sampling error, results favored equal variances in recruit classes.

The last step to determine whether the recruit classes were a homogenous set, permitting combination of the classes, was to conduct separate ANOVAs for the labeling and provocation tasks. In this research, however, the recruit class sizes were unequal. Consequently, the actual discrepancy in variances might be magnified, which would affect the probability of making a Type 1 error (Keppel, 1991). To correct for inflated variance heterogeneity, I applied a more conservative significance level of $\alpha = .025$ as the criterion value for ANOVA tests. On the labeling task, there was no statistical difference among recruit classes at the $p > .025$ level ($F(3, 139) = 2.83$). Likewise, there was no statistical difference on the provocation task by recruit classes at the $p > .025$ level ($F(3, 136) = 1.93$).

In summary, a review of the exploratory data analyses suggested that the recruit class data sets appear homogenous and that I could analyze them as a single set.

**Suspect Behavior as Unit of Analysis**

*Labeling task.* Appendix M gives the average severity ratings and corresponding ranks of the 37 suspect behaviors. An intraclass correlation (ICC), two-
way mixed effect model (consistency definition), revealed that the extent of consensus on rating the severity of suspect behaviors was excellent, alpha = 0.99. Figure 1 shows the average ratings of severity (mean average rating = 3.67) plotted against the rank position corresponding to each behavior as listed in Appendix M. “Suspect swears at officer” anchors the far left of the scale, and “suspect fires a handgun at officer” anchors the far right. Also shown are two individual labeling tendencies (individual recruit as unit of analysis), one for a “high rater” \( M = 5.35 \), and one for a “low rater” \( M = 2.49 \). The correlation of average behavior ratings (or vertical measures) for the labeling task with those for the provocation task was almost perfect \( r = +0.99, n = 37, p < .01 \) 1-tailed.

Diagnostics on the labeling data showed that the suspect behavior items were suitable for factor analysis: the \( 37 \times 37 \) sample inter-item correlation matrix showed evidence of coefficients greater than 0.30; the Kaiser-Meyer-Olkin index was 0.90; and Bartlett’s test of sphericity was significant, \( \chi^2 (666) = 4016.11 \) at \( p < .01 \).

On subject matter and statistical grounds, a five-factor solution best explained the pattern of interrelatedness among behavior items along a severity dimension. Retained factors had moderate to high saturation levels (.60 and .80) with five or more behavior items per factor, which suggested that the sample size was sufficient to obtain a stable factor pattern that approximated the population pattern (Guadagnoli & Velicer, 1988). All item loadings averaged 0.65 on their respective factors, whereas loadings on other factors averaged 0.17 (see Appendix N). The five factors explained 85% of the estimated common variance: an acceptable percentage of the total initial communality estimates of the measured items (Floyd & Widaman, 1995). I used an
item's largest correlation with other items as its initial communality estimate. My choices of factor labels summarized the apparent severity meaning of item loadings and associations along a severity dimension: low, moderate, intermediate, high, and maximum. The average mean ratings of behavior items for these factors were 1.95 (low), 2.71 (moderate), 3.35 (intermediate), 4.39 (high), and 5.89 (maximum).

A one-sample t-test found recruits’ average behavior ratings ($M = 3.67, SD = 1.45$) were not significantly different from the One-for-One comparison model of conservative responses ($M = 3.81$), $t(36) = -0.58$ at $p > 0.05$ 2-tailed (95% confidence interval = [-.62, .34]). Figure 2 shows recruits’ average severity ratings of suspect behaviors plotted against the One-for-One model.

_Provocation task._ Appendix M gives the average responses to provocation and corresponding ranks of the 37 suspect behaviors. Ranks anchoring the top and bottom portions of the list were very similar to the severity ranks. Using the same abscissa as in Figure 1, Figure 3 shows the average responses to provocation. Also shown are two individual provocation tendencies (individual recruit as unit of analysis), one for a high rater ($M = 5.65$), and one for a low rater ($M = 2.43$).
Figure 1. Average severity ratings of suspect behaviors by their corresponding severity rank. Also shown are two individual labeling tendencies.
Figure 2. Average severity ratings of suspect behaviors plotted against the One-for-One comparison model of conservative responses.
Figure 3. Average responses to provocation plotted against the severity rank position corresponding to each suspect behavior (same abscissa as in Figure 1). Also shown are two individual provocation tendencies.
Individual Recruit as Unit of Analysis

The correlation of individuals' average ratings for the labeling task with those for the provocation task was very strong \( r = +0.87, n = 139, p < .01 \) 1-tailed. One could predict very well a recruit's provocation ratings from knowing his or her labeling ratings and vice versa, \( r^2 = 0.76 \). Figure 4 shows recruits’ provocation ratings (or average responses to suspect behaviors) plotted by their labeling ratings (or average severity ratings of suspect behaviors). Using the scale score of 3 to mark the point at which recruits begin to respond with high ratings, the data defined four aggression types: recruits with high labeling and high provocation ratings (HH; \( n = 106 \)), high labeling and low provocation (HL; \( n = 17 \)), low labeling and low provocation (LL; \( n = 13 \)), and low labeling and high provocation ratings (LH; \( n = 3 \)).

Recruits' labeling and provocation ratings for group separation were as follow: HH = labeling > 3 and provocation > 3; HL = labeling > 3 and provocation ≤ 3; LL = labeling ≤ 3 and provocation ≤ 3; and LH = labeling ≤ 3 and provocation > 3. There were five missing cases.

There was a convergence between recruits' life events and their aggressive tendencies. Appendix O shows recruits having high labeling and high provocation ratings (HH) had a life history of experiences that involved both the use of aggression and possible use of aggression. Although recruits of the HL, LL, and LH aggression types showed some history of aggressive behaviors, further study of these types and linkage to actual life events require more occupants for each type.

Table 3 gives the relation of recruits' labeling and provocation ratings with other previously established measures of aggression. Labeling and provocation ratings
were significantly tied to criterion measures of physical aggression (Buss-Perry AQ long-form physical aggression subscale), anger (Buss-Perry AQ short-form anger subscale, Multidimensional Anger Inventory anger arousal subscale) and verbal aggression (Verbal Aggression Scale aggressiveness subscale).

Scores on the Marlowe-Crowne social desirability scale exposed recruits' tendencies for giving guarded responses to appear more acceptable or desirable for police work, $M = 10.19$. A mean score of 13 would have indicated that recruits were extreme in a way that favored making a good impression, but a mean score of zero would have indicated that recruits were not motivated to "fake good."
Figure 4. Recruits' provocation ratings plotted by their labeling ratings.
Table 3

Correlations of Labeling and Provocation Ratings with other Measures

|                | Labeling | Provocation |
|----------------|----------|-------------|
| PA₁            | 0.06     | 0.00        |
| VA₁            | 0.08     | 0.12        |
| AN₁            | 0.17*    | 0.22**      |
| HO₁            | 0.10     | 0.13        |
| PA₂            | 0.40**   | 0.25*       |
| VA₂            | 0.32*    | 0.24        |
| AN₂            | 0.32*    | 0.27*       |
| HO₂            | 0.18     | 0.13        |

Note. PA₁ = Buss-Perry AQ short-form physical aggression subscale; VA₁ = Buss-Perry AQ short-form verbal aggression subscale; AN₁ = Buss-Perry AQ short-form anger subscale; HO₁ = Buss-Perry AQ short-form hostility subscale; PA₂ = Buss-Perry AQ long-form physical aggression subscale; VA₂ = Verbal Aggression Scale aggressiveness subscale; AN₂ = Multidimensional Anger Inventory anger arousal subscale; HO₂ = Cook-Medley Hostility Scale; * signifies p < .05; ** signifies p < .01; all tests were 1-tailed; PA₁, VA₁, AN₁, and HO₁ results based on 143 cases; and PA₂, VA₂, AN₂, and HO₂, results based on 48 cases.
Chapter 4: Discussion

This paper proposes a self-report approach to screen police candidates' aggressive tendencies. The research finds some groundwork evidence that supports the utility of the approach in both conditional and pre-conditional police employment evaluations.

Who are the stellar officers? What are their characteristics and how do psychological screeners measure them? While there may be some disagreement on the characteristics of an ideal officer, police stakeholders would agree that there is no place in police work for a candidate who fails to demonstrate a willingness to be aggressive. Force is an essential aspect of policing. Psychological screeners recognize this important job task. Their measurement methods must be sensitive to capturing candidates' ability to (a) recognize violent citizen behavior and (b) manage such behavior with reasonable responses that might involve different degrees of force. The idea of measuring recruits' labeling and provocation judgments on instances of suspect behaviors during hypothetical arrest situations shows evidence of selecting-in these two desirable police characteristics. Figure 4's visual impression suggests that police agencies and psychologists are successful at selecting-in candidates' who are able to detect aggression and who are responsive to directed aggression: even if their evaluation methods use a screen-out strategy. Police agencies and psychologists usually see in hindsight (e.g., candidates' on-the-job performance) their success at accepting candidates who have such wanted traits. However, both the labeling and provocation tasks are sensitive to measuring this sought-after candidate profile.
Perceptions of Aggression

Labeling and provocation judgments are facets of recruits' behavior in which there is agreement on an underlying dimension of perceived aggression (or behavior severity). Average ratings of suspect behaviors for the labeling task with those for the provocation task correlate almost perfectly. When responding to situations such as arrests that might require some degree of coercive action by recruits, they see in like ways the severity of different suspect behaviors at arrest. They also tend to see clusters of suspect behaviors along a severity dimension. A factor analysis revealed that recruits group together different suspect behaviors that they perceive to be related (see Appendix N). Low, moderate, intermediate, high, and maximum labels are conceptually appropriate to describe the pattern of behavior associations in terms of severity or relative degree of potential injury to the recruit. Recruits' sketch of suspect behaviors along this severity dimension, however, raises some concern for police trainers.

Recruits grouped what police experts and trainers would consider a collection of dissimilar behaviors. For example, recruits saw suspects who raised their arms and made fists, clenched their fists, or stood in fighting stances as displaying the same level of threat as those suspects who pushed, kicked, or punched. Although nonverbal types of behavior might serve as preparatory cues of active resistance, qualitatively and quantitatively they might call for different responses. We can see complex factor loadings where suspect behavior items correlate with more than one factor. For example, recruits thought suspects who fold and lock their arms demonstrate the same willingness to avoid arrest as those suspects who shout and curse. We also see cross
factor loadings for the suspect behavior “hitting neck with baseball bat,” signaling recruits’ insensitivity to threat level.

Factor analysis procedures uncover recruits’ insensitivity to the finer distinctions of some potential citizen behaviors during a foreseeable task of policing. Even so, psychological screeners can arm police trainers with such information so that recruits receive training on police continua and avoid using responses that are physical where verbal ones may be reasonable alternatives.

The labeling task is sensitive to detecting police recruits’ conception of behavior severity. Overall, recruits’ average ratings of behavior severity are not very different from data generated for a One-for-One comparison model of conservative responses (see Figure 2). Recruits’ average behavior ratings can serve equally well as a comparison model against which to test individual labeling and provocation differences.

*Threshold Measures*

How sensitive are recruits at detecting the severity level of behaviors suspects might use during arrest situations? What is the minimum amount of suspect resistance needed to trigger a forceful response? Given that recruits have a commonly understood scale of behavior severity, psychological screeners can treat it as a stable stimulus property. They can observe and speak about where a recruit begins to detect differences in behaviors located along the severity scale and where the recruit begins to respond to behaviors by using tactics that are more forceful. Average labeling and provocation ratings represent the best estimates of where recruits draw these lines (or thresholds) with respect to detecting aggression and using physical tactics. Average
acts of aggression. The data on which I argue the use of typologies is not complete, but is suggestive of certain kinds of candidates that emerge from the data. My proposed typologies are a conceptual speculation informed by police practices, informed and limited by empirical evidence.

Methodological Conclusions, Limitations, and Future Directions

This paper proposes an instrument with which psychological screeners can derive meaningful measures of aggression and their interpretations that can assist in police selection and training. Screeners may use the instrument for conditional and pre-conditional employment evaluations. Labeling and provocation ratings are easily measurable using user-friendly statistical methods. Outcome measures direct screeners attention to candidates' emerging patterns of aggressive tendencies. Typologies of low and high labeling and provocation ratings offer an organizational framework having analytical utility. Threshold measures give screeners a visual impression of where candidates begin to see behaviors as aggressive and where they begin to respond with behaviors that are more forceful.

Labeling and provocation tasks are sensitive to identifying patterns of aggressive tendencies that other tests might miss. For example, police candidates may respond cautiously or in a guarded fashion to a battery of psychological tests. They may try to uncover what the test questions appear to be measuring and then respond with answers that create a favorable impression. Candidates are careful not to reveal any undesirable characteristics that might exclude them from pre-employment consideration. Inflated “lie scores” are common. Under post-employment circumstances, this paper finds evidence that hired candidates have a tendency to “fake
good." On average, police recruits scored 10.19 on the Marlowe-Crowne 13-item short-form social desirability scale. This apparent unavoidable social desirability strategy can truncate psychological test data and attenuate predictive relations, and thus, the labeling and provocation ratings might have exhibited stronger relations with criterion measures of AN, PA, VA, and HO.

Under the labeling and provocation test situation, given candidates’ predilections for self-protection and social approval, they would be happy to give correct responses, but they are relatively unaware of police force practices. Candidates recognize police work requires some use of aggression, but the extent to which their choices of aggression against hypothetical vignettes might be reasonable is the test dilemma. Without formal police training, candidates must use their experience to inform them what to do. Consequently, their responses to the labeling and provocation tasks represent the best estimates of how they might actually perform in the work field. Details of candidates’ prior life events involving aggression and psychological test data sources complete the picture of fitness or unfitness to perform enforcement tasks.

Police professionalism implies screening the fitness of candidates to manage force events. The HH, HL, LL, and LH typology framework is useful as a way of directing psychological screeners’ attention to patterns of aggressive tendencies revealed in the test data. Each aggression type is specifiable by the perceptions and by the behaviors of its occupants. Although the typology framework has screening utility, its theoretical justification requires further empirical investigation. Predictive validation procedures allow screeners to render an at-risk assessment of eventual uses
of job-related aggression. Such procedures are best when screeners use longitudinal studies (Beutler, Nussbaum, & Meredith, 1988; and Bartol, 1991). Future work involves studying how police recruits' labeling and provocation ratings behave relative to job performance data such as academy class rank, department disciplinary action, and citizen complaints of verbal discourtesy and excessive physical force. Forming connections between the different combinations of labeling and provocation ratings and job performance measures will begin to round out the theoretical utility (construct and criterion validity) of distinguishing aggression types. Well-populated aggression types might emerge through continuing data collection: Further analytical descriptions of the aggression types and better discrimination among them are possible. At present, the typology framework is tentative.

The foregoing look at screening police candidates' aggressive tendencies using a labeling and provocation task is encouraging. My proposal has some empirical support and practical justification. The self-report screening approach provides aggression measurements that are very useful to police practitioners. Future work should provide estimates of the predictive validity of the labeling and provocation framework in police employment evaluations. Future directions may well include measuring provocation responses by alternative methods such as a paper-and-pencil test, computer test, and an interactive video situational test.
Appendices

Appendix A: Labeling Task

Using the 6 point scale shown below, indicate how much force you think you would associate with each of the following suspect behaviors during an arrest. Place your rating in the response space to the right of the behavior. There are no right or wrong answers.

1 = No Force  
2 = Low Force  
3 = Moderate Force  
4 = Intermediate Force  
5 = High Force  
6 = Maximum Force

1 Suspect fires a handgun at officer  
2 Suspect runs out of their house away from officer  
3 Suspect screams at officer  
4 Suspect grabs officer and throws officer to the ground  
5 Suspect stands in a fighting stance toward officer  
6 Suspect folds and locks their arms  
7 Suspect scratches officer’s face  
8 Suspect yells at officer  
9 Suspect pulls away from officer  
10 Suspect raises their arms and makes a fist toward officer  
11 Suspect fires a shotgun at officer  
12 Suspect wraps their arms around the car steering wheel and holds on  
13 Suspect punches officer  
14 Suspect shouts and curses at officer  
15 Suspect flees in a car from officer  
16 Suspect fires a rifle at officer  
17 Suspect grabs around a pole and holds tightly  
18 Suspect kicks officer  
19 Suspect clenches their fists at officer  
20 Suspect sits in a chair and tucks their arms tightly  
21 Suspect pushes officer to the ground  
22 Suspect drives a car at officer to hit officer  
23 Suspect aggressively points their finger at officer  
24 Suspect flees on foot from officer  
25 Suspect stabs officer with a knife  
26 Suspect flees on a bicycle from officer  
27 Suspect swears at officer  
28 Suspect slaps officer’s face  
29 Suspect hits officer’s neck with a baseball bat  
30 Suspect lays on the ground and stifens their body  
31 Suspect clenches their hands and stares at officer  
32 Suspect falls to the ground and curls their arms under their body  
33 Suspect strikes officer with their elbow  
34 Suspect raises their voice and argues with officer  
35 Suspect strikes officer’s head with a baton  
36 Suspect jumps out of their car and runs away from officer  
37 Suspect glares at officer and clenches their teeth

Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _
Response _ _ _

Appendix B: Items Constituting the Labeling and Provocation Tasks

| Category               | Constituent items                                                                 |
|------------------------|-----------------------------------------------------------------------------------|
| **Nonverbal Resistance** | 5 Suspect stands in a fighting stance toward officer                                |
|                        | 10 Suspect raises their arms and makes a fist toward officer                         |
|                        | 19 Suspect clenches their fists at officer                                           |
|                        | 23 Suspect aggressively points their finger at officer                                |
|                        | 31 Suspect clenches their hands and stares at officer                                 |
|                        | 37 Suspect glares at officer and clenches their teeth                                 |
| **Verbal Resistance**  | 3 Suspect screams at officer                                                         |
|                        | 8 Suspect yells at officer                                                          |
|                        | 14 Suspect shouts and curses at officer                                              |
|                        | 27 Suspect swears at officer                                                        |
|                        | 34 Suspect raises their voice and argues with officer                                 |
| **Passive Resistance** | 6 Suspect folds and locks their arms                                                 |
|                        | 12 Suspect wraps their arms around the car steering wheel and holds on                |
|                        | 17 Suspect grabs around a pole and holds tightly                                     |
|                        | 20 Suspect sits in a chair and tucks their arms tightly                               |
|                        | 30 Suspect lays on the ground and stiffens their body                                 |
|                        | 32 Suspect falls to the ground and curls their arms under their body                  |
| **Defensive Resistance** | 2 Suspect runs out of their house away from officer                                  |
|                        | 9 Suspect pulls away from officer                                                    |
|                        | 15 Suspect flees in a car from officer                                                |
|                        | 24 Suspect flees on foot from officer                                                |
|                        | 26 Suspect flees on a bicycle from officer                                           |
|                        | 36 Suspect jumps out of their car and runs away from officer                          |
| **Assaultive Resistance** | 4 Suspect grabs officer and throws officer to the ground                               |
|                        | 7 Suspect scratches officer’s face                                                   |
|                        | 13 Suspect punches officer                                                           |
|                        | 18 Suspect kicks officer                                                             |
|                        | 21 Suspect pushes officer to the ground                                              |
|                        | 28 Suspect slaps officer’s face                                                      |
|                        | 33 Suspect strikes officer with their elbow                                          |
| **Deadly Force Resistance** | 1 Suspect fires a handgun at officer                                                |
|                        | 11 Suspect fires a shotgun at officer                                               |
|                        | 16 Suspect fires a rifle at officer                                                 |
|                        | 22 Suspect drives a car at officer to hit officer                                     |
|                        | 25 Suspect stabs officer with a knife                                                |
|                        | 29 Suspect hits officer’s neck with a baseball bat                                   |
|                        | 35 Suspect strikes officer’s head with a bat                                         |
Appendix C: Provocation Task

Instructions:

Please stand and remain standing six feet from the projection screen during this portion of the survey. We have provided a six-foot floor marker for you.

Using the 6 point scale shown below, indicate how you think you would respond to each of the following suspect behaviors during an arrest.

Each suspect behavior will appear on the screen. You will have only three (3) seconds to read the behavior.

Following each suspect behavior, the 6 point scale shown below will appear on the screen. You will have only three (3) seconds to choose and state your response aloud. There are no right or wrong answers.

Following the 6 point scale, a blank screen will appear. You will have only three (3) seconds to prepare for the next suspect behavior. There are thirty-seven (37) suspect behaviors.

1 = No Force
2 = Low Force
3 = Moderate Force
4 = Intermediate Force
5 = High Force
6 = Maximum Force
1 Suspect fires a handgun at me  
2 Suspect runs out of their house away from me  
3 Suspect screams at me  
4 Suspect grabs me and throws me to the ground  
5 Suspect stands in a fighting stance toward me  
6 Suspect folds and locks their arms  
7 Suspect scratches my face  
8 Suspect yells at me  
9 Suspect pulls away from me  
10 Suspect raises their arms and makes a fist toward me  
11 Suspect fires a shotgun at me  
12 Suspect wraps their arms around the car steering wheel and holds on  
13 Suspect punches me  
14 Suspect shouts and curses at me  
15 Suspect flees in a car from me  
16 Suspect fires a rifle at me  
17 Suspect grabs around a pole and holds tightly  
18 Suspect kicks me  
19 Suspect clenches their fists at me  
20 Suspect sits in a chair and tucks their arms tightly  
21 Suspect pushes me to the ground  
22 Suspect drives a car at me to hit me  
23 Suspect aggressively points their finger at me  
24 Suspect flees on foot from me  
25 Suspect stabs me with a knife  
26 Suspect flees on a bicycle from me  
27 Suspect swears at me  
28 Suspect slaps my face  
29 Suspect hits my neck with a baseball bat  
30 Suspect lays on the ground and stiffens their body  
31 Suspect clenches their hands and stares at me  
32 Suspect falls to the ground and curls their arms under their body  
33 Suspect strikes me with their elbow  
34 Suspect raises their voice and argues with me  
35 Suspect strikes my head with a baton  
36 Suspect jumps out of their car and runs away from me  
37 Suspect glares at me and clenches their teeth

Note: Numbering specifies only the ordering of behaviors presented. They did not appear with the behaviors during the survey.
Appendix D: Response to Provocation Short-Form

Using the 6 point scale shown below, indicate how you think you would respond to each of the following suspect behaviors during an arrest. Place your rating in the response space to the right of the behavior.

1 = No Force
2 = Low Force
3 = Moderate Force
4 = Intermediate Force
5 = High Force
6 = Maximum Force

1. Suspect scratches my face
   Response ______

2. Suspect yells at me
   Response ______

3. Suspect pulls away from me
   Response ______

4. Suspect raises their arms and makes a fist toward me
   Response ______

5. Suspect fires a shotgun at me
   Response ______

6. Suspect wraps their arms around the car steering wheel and holds on
   Response ______
ratings are conceptually similar to the use of thresholds in psychophysics (Collyer et al., 2004).

Consider how psychological screeners could index recruits' labeling thresholds using rating functions in Figure 1. Locate the function for the high labeling rater. This recruit sees behaviors that anchor the far left of the scale (low aggression) as mostly intermediate aggression. The recruit's detection rate gradually increases such that the recruit begins to detect differences in severity near behavior 28 (suspect clenches their hands and stares). Screeners can locate this recruit's threshold for detecting aggression by finding the recruit's average behavior rating (5.35) on the function and reading down to locate the threshold behavior on the scale. For the low labeling rater, this recruit sees low aggression behaviors as nonaggression. The recruit's threshold (2.49) is located near behavior 25 (suspect grabs around a pole).

Psychological screeners could also index recruits' provocation thresholds using rating functions in Figure 3. Locate the function for the high provocation rater. This recruit responds with intermediate levels of aggression against low-level aggressive behaviors. The recruit's threshold (5.65) is near behavior 30 (suspect folds and locks their arms) at which the recruit begins to respond using maximum force. For the low provocation rater, this recruit's threshold (2.43) is located near behavior 15 (suspect stands in a fighting stance).

Threshold values are informative. They allow psychological screeners to study the fitness of police candidates to engage in the performance of force. What screeners seek is a balance of perceptual sensitivity and response correctness. Recruits' mean average rating of behavior severity ($M = 3.67$) or the One-for-One model's mean
average response \((M = 3.81)\) gives screeners reasonable test values against which they can compare police candidates' labeling and provocation ratings. The one-sample \(t\)-test is a simple procedure screeners can use for evaluating the average difference between each response for a given task and the selected test value. Screeners can also specify confidence intervals for this difference. For example, Figure 3 shows a high provocation rater's responses plotted against the severity rank position corresponding to each suspect behavior. A one-sample \(t\)-test reveals that the individual's average responses \((M = 5.65, SD = 0.72)\) are significantly different from the test function \((M = 3.67), t(36) = 16.82\) at \(p < .01\) 2-tailed \((95\% \text{ confidence interval} = [1.74, 2.22])\).

What police trainers hope to accomplish in academy training sessions is raising aggression detection among police candidates (high labeling raters) and offering reasonable guidelines for choosing appropriate responses to threatening situations. Response correctness might mean raising the average provocation ratings of those candidates who are difficult to provoke into being aggressive and lowering it for those who are more easily provoked. What should be the condition of candidates who have received training is that they are perceptive to threat, but not hyper-receptive, and they are judicious in their forceful responses to threat, but not guarded.

**Responses to Aggression**

Although recruits agree on a common scale of behavior severity, the labeling and provocation tasks provide evidence that not every recruit perceives or responds to a given suspect behavior in the same way (see Figure 1 and 3). How recruits perceive aggression, however, is a strong predictor of how they will respond when provoked, \(r^2 = 0.76.\) Prediction was not as strong in Collyer et al. (2004) sample of college
students, $r^2 = 0.13$. Perhaps this difference is a selection phenomenon that is especially true in the police employment setting. Candidates have met strict entry standards and might be like-minded and better equipped to both detect and respond to acts of aggression.

Figure 4 shows a linear function, higher average labeling ratings generally imply higher average provocation ratings. Consider a high labeling rater, $M = 4$. On average, this recruit sees many behaviors as aggressive and discounts few of them as nonaggressive. If we read up to locate the recruit’s average response to different provocative behaviors, we find that the recruit generally uses responses that are more forceful. The recruit is physically aggressive to get the job done. Provocation ratings provide a measurement of physical aggression similar to Buss and Perry’s Aggression Questionnaire (1992; see Table 3).

Provocation ratings assess the readiness of recruits to be aggressive against a threat of force, but also reveal a degree of anger — “physiological arousal in preparation for aggression” (see Table 3). Anger is facilitative; it primes recruits’ defensive mechanisms when they interpret the behaviors of others as threatening or dangerous. Anger prepares recruits to fight. It might also trigger a hot reactive aggression (e.g., Beck, 1999). With this kind of angry aggression, recruits might hold suspects who resist their authority – showing disrespect – more culpable. Recruits’ forceful responses might take the form of punishment (or unreasonable force); suspects “must be taught a lesson for being disrespectful.” Hot reactive aggression might echo recruits’ cultural experiences (e.g., parenting styles, peer group pressure). Although anger is a normal reaction to threatening conditions, police trainers can
educate recruits on anger’s functional disadvantages and provide them with techniques to avoid its harmful effects.

*LL, LH, HH, and HL Typologies*

This paper proposes a framework for distinguishing police candidates according to the way they might deal with citizens. Psychological screeners can differentiate candidates on their labeling and provocation ratings. On matters of aggression, determining what is acceptable or unacceptable is inherent in the police selection process. The HH, HL, LL, and LH aggression types offer a useful framework for detecting emerging patterns of aggressive tendencies as a basis for (a) generating hypotheses about future work performance and (b) giving police trainers formative feedback on relevant aggression issues and problem solving skills that candidates should work on in academy training sessions.

Each aggression type suggests a set of perceptions and set of response behaviors of its occupants. These two sets are measurable independently and this investigation finds some provisional evidence suggestive of analytical descriptions of the typologies.

Consider how a change in recruits’ labeling ratings implies a change in their provocation ratings. The occupants of the HH aggression type are high labeling raters and high provocation raters. They see many types of suspect force as aggression, and are easily provoked into being aggressive. Their life experience tells a story of events involving the use and possible use of aggressive behaviors (see Appendix O).

Conceptually, the role of force in policing encourages measuring candidates’ conceptions of the continuum of force itself and their potential responses to citizens’
Appendix E: Interview Protocol

A) Purpose
Give the interviewee the purpose of the questionnaire.
1. To understand how police use force.
2. To find out how potential respondents understand questions and perform the response tasks. (Tell the interviewee you do not intend to use his or her answers to the questionnaire)

B) Administer a Short-Form of the Response to Provocation Inventory
Have the interviewee complete the inventory in the usual way.

C) Questions on Instructions
Have the interviewee read the inventory instructions aloud. Ask the following questions:
1. Would you summarize in your own words what you think the instructions meant?
2. Did you have any uncertainties or confusions about what the instructions meant?
3. Would you summarize in your own words what you think the term force meant?
4. Did you have any uncertainties or confusions about what force meant?
5. Would you summarize in your own words what you think each category of the response scale meant? (Please give examples of behaviors that describe each category).
6. Did you have any uncertainties or confusions about what the categories of the response scale meant?

D) Question-by-Question Review
Have the interviewee read each behavior item aloud. Ask the following questions:
1. Would you summarize in your own words what you think the item meant?
2. Did you have any uncertainties or confusions about the item?
3. Tell me how you went about calculating the answer you gave for the amount of force you would use?
4. How confident were you that you could give an accurate answer?
Appendix F: Buss-Perry Short-Form Aggression Questionnaire

Instructions:
Using the 6 point scale shown below, indicate how uncharacteristic or characteristic each of the following statements is in describing you. Place your rating in the space to the right of the statement.

| 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| extremely uncharacteristic of me | | | | | | extremely characteristic of me |

1. I often find myself disagreeing with people.  Rate______
2. At times I feel I have gotten a raw deal out of life.  Rate______
3. I have threatened people I know.  Rate______
4. I wonder why sometimes I feel so bitter about things.  Rate______
5. I have trouble controlling my temper.  Rate______
6. My friend’s say that I’m somewhat argumentative.  Rate______
7. I flare up quickly but get over it quickly.  Rate______
8. Given enough provocation, I may hit another person.  Rate______
9. I can’t help getting into arguments when people disagree with me.  Rate______
10. Other people always seem to get the breaks.  Rate______
11. There are people who pushed me so far that we came to blows.  Rate______
12. Sometimes I fly off the handle for no good reason.  Rate______
Appendix G: Buss-Perry Physical Aggression Subscale

Instructions:
Using the 5 point scale shown below, indicate how uncharacteristic or characteristic each of the following statements is in describing you. Place your rating in the space to the right of the statement.

1. Once in a while I can’t control the urge to strike another person. Rate ______
2. Given enough provocation, I may hit another person. Rate ______
3. If somebody hits me, I hit back. Rate ______
4. I get into fights a little more than the average person. Rate ______
5. If I have to resort to violence to protect my rights, I will. Rate ______
6. There are people who pushed me so far that we came to blows. Rate ______
7. I can think of no good reason for ever hitting a person. Rate ______
8. I have threatened people I know. Rate ______
9. I have become so mad that I have broken things. Rate ______
Appendix H: Aggressiveness Subscale from the Verbal Aggressiveness Scale

Instructions:
Please respond to the following statements by choosing the number from the list below which most accurately represents you.

1-Almost never true
2-Rarely true
3-Occasionally true
4-Often true
5-Almost always true

_____ 1. When individuals are very stubborn, I use insults to soften their stubbornness.
_____ 2. When people refuse to do a task I know is important, without good reason, I tell them they are unreasonable.
_____ 3. If individuals I am trying to influence really deserve it, I attack their character.
_____ 4. When people behave in ways that are in very poor taste, I insult them in order to shock them into proper behavior.
_____ 5. When people simply will not budge on a matter of importance, I lose my temper and say rather strong things to them.
_____ 6. When individuals insult me, I get a lot of pleasure out of really telling them off.
_____ 7. I like poking fun at people who do things which are very stupid in order to stimulate their intelligence.
_____ 8. When people do things which are mean or cruel, I attack their character in order to help correct their behavior.
_____ 9. When nothing seems to work in trying to influence others, I yell and scream in order to get some movement from them.
_____ 10. When I am not able to refute others’ positions, I try to make them feel defensive in order to weaken their positions.
Appendix I: Anger Arousal Subscale from the Multidimensional Anger Inventory

Instructions:
Everybody gets angry from time to time. A number of statements that people have used to describe the times that they get angry are included below. Read each statement and circle the number to the right of the statement that best describes you. There are no right or wrong answers.

If the statement is completely undescriptive of you, circle a 1.
If the statement is mostly undescriptive of you, circle a 2.
If the statement is partly undescriptive and partly descriptive of you, circle a 3.
If the statement is mostly descriptive of you, circle a 4.
If the statement is completely descriptive of you, circle a 5.

Please answer every item.

1. I tend to get angry more frequently than most people. 1 2 3 4 5
2. It is easy to make me angry. 1 2 3 4 5
3. Something makes me angry almost every day. 1 2 3 4 5
4. I often feel angrier than I think I should. 1 2 3 4 5
5. I am surprised at how often I feel angry. 1 2 3 4 5
6. At times, I feel angry for no specific reason. 1 2 3 4 5
7. When I get angry, I stay angry for hours. 1 2 3 4 5
8. I get so angry, I feel like I might lose control. 1 2 3 4 5
Appendix J: Cook-Medley Hostility Scale

Directions:
If a statement is true or mostly true, as pertaining to you, circle the letter T. If a statement is false, or usually not true about you, circle the letter F. Try to give a response to every statement.

1. When I take a new job, I like to be tipped off on who should be gotten next to. T F
2. When someone does me wrong, I feel I should pay him back if I can, just for the principle of the thing. T F
3. I prefer to pass by school friends, or people I know but have not seen for a long time, unless they speak to me first. T F
4. I often had to take orders from someone who did not know as much as I did. T F
5. I think a great many people exaggerate their misfortunes in order to gain the sympathy and help of others. T F
6. It takes a lot of argument to convince most people of the truth. T F
7. I think most people lie to get ahead. T F
8. Someone has it in for me. T F
9. Most people are honest chiefly through the fear of getting caught. T F
10. Most people will use somewhat unfair means to gain profit or an advantage, rather than lose it. T F
11. I commonly wonder what hidden reason another person may have for doing something nice for me. T F
12. It makes me impatient to have people ask my advice or otherwise interrupt me when I am working on something important. T F
13. I feel that I have often been punished without cause. T F
14. I am against giving money to beggars. T F
15. Some of my family has habits that bother me very much. T F
16. My relatives are nearly all in sympathy with me. T F
17. My way of doing things is apt to be misunderstood by others. T F
18. I don’t blame anyone for trying to grab everything they can get in this world. T F
19. No one cares what happens to you. T F
20. I can be friendly with people who do things I consider wrong. T F
21. It is safer to trust nobody. T F
22. I do not blame a person for taking advantage of someone who lays himself open to it. T F
23. I have often felt that strangers were looking at me critically. T F
24. Most people make friends because friends are likely to be useful to them. T F
25. I am sure that I am being talked about. T F
26. I am likely not to speak to people until they speak to me. T F
27. Most people inwardly dislike putting themselves out to help other people. T F
28. I tend to be on guard with people who are somewhat more friendly than I had expected. T F
29. I have sometimes stayed away from another person because I feared doing or saying something that I might regret afterwards. T F
30. People often disappoint me.  
31. I like to keep people guessing what I'm going to do next.  
32. I frequently ask people for advice.  
33. I am not easily angered.  
34. I have often met people who are supposed to be experts who were no better than I.  
35. It makes me think of failure when I hear of the success of someone I know well.  
36. I would certainly enjoy beating a crook at his own game.  
37. I have at times had to be rough with people who were rude or annoying.  
38. People generally demand more respect for their own rights than they are willing to allow for others.  
39. There are certain people whom I dislike so much I am inwardly pleased when they are catching it for something they have done.  
40. I am often inclined to go out of my way to win a point with someone who has opposed me.  
41. I am quite often not in on the gossip and talk of the group I belong to.  
42. The man who had the most to do with me when I was a child (such as my father, step-father, etc.) was very strict with me.  
43. I have often found people jealous of my good ideas just because they had not thought of them first.  
44. When a man is with a woman, he is usually thinking of things related to her sex.  
45. I do not try to cover up my poor opinion or pity of a person so that he won't know how I feel.  
46. I have frequently worked under people who seem to have things arranged so that they get credit for good work, but are able to pass off mistakes to those under them.  
47. I strongly defend my own opinions as a rule.  
48. People can pretty easily change me even though I thought that my mind was made up on a subject.  
49. Sometimes I am sure that other people can tell what I'm thinking.  
50. A large number of people are guilty of bad sexual conduct.
Appendix K: Marlowe-Crowne Short-Form Social Desirability Scale

Listed below are a number of statements concerning personal attitudes and traits. Read each item and decide whether the statement is true or false as it pertains to you personally. Place your response in the space before each statement.

1. It is sometimes hard for me to go on with my work if I am not encouraged.
2. I sometimes feel resentful when I don’t get my way.
3. On a few occasions, I have given up doing something because I thought too little of my ability.
4. There have been times when I felt like rebelling against people in authority even though I knew they were right.
5. No matter who I’m talking to, I’m always a good listener.
6. There have been occasions when I took advantage of someone.
7. I’m always willing to admit it when I make a mistake.
8. I sometimes try to get even rather than forgive and forget.
9. I am always courteous, even to people who are disagreeable.
10. I have never been irked when people expressed ideas very different from my own.
11. There have been times when I was quite jealous of the good fortune of others.
12. I am sometimes irritated by people who ask favors of me.
13. I have never deliberately said something that hurt someone’s feelings.
Appendix L: Behavioral Experience Questionnaire

For each Activity or Event listed below, enter under “Times,” the total number of instances that occurred to you or was something that you did. If you are unsure of the exact number of times, enter a “?” next to your best estimate.

| Times | Item                                                                 | Activity or Event                                                                                       |
|-------|----------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------|
| 1     | As an employee, caused more than $50 in loss or damage to property or equipment belonging to an employer |                                                                                                         |
| 2     | Set a fire or committed an act of vandalism that resulted in any injury that required professional medical attention and/or property damage of $50 or more |                                                                                                         |
| 3     | Charged with or convicted of any misdemeanor even if the charges were dropped or expunged |                                                                                                         |
| 4     | Charged with or convicted of driving under the influence of drugs or alcohol even if the charges were dropped or expunged |                                                                                                         |
| 5     | Charged with or convicted of a firearms or explosives offense even if the charges were dropped or expunged |                                                                                                         |
| 6     | Charged with or convicted of the possession, sale, or manufacture of illegal drugs even if the charges were dropped or expunged |                                                                                                         |
| 7     | Engaged or participated in an act of domestic violence |                                                                                                         |
| 8     | Have been accused or convicted of domestic violence even if the charges were dropped or expunged |                                                                                                         |
| 9     | Use of alcohol has caused you some kind of trouble |                                                                                                         |
| 10    | Use of drugs has caused you some kind of trouble |                                                                                                         |
| 11    | Had a lawsuit (legal or civil) either filed or rendered against you |                                                                                                         |
| 12    | Had a license suspended or revoked (e.g. a driver’s license or license to practice a trade or profession) |                                                                                                         |
| 13    | Have been denied custody of one or more of your children |                                                                                                         |
| 14    | Had a physical altercation (fight which could include just pushing and shoving) with coworker(s), supervisor(s), or subordinate(s) |                                                                                                         |
| 15    | Had a physical altercation (fight which could include just pushing and shoving) with someone, other than a person at your place of employment, as an adult (18 years of age or older) |                                                                                                         |
| Times | Item | Activity or Event |
|-------|------|------------------|
| 16    | 16   | Had a physical altercation (fight which could include just pushing and shoving) with someone as a juvenile (17 years of age or younger) |
| 17    | 17   | Had to warn or threaten someone because you felt in danger or provoked |
| 18    | 18   | Had to defend yourself or someone else because you felt in danger or provoked |
| 19    | 19   | Have had to point a firearm at someone |
| 20    | 20   | Someone has pointed a firearm at you |
| 21    | 21   | Have carried a firearm outside the line of duty |
| 22    | 22   | Have done something about an animal that was annoying you |
| 23    | 23   | Have had to react to an animal that was threatening you or someone else |
| 24    | 24   | Have been involved in an accident while driving a motorized vehicle |
| 25    | 25   | While driving a motorized vehicle, have yelled, cursed or used nonverbal gestures at another driver or passenger that was annoying you |
| 26    | 26   | While a passenger in a motorized vehicle, have yelled, cursed or used nonverbal gestures at another driver or passenger that was annoying you |
| 27    | 27   | Have been the subject of an internal investigation while serving in a security, law enforcement or other public safety position |
| 28    | 28   | Have had a sustained complaint regarding your performance or conduct while serving in a security, law enforcement or other public safety position |
| 29    | 29   | Have illegally used controlled substances while employed as security officer, law enforcement officer, prosecutor, courtroom official, or other public safety official |
| 30    | 30   | Left a job following allegations of misconduct |
| 31    | 31   | Quit a job after a dispute |
| 32    | 32   | Received a verbal reprimand or verbal warning regarding your on-the-job misconduct |
| 33    | 33   | Received a letter of reprimand or letter of warning regarding your on-the-job misconduct |
| Times | Item | Activity or Event |
|-------|------|-------------------|
| 34    |      | Received a verbal or written reprimand for directing insulting, abusive, or obscene language at coworker(s), supervisor(s) or subordinate(s) |
| 35    |      | Received a verbal or written reprimand for reporting to work under the influence of alcohol and/or any illegal, controlled substance |
| 36    |      | Have been subject to court-martial or disciplinary proceedings under the Uniform Code of Military Justice (include non-judicial, Captain’s Mast, etc.) |
| 37    |      | Received less than an honorable discharge from the military (provide the type of discharge in the following Description of Circumstances and Outcome section) |

In the space below, enter the year (or approximate year), the item number and a brief description of the circumstances and outcome surrounding each activity or event that you listed above as happening to you or was something that you did.

| Year | Item | Description of Circumstances and Outcome |
|------|------|------------------------------------------|
Appendix M: Average Ratings and Rankings of Suspect Behaviors by Labeling and Provocation Tasks

| Item | Suspect Behavior                              | Mean Severity Rating | Mean Provocation Rating | Mean One-for-One Rating |
|------|-----------------------------------------------|----------------------|-------------------------|-------------------------|
| 1    | fires a handgun                               | 5.97                 | 5.96                    | 6                       |
| 11   | fires a shotgun                               | 5.97                 | 5.99                    | 6                       |
| 16   | fires a rifle                                 | 5.95                 | 5.96                    | 6                       |
| 25   | stabs with knife                              | 5.92                 | 5.92                    | 6                       |
| 22   | drives a car at                               | 5.82                 | 5.79                    | 6                       |
| 29   | hits neck with a baseball bat                 | 5.73                 | 5.68                    | 6                       |
| 35   | strikes head with a baton                     | 5.68                 | 5.62                    | 6                       |
| 4    | grabs and throws to ground                    | 5.10                 | 4.90                    | 5                       |
| 13   | punches                                       | 4.97                 | 4.74                    | 5                       |
| 21   | pushes to ground                              | 4.80                 | 4.71                    | 5                       |
| 33   | strikes with elbow                            | 4.73                 | 4.49                    | 5                       |
| 18   | kicks                                         | 4.59                 | 4.40                    | 5                       |
| 7    | scratches face                                | 4.31                 | 4.11                    | 5                       |
| 28   | slaps face                                    | 4.20                 | 4.12                    | 5                       |
| 5    | stands in a fighting stance                   | 4.07                 | 3.63                    | 2                       |
| 10   | raises arms and makes a fist                   | 3.80                 | 3.45                    | 5                       |
| 15   | flees in a car                                | 3.76                 | 3.64                    | 4                       |
| 2    | runs out of house and away                    | 3.37                 | 2.74                    | 4                       |
| 36   | jumps out of car and runs away                | 3.33                 | 3.23                    | 4                       |
| 12   | wraps arms around steering wheel              | 3.25                 | 2.83                    | 3                       |
| 9    | pulls away                                    | 3.24                 | 3.16                    | 4                       |
| 24   | flees on foot                                 | 3.21                 | 3.22                    | 4                       |
| 19   | clenches fists                                | 3.15                 | 2.82                    | 2                       |
| 26   | flees on bycyle                               | 3.09                 | 3.04                    | 4                       |
| 17   | grabs around a pole                           | 3.08                 | 2.70                    | 3                       |
| 30   | lays on ground and stiffens body              | 2.52                 | 2.49                    | 3                       |
| 32   | falls to ground curl arms under body          | 2.52                 | 2.50                    | 4                       |
| 31   | clenches hands and stares                     | 2.46                 | 2.15                    | 2                       |
| 20   | sits in chair and tucks arms tightly          | 2.17                 | 2.01                    | 3                       |
| 6    | folds and locks arms                         | 2.11                 | 2.11                    | 3                       |
| 23   | aggressively points finger                    | 2.11                 | 1.85                    | 2                       |
| 34   | raises voice and argues                       | 1.91                 | 1.81                    | 3                       |
| 37   | glares and clenches teeth                     | 1.89                 | 1.82                    | 3                       |
| 14   | shouts and curses                            | 1.87                 | 1.72                    | 4                       |
| 8    | yells                                         | 1.79                 | 1.60                    | 3                       |
| 3    | screams                                       | 1.76                 | 1.57                    | 4                       |
| 27   | swears                                        | 1.63                 | 1.44                    | 3                       |

Mean Average Rating: 3.67 (Sev. Rank) 3.51 (Prov. Rank) 3.81 (One-for-One Rank)

Note. See Appendix A and Appendix C for the complete wording of suspect behavior items for the labeling and provocation tasks.
Appendix N: Labeling Task Behavior Items and Varimax Pattern

| Item | Suspect Behavior | Low   | Moderate | Intermediate | High | Maximum |
|------|------------------|-------|----------|--------------|------|---------|
| 21   | push to ground   | 0.07  | 0.08     | 0.24         | 0.79 | 0.15    |
| 18   | kicks            | 0.08  | 0.15     | 0.19         | 0.77 | 0.11    |
| 13   | punches          | 0.13  | 0.06     | 0.24         | 0.74 | 0.31    |
| 7    | scratches face   | 0.13  | 0.20     | 0.11         | 0.74 | 0.16    |
| 33   | strikes with elbow | 0.15 | 0.18    | 0.14         | 0.72 | 0.22    |
| 28   | slaps face       | 0.25  | 0.17     | 0.13         | 0.64 | 0.20    |
| 4    | grabs and throws to ground | 0.11 | 0.08    | 0.25         | 0.61 | 0.21    |
| 9    | pulls away       | 0.34  | 0.22     | 0.21         | 0.59 | -0.01   |
| 10   | raises arms and makes a fist | 0.32 | 0.18    | 0.17         | 0.56 | -0.01   |
| 19   | clenches fists   | 0.34  | 0.32     | 0.14         | 0.53 | 0.08    |
| 5    | stands in a fighting stance | 0.25 | 0.09    | 0.19         | 0.50 | 0.05    |
| 29   | hunches neck with a baseball bat | 0.07 | 0.10    | -0.03        | 0.43 | 0.38    |
| 8    | yells            | 0.87  | 0.12     | 0.18         | 0.17 | 0.03    |
| 14   | shouts and curses | 0.84 | 0.16    | 0.24         | 0.18 | 0.07    |
| 27   | swears           | 0.84  | 0.20     | 0.17         | 0.20 | -0.06   |
| 3    | screams          | 0.81  | 0.15     | 0.27         | 0.17 | 0.03    |
| 34   | raises voice and argues | 0.77 | 0.23    | 0.13         | 0.16 | 0.06    |
| 37   | glares and clenches teeth | 0.61 | 0.40    | 0.11         | 0.20 | 0.05    |
| 23   | aggressively points finger | 0.58 | 0.30    | 0.21         | 0.13 | 0.08    |
| 31   | clenches hands and stares | 0.47 | 0.45    | 0.07         | 0.25 | 0.15    |
| 6    | folds and locks arms | 0.44 | 0.43    | 0.14         | 0.17 | 0.13    |
| 32   | falls to ground curls arms under body | 0.30 | 0.74    | 0.13         | 0.11 | 0.04    |
| 50   | lays on ground and stiffens body | 0.32 | 0.73    | 0.18         | 0.09 | 0.11    |
| 17   | grabs around a pole | 0.21 | 0.65    | 0.27         | 0.20 | -0.03   |
| 12   | wraps arms around car steering wheel | 0.09 | 0.60    | 0.27         | 0.28 | 0.01    |
| 20   | sits in chair and tucks arms tightly | 0.43 | 0.57    | 0.09         | 0.24 | 0.07    |
| 26   | flees on bycycle | 0.30  | 0.28     | 0.79         | 0.26 | 0.09    |
| 24   | flees on foot    | 0.29  | 0.31     | 0.77         | 0.24 | 0.04    |
| 36   | jumps out of car and runs away | 0.35 | 0.25    | 0.68         | 0.28 | 0.05    |
| 2    | runs out of house and away | 0.22 | 0.21    | 0.65         | 0.31 | 0.02    |
| 15   | flees in a car   | 0.26  | 0.08     | 0.62         | 0.32 | 0.14    |
| 1    | fires a handgun  | 0.02  | -0.01    | 0.14         | 0.00 | 0.81    |
| 11   | fires a shotgun  | -0.02 | 0.06     | 0.10         | -0.01| 0.79    |
| 16   | fires a rifle    | 0.04  | -0.05    | 0.06         | 0.21 | 0.46    |
| 25   | stabs with knife | 0.07  | 0.08     | -0.08        | 0.28 | 0.43    |
| 35   | strikes head with a baton | 0.07 | 0.07    | -0.06        | 0.28 | 0.43    |
| 22   | drives a car at  | 0.10  | 0.09     | 0.06         | 0.34 | 0.40    |

Note. The bold type loadings indicate the highest loading and assignment to suggested factors. See Appendix A for the complete wording of the suspect behavior items.
Appendix O: Frequency Count of Life Events from the
Behavioral Experience Questionnaire

| Life Event | HH (n = 36) Count | HL (n = 6) Count | LL (n = 4) Count | LH (n = 2) Count |
|------------|-------------------|------------------|------------------|-----------------|
| 1          | 45                | 1                | 3                | 2               |
| 2          | 0                 | 1                | 0                | 0               |
| 3          | 8                 | 0                | 0                | 0               |
| 4          | 1                 | 0                | 0                | 0               |
| 5          | 0                 | 0                | 0                | 0               |
| 6          | 0                 | 0                | 0                | 0               |
| 7          | 2                 | 0                | 0                | 0               |
| 8          | 0                 | 0                | 0                | 0               |
| 9          | 10                | 2                | 0                | 0               |
| 10         | 0                 | 0                | 0                | 0               |
| 11         | 0                 | 0                | 0                | 0               |
| 12         | 4                 | 1                | 2                | 1               |
| 13         | 0                 | 0                | 0                | 0               |
| 14         | 2                 | 0                | 0                | 0               |
| 15         | 52                | 2                | 5                | 0               |
| 16         | 55                | 8                | 3                | 0               |
| 17         | 102               | 4                | 3                | 0               |
| 18         | 65                | 2                | 2                | 0               |
| 19         | 66                | 0                | 0                | 0               |
| 20         | 21                | 0                | 0                | 0               |
| 21         | 50                | 0                | 0                | 0               |
| 22         | 7                 | 0                | 0                | 0               |
| 23         | 13                | 0                | 1                | 0               |
| 24         | 50                | 8                | 4                | 3               |
| 25         | 90                | 17               | 6                | 0               |
| 26         | 63                | 13               | 1                | 0               |
| 27         | 0                 | 0                | 0                | 0               |
| 28         | 0                 | 0                | 0                | 0               |
| 29         | 0                 | 0                | 0                | 0               |
| 30         | 0                 | 0                | 0                | 0               |
| 31         | 0                 | 0                | 1                | 0               |
| 32         | 5                 | 0                | 0                | 0               |
| 33         | 3                 | 0                | 1                | 0               |
| 34         | 0                 | 0                | 0                | 0               |
| 35         | 0                 | 0                | 0                | 0               |
| 36         | 0                 | 0                | 2                | 0               |
| 37         | 0                 | 0                | 0                | 0               |

Note. HH = high labeling high provocation raters; HL = high labeling low provocation raters; LL = low labeling low provocation raters; LH = low labeling high provocation raters. See Appendix L for a complete description of life event items.
Bibliography

Adams, K. (1995). Measuring the prevalence of police abuse of force. In W. A. Geller & H. Toch (Eds.), *And justice for all: Understanding and controlling police abuse of force* (pp. 61-98). Washington, DC: Police Executive Research Forum.

Adams, K. (1996). What we know about police use of force. In, *Use of force by police: Overview of national and local data* (pp. 1-14). Washington, DC: National Institute of Justice.

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Bartol, C. R. (1982). Psychological characteristics of small town police officers. *Journal of Police Science and Administration, 10*, 58-63.

Bartol, C. R. (1991). Predictive validation of the MMPI for small-town police officers who fail. *Professional Psychology Research and Practice, 22*(2), 127-132.

Beatty, M. J., Rudd, J. E., & Valencic, K. M. (1999). A re-examination of the verbal aggressiveness scale: One factor or two? *Communication Research Reports, 16*, 10-17.

Beck, A. T. (1999). *Prisoners of hate: The cognitive basis of anger, hostility, and violence*. New York, NY: Harper Collins.

Benner, A. W. (1986). Psychological screening of police applicants. In J. T. Reese & H. A. Goldstein (Eds.), *Psychological services for law enforcement* (pp. 11-18). Washington, DC: U.S. Government printing Office.
Beutler, L. E., Nussbaum, P. D., & Meredith, K. E. (1998). Changing personality patterns of police officers. *Professional Psychology Research and Practice, 19*(5), 503-507.

Biaggio, M. K., & Maiuro, R. D. (1985). Recent advances in anger assessment. In C. D. Spielberger, & J. N. Butcher (Eds.), *Advances in personality assessment* (pp. 71-101). Hillsdale, NJ: Lawrence Erlbaum.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: John Wiley.

Bonsignore v. City of New York. (1981). In *New York Law Journal, 78-0240.*

Bushman, B. J., Cooper, H. M., & Lemke, K. M. (1991). Meta-analysis of factor analyses: An illustration using the Buss-Durkee Hostility Inventory. *Personality and Social Psychology Bulletin, 17,* 344-349.

Buss, A. H. (1961). *The psychology of aggression.* New York: Wiley.

Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology, 63,* 452-459.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota Multiphasic Personality Inventory-2: Manual for administration and scoring.* Minneapolis: University of Minnesota Press.

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,* 138-167.

Cohen, B., & Chaiken, J. M. (1973). *Police background characteristics and performance.* Lexington, MA: Lexington Books.
Collyer, C. E., Gallo, F. J., & Boney-Mcloy, S. (2004). Individual perceptions of violence: Labeling and provocation thresholds. Unpublished manuscript, University of Rhode Island.

Conte v. Horcher, 50 Ill. App. 3rd 151 (1977).

Contrada, R. J. & Jussim, L. (1992). What does the Cook-Medley Hostility Scale measure? In search of an adequate measure model. Journal of Applied Social Psychology, 22, 615-627.

Converse, J. M., & Presser, S. (1986). Survey questions: Handcrafting the standardized questionnaire. Thousand Oaks, CA: Sage Publications.

Costello, R. M., Schneider, S. L., & Schoenfeld, L. S. (1996). Validation of a preemployment MMPI index correlated with disciplinary suspension days of police officers. Psychology, Crime and Law, 2, 299-306.

Craig, R. J. (2005). Personality-guided forensic psychology. Washington, DC: American Psychological Association.

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. Journal of Consulting Psychology, 24, 349-354.

Curran, S. F. (1998). Revised IACP guidelines: Pre-employment psychological evaluation of law enforcement applicants. The Police Chief, 65(10), 88-95.

Daley, R. (1982). Psychological screening of police candidates. Police Chief, 49, 53-54.

Detrick, P., Chibnall, J. T., & Rosso, M. (2001). Minnesota Multiphasic Inventory-2 in police officer selection: Normative data and relation to the Inwald Personality Inventory. Professional Psychology: Research and Practice, 32(5), 484-490.
Evan, D. R., & Stangeland, M. (1971). Development of the reaction inventory to measure anger. *Psychological Reports, 21*, 412-414.

Fitzgerald, P. R. (1987). The prediction of police performance using the MMPI and CPI. (Doctoral dissertation, St. Louis University, 1987). *Dissertation Abstracts International, 47*, 3519.

Forsyth, B. H., & Lessler, J. T. (1991). Cognitive laboratory methods: A taxonomy. In P. N. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), *Measurement errors in surveys* (pp. 393-418). New York: John Wiley.

Fowler, F. J. (1995). *Improving survey questions*. Thousand Oaks, CA: Sage Publications.

Garner, J., Buchanan, J., Schade, T., & Hepburn, J. (1996). *Understanding the use of force by and against police*. Washington, DC: National Institute of Justice.

Garner, J. H., & Maxwell, C. D. (1996). Measuring the amount of force used by and against the police in six jurisdictions. In, *Use of force by police: Overview of national and local data* (pp. 25-44). Washington, DC: National Institute of Justice.

Gough, H. G. (1975). *Manual for the California Psychological Inventory*. Palo Alto, CA: Consulting Psychological Press.

Greenberg, B. E., Riggs, M., & Bryant, F. B. (2003). Validation of a short aggression inventory for law enforcement. *Journal of Police and Criminal Psychology, 18*(2), 12-19.
Grant, J. D., & Grant, J. (1995). Officer selection and the prevention of abuse of force. In W. A. Geller & H. Toch (Eds.), *And justice for all: Understanding and controlling police abuse of force* (pp. 151-162). Washington, DC: Police Executive Research Forum.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103*(2), 265-275.

Hargrave, G. E., Hiatt, D. (1989). Use of the California psychological inventory in law enforcement officer selection. *Journal of Personality Assessment, 53*(2), 267-277.

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 and Administration, 16*(4), 268-273.

Hathaway, S. R., & Mckinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory* (rev. ed.). Minneapolis: University of Minnesota Press.

Hiatt, D. P. (1986). The benefits of psychological screening to the applicants. In J. T. Reese & H. A. Goldstein (Eds.), *Psychological services for law enforcement* (pp. 43-45). Washington, DC: U.S. Government printing Office.

Hibler, N. S., & Kurke, M. I. (1995). Ensuring personal reliability through selection and training. In M. I. Kurke & E. M. Scrivner (Eds.), *Police psychology into the 21st century* (pp. 71-91). Hillsdale, NJ: Erlbaum.

Infante, D. A., & Wigley, C. J., III. (1986). Verbal aggressiveness: An interpersonal model and measure. *Communication Monographs, 53*, 63–69.
International Association of Chiefs of Police. (2004). *Psychological evaluation of public safety applicants: The 2004 revised guidelines of the police psychological services section*. Alexandria, VA: International Association of Chiefs of Police.

Independent Commission on the Los Angeles Police Department. (1991). *Report of the independent commission on the Los Angeles police department*. Los Angeles, CA: Independent Commission on the Los Angeles Police Department.

Inwald, R. E. (1986). Law enforcement officer screening: A description of one pre-employment psychological testing program. In J. T. Reese & H. A. Goldstein (Eds.), *Psychological services for law enforcement* (pp. 11-18). Washington, DC: U.S. Government printing Office.

Inwald, R. E., Knatz, H., & Shusman, L. (1983). *Inwald Personality Inventory manual*. New York: Hilson Research.

Inwald, R. E., & Shusman, E. J. (1984). The IPI and MMPI as predictors of academy performance for police recruits. *Journal of Police Science and Administration*, 12(1), 1-11.

Keppel, G. (1991). *Design and analysis: A researcher’s handbook* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Levine, T. R., Beatty, M. J., Limon, S., Hamilton, M. A., Buck, R., & Chory-Assad, R. M. (2004). The dimensionality of the Verbal Aggressiveness Scale. *Communication Monographs*, 71(3), 245-268.

McKenna v. Fargo, 451 F. Supp. 1355 (U.S. District Courts, NJ, May 25, 1978).
McLaughlin, V. (1992). Police and the use of force: The savannah study. Westport, CT: Praeger.

Megargee, E. I. (1970). The prediction of violence with psychological tests. In C. D. Spielberger (Ed.), *Current topics in clinical and community psychology volume 2* (pp. 97-156). New York: Academic Press.

Mills, R. B., McDevitt, R. J., & Tonkin, S. (1966). Situational tests in metropolitan police recruit selection. *The Journal of Criminal Law, Criminology and Police Science, 57*(1), 99-106.

Mills, M. C., & Stratton, J. G. (1982). The MMPI and the prediction of police job performance. *FBI Law Enforcement Bulletin, 51*(2), 10-15.

Mufson, D. W., & Mufson, M. A. (1998). Predicting police officer performance using the Inwald Personality Inventory: An illustration from Appalachia. *Professional Psychology: Research Practice, 29*(1), 59-62.

Neal, B. (1986). The K scale (MMPI) and job performance. In J. T. Reese & H. A. Goldstein (Eds.), *Psychological services for law enforcement* (pp. 11-18). Washington, DC: U.S. Government printing Office.

Poland, J. M. (1978). Police selection methods and performance. *Journal of Police Science and Administration, 6*(4), 374-393.

Reynolds, W. (1982). Development of reliable and valid short forms of the Marlowe-Crowne social desirability scale. *Journal of Clinical Psychology, 38*(1), 119-125.
Sarchione, C. D., Cutler, M. J., Muchinsky, P. M., & Nelson-Gray, R. O. (1998). Prediction of dysfunctional job behaviors among law enforcement officers. *Journal of Applied Psychology, 83*(6), 904-912.

Scogin, F., Schumacher, J., Gardner, J., & Chaplin, W. (1995). Predictive validity of psychological testing in law enforcement settings. *Professional Psychology Research and Practice, 26*(1), 68-71.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*, 420-428.

Shusman, E. J., Inwald, R. E., & Knatz, H. F. (1987). A cross-validation study of police recruit performance as predicted by the IPI and MMPI. *Journal of Police Science and Administration, 15*(2), 162-169.

Silverstein, A. B. (1983). Validity of random short forms: II. The Marlowe-Crowne social desirability scale. *Journal of Clinical Psychology, 39*(4), 582-584.

Skolnick, J. H., & Fyfe, J. J. (1994). *Above the law: Police and the excessive use of force*. New York, NY: Free Press.

Smith, D. A., Klein, J. R. (1984). Police control of interpersonal disputes. *Social Problems, 31*(4), 467-481.

Streiner, D.L. (1994). Figuring out factors: The use and misuse of factor analysis. *Canadian Journal of Psychiatry, 39*, 135-140.

Sudman, S., & Bradburn, N. M. (1982). Asking questions: A practical guide to questionnaire design. San Francisco, CA: Jossey-Bass Publishers.
Suris, A., Lind, L., Emmett, G., Borman, P. D., Kashner, M., & Barratt, E. S. (2004). Measures of aggressive behavior: Overview of clinical and research instruments. *Aggression and Violent Behavior, 9*, 165-227.

Surrette, M. A., & Serafino, G. (2003). Relationship between personality and law enforcement performance. *Applied HRM Research, 8*(2), 89-92.

Suzuki, S., & Rancer, A. S. (1994). Argumentativeness and verbal aggressiveness: Testing for conceptual and measurement equivalence across cultures. *Communication Monographs, 61*, 256–279.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Needham Heights, MA: Allyn & Bacon.

Tenopyr, M. L. (1994). Big five, structural modeling, and item response theory. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction* (pp. 519-534). Palo Alto, CA: Consulting Psychologists Press.

Terry v. Ohio, 392 U.S. 1 (1968).

Toch, H. (1995). The violence-prone police officer. In W. A. Geller & H. Toch (Eds.), *And justice for all: Understanding and controlling police abuse of force* (pp. 99-112). Washington, DC: Police Executive Research Forum.

Varela, J. G., Boccaccini, M. T., Scogin, F., Stump, J., & Caputo, A. (2004). Personality testing in law enforcement employment settings. *Criminal Justice and Behavior, 3*(6), 649-675.
Vetter, G. R. (1999, Winter). Comment: Is a personality test a pre-job-offer medical examination under the ADA? [Electronic version]. *Northwestern University Law Review*, 93, Nw. U. L. Rev. 597.

Wells, V. K. (1992). The MMPI and CPI as predictors of police performance (Doctoral dissertation, St. Louis University, 1992). *Dissertation Abstracts International*, 53, 597.

Wrightsman, L. S., & Fulero, S. M. (2005). *Forensic psychology* (2nd ed.). Belmont, CA: Wadsworth.

Zook, A. II, & Sipps, G. J. (1985). Cross-validation of a short-form of the Marlowe-Crowne social desirability scale. *Journal of Clinical Psychology, 41*(2), 236-238.