DEVELOPMENT OF A MEASURE OF AGGRESSIVE BEHAVIOR EXPECTANCIES IN ADULTS: THE AGGRESSION EXPECTANCY QUESTIONNAIRE

by

COURTLAND STEVENS HYATT

(Under the Direction of JOSHUA MILLER)

ABSTRACT

Social learning theory postulates that aggression is learned and elicited through a series of cognitive processes including expectancies, or the various consequences that an individual views as likely following aggressive behavior. The current manuscript describes a measurement development project that ultimately yielded a 16-item measure of positive and negative aggression expectancies suitable for use in adult populations. Across six studies, we took an iterative approach by gathering content from laypersons and experts in aggression research, administering large item pools to several samples, and refining item content through a combination of empirical (i.e., factor loadings, model fit) and conceptual (i.e., content breadth, non-redundancy) considerations. The Aggression Expectancy Questionnaire (AEQ) displays a robust four-factor structure, as well as preliminary evidence of convergent and divergent validity with self-reported aggression and relevant personality variables (i.e., antagonism). It is posited that this type of cognitive mechanism may serve as an intermediary link between distal predictors of aggression (e.g., psychopathy, narcissism) and its proximal manifestation.

INDEX WORDS: Aggression, Expectancies, Social Learning Theory, Personality.

DEVELOPMENT OF A MEASURE OF AGGRESSIVE BEHAVIOR EXPECTANCIES IN ADULTS: THE AGGRESSION EXPECTANCY QUESTIONNAIRE

by

COURTLAND STEVENS HYATT

B.S., The University of Georgia, 2014

M.S., The University of Georgia, 2017

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2021

DEVELOPMENT OF A MEASURE OF AGGRESSIVE BEHAVIOR EXPECTANCIES IN ADULTS: THE AGGRESSION EXPECTANCY QUESTIONNAIRE

by

COURTLAND STEVENS HYATT

Major Professor: Joshua Miller

Committee: Keith Campbell
Nathan Carter
Amos Zeichner

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2021

ACKNOWLEDGEMENTS

I would like to acknowledge and thank my loved ones who have supported me through all my personal and professional pursuits. These include: my wife and best friend, Callie Hyatt; my parents, Daphne and Todd Hyatt; my brother, Alexander Hyatt; my grandparents, Linda and Wendell Burton, Charles Hill, Allie Kilpatrick-Hill, and Margaret and Morley Hyatt; my schnauzer, Chance; and to all my cherished friends. I am also deeply indebted to my major professor, Josh Miller, for his invaluable guidance in academia and beyond. Finally, I am grateful to my doctoral committee for sharing their expertise and wisdom. My sincerest gratitude to you all.

TABLE OF CONTENTS

LIST OF TABLES

Page

CHAPTER 1

INTRODUCTION

Aggression is defined as intentional harm inflicted upon an individual who is motivated to avoid this harm (Anderson & Bushman, 2002), and it can take many forms (e.g., physical, verbal, relational) and be enacted with many motivations (e.g., proactive, reactive). Beyond the intrapersonal impact on victims, aggression and related conduct problems are associated with substantial societal costs that burden economic, healthcare, educational, and criminal justice systems (Burt et al., 2018; Rivenbark et al., 2018). Given long-standing interest in understanding this important behavior, there are decades of empirical literature aimed at identifying the situation- and person-level variables that contribute to aggressive behavior (e.g., Bettencourt, Talley, Benjamin, & Valentine, 2006). The focus of the current article is on person-level variables, but it is important to mention that numerous situational precursors to aggression have been identified, including but not limited to provocation (e.g., Bettencourt & Miller, 1996), threat (e.g., Taylor, Gammon, & Capasso, 1976), and acute alcohol consumption (e.g., Parrott & Eckhardt, 2018).

Broadly, there are several person-level demographic factors that have been epidemiologically linked to aggression. First, there is general consensus that male gender is a distal risk factor for aggressive behavior, particularly more extreme forms of aggression (Bettencourt & Miller, 1996; Knight, Guthrie, Page, & Fabes, 2002; Hyde, 1984; Moffitt, 2003). Indeed, since 2000, over 90% of the homicides in the United States were perpetrated by men (Fox & Fridel, 2017). Additionally, the relationship between age and aggression is more

nuanced, but cross-cultural studies have consistently found associations between aggression and age, such that the incidence of aggression and other forms of antisocial behavior peaks in late adolescence and early adulthood, and then gradually declines over the remainder of the lifespan (e.g., Blonigen, 2010; Moffitt, 2003). Of note, this negative relation between age and aggression is mirrored by an analogous positive relation between age and the personality trait Five Factor Model Agreeableness (Donnellan & Lucas, 2008), which is one of the most robust personality correlates of aggression (Jones, Miller, & Lynam, 2011; Miller & Lynam, 2001).

Although variables like age and gender are associated with a relative increased prevalence of aggression, virtually every human has the capacity to behave aggressively under certain circumstances, such as in acutely life-threatening situations or when necessary resources are scarce. This is sensible from an evolutionary perspective (e.g., Wrangham, 2018), and organisms across the animal kingdom are liable to behave aggressively when necessary for self-preservation or survival of their offspring (Blanchard & Blanchard, 2003). Given this imbalance between the universal potential for aggression and its relative infrequency, aggression theorists have sought to identify more proximal psychological processes that can help explain the occurrence of aggressive behavior.

Social Learning Theory and Aggression

Albert Bandura's social learning theory (Bandura, Ross, & Ross, 1961, 1963) was partially founded on the observation that children display behavior that they have not been directly reinforced to exhibit. Rather, they can learn complex social behaviors through observing an actor, and importantly, by encoding the reinforcement or punishment that the observed actor receives in response to their behavior (i.e., vicarious reinforcement; Bandura, Ross, & Ross, 1963). As individuals grow and learn, they encounter aggression-related experiences including

"exposure to models, verbal discussions, and discipline encounters," which allows them to "mentally represent their environments and themselves in terms of certain crucial classes of cognitions that include response-outcome expectancies, perceptions of self-efficacy, and standards for evaluative self-reactions" (p.781; Grusec, 1992).

Critically, Bandura postulated that there were multiple steps that took place in between the observation of a behavior and its imitation (Bandura, 1969). In brief, these components include paying attention to the behavior and its consequences, retaining the observation in memory, converting the observation from a symbolic representation to potential for action, and finally, the motivational component, wherein the situational demands may bring about the observed behavior. For example, even when a child has observed an aggressive action punished in the past, the proximal, appetitive influence of a reward may be sufficient to invoke this action in pursuit of a reward despite the potential for punishment (Bandura, 1965).

Bandura's proposition that the internalization and enactment of aggressive behavior happens over the course of several stages portended later efforts to take a granular perspective on the series of cognitive processes that can result in aggressive behavior.

Social Information Processing

A predominant model of social behavior that has emerged as a descendant of Bandura's social learning theory is the Social Information-Processing model (e.g., see Dodge & Crick, 1990). In original formulations of this model, an aggressive behavioral response represented the outcome of a series of cognitive processes that involved encoding and interpreting a situation, mentally searching for responses, and selecting the response. In contrast to the sequence proposed by Bandura, the focus of this model is less directly on how aggressive behavior is learned, but rather on the cognitive "steps" that occur between learning and the idiographic

enactment of aggressive behavior specific to an individual in a situation. Furthermore, this model represents an elaboration on Bandura's foundational steps by expounding upon the subcomponents that constitute each step. For example, a Social Information-Processing model of aggression separates Bandura's final "motivational step" into a much more individuated set of considerations. For example, in order to understand a child's motivation to aggress, one must first understand the interpretations of the situation that they have made, the response options that are available to them, and the identification of potential consequences that are likely to result from each option, which are likely multi-faceted and complex based on the child's social knowledge and the schemas that they have acquired (Crick & Dodge, 1994).

Expectancies of Aggressive Behavior

An important, but understudied component of the Social Information-Processing model that occurs during the response decision phase is called *expectancies* (or outcome expectancies in the language of the Social-Information Processing model). Expectancies are "cognitive representations summarizing an individual's learning about their environment," which in turn "guide behavior by allowing people to anticipate changes in the environment or predict potential outcomes of their behavior" (p. 120; Treloar, Pedersen, & McCarthy, 2015). Since individuals are motivated to act in ways that facilitate goal attainment, aggression can be understood as the behavioral result of a complex adjudication process, during which individuals reference their past experiences to determine the potential outcomes that they might expect as a result of behaving aggressively (or from not doing so). Based on the perceived likelihood of these outcomes and the subjective value placed on them (Rotter, 1954), individuals may enact an aggressive response as a behavioral strategy in pursuit of a desired goal.

Before reviewing the literature on outcome expectancies in aggression, it is important to mention that there is a larger literature on the role of the expectancies in substance use research (e.g., Jones, Corbin, & Fromme, 2001; Hull & Bond, 1986) that is useful to bring to bear on the current work. In this literature, it is common to parse several distinct types of expectancies that can occur for a given behavior: stimulus expectancies, response expectancies, and outcome expectancies (Vogell-Sprott & Fillmore, 1999). Stimulus expectancies are more specific to the substance use literature, as they are primarily concerned with properties about a substance being received. However, response expectancies and outcome expectancies are important to differentiate when considering aggressive behavior. Response expectancies refer to subjective, internal responses, while outcome expectancies refer to external, environmental contingencies (Treloar et al., 2015). In the context of substance use, an individual might expect that consuming hot coffee from a mug (stimulus expectancy) will result in increased energy and alertness (response expectancy), which may help facilitate a better presentation to co-workers (outcome expectancy).

While the Social-Information Processing Model uses the term outcome expectancies, both types of expectancies are pertinent to aggressive behavior and both occupy the same position within this model. As with substance use, there are a multitude of expectancies that can be forecasted as a consequence of aggression that can be perceived as either adaptive or maladaptive based on the degree to which they are related to goal attainment. For example, a positive response expectancy may be “if I push this person, then I will feel strong,” and a related positive outcome expectancy could be “if I push this person, then they will learn that they shouldn’t mess with me in the future.” In contrast, a negative response expectancy may be “if I

hit this person, then I will feel guilty," which could be related to the negative outcome expectancy "if I hit this person, then I will be looked down on by my peers."

In studies with children, positive effects have been observed such that more aggressive individuals tend to have more positive outcome expectancies for aggression (e.g., "you get what you want if you're a bully"; Bentley & Li, 1996; Crick & Dodge, 1996). Although the semantic distinction between types of expectancies was not incorporated at the time of this research, both types of expectancies were recognized as important precipitants of aggression. For example, Perry, Perry, and Rasmussen (1986) developed an outcome expectancy questionnaire for elementary school children that included response expectancies (e.g., self-reward – "if I shouted at this person, then I would feel very good") and outcome expectancies (e.g., tangible reward – "if I push my way to the front of the line, then I will get to drink water first"). When they administered this questionnaire to children separated into aggressive and non-aggressive groups, the aggressive group endorsed that after behaving aggressively, they expected more tangible reward, more peer approval, and less aversive treatment by peers in the future compared to their non-aggressive counterparts. Similarly, Hall, Herzberger, and Skowronski (1998) found that self-reported aggression is negatively related to expectations of feeling bad after aggressing (response expectancy), and negatively related to expectations about being punished and positively related to expecting that one will be respected by one's peers after aggressing (outcome expectancies). Importantly, these effects have also been demonstrated across methods: juvenile delinquent case managers provided ratings of proactive and reactive aggression in the incarcerated adolescent boys, and these ratings correlated positively with the adolescents' reports of feeling good about oneself after aggressing (response expectancy) and being liked by others after aggressing (outcome expectancy; Smithmyer, Hubbard, & Simons, 2000).

Although this model of aggressive behavior allows for an understanding of how aggression manifests in all individuals, it is inevitable that across the population, some individuals behave aggressively more often than others. As such, a primary way that these studies provide support for this model is through evidence that expectancies regarding aggression are generally more positive (and less negative) among certain groups, such as aggressive children (vs. non-aggressive children; Boldizar, Perry, & Perry, 1989; Perry, Perry, & Rasmussen, 1986), bullies (vs. non-bullies; Bentley & Li, 1996), and incarcerated youth (vs non-incarcerated youth; Barriga, Landau, Stinson II, Liau, & Gibbs, 2000). There is also evidence that individuals who are higher in trait irritability and aggressiveness are more likely to report that characters in a vignette are having aggression-related thoughts and imagining aggressive outcomes (Dill, Anderson, Anderson, & Deuser, 1997). However, besides this study by Dill and colleagues (1997), there is virtually no research on individual differences that predict variation in expectancies associated with aggression. This is unfortunate, since research on individual difference variables like personality traits has made a significant contribution to the literature on aggressive behavior. Moreover, there is support for hypothesized relations between personality traits like *Antagonism* and other components of the Social Information-Processing model, including the tendency to make hostile attributions when viewing neutral stimuli and generating a larger proportion of aggressive responses in a hypothetical social situation relative to non-aggressive ones (Miller, Lynam, & Jones, 2008). Taken together, these findings suggest that personality is a key individual difference that underlies patterns of aggression-related cognitions, but more specific links remain untested.

Individual Trait Differences as Predictors of Aggression

Up to this point, aggression has been discussed primarily as a behavioral construct, but aggression can also be operationalized as a trait (i.e., aggressiveness), which lies on a continuum and can be captured by self-report measures (e.g., Buss-Perry Aggression Questionnaire; Buss & Perry, 1992). Empirical support for the convergent validity of self-report trait aggressiveness is relatively strong: it has been linked to aggression in a laboratory setting (Bettencourt, Talley, Benjamin, & Valentine, 2006), parent, teacher, and peer reports of aggression (Österman et al., 1994; Tackett & Ostrov, 2010; Tackett, Waldman, & Lahey, 2009), and aggressive incidents in a forensic institution (Walters, 2007).

Beyond the tautological relations between trait aggressiveness and other indices of aggressive behavior, there are numerous other personality variables that have been closely linked to aggression, including psychopathy (e.g., Hare & McPherson, 1984), narcissism (e.g., Baumeister, Bushman, & Campbell, 2000), and sadism (e.g., Reidy, Zeichner, & Seibert, 2011). Rather than review the relations between each individual trait construct and aggression, it is parsimonious to invoke the Five Factor Model (FFM) of personality, since this structural model of personality was explicitly developed in order to be able to capture and organize all important trait-related individual differences (see John & Srivastava, 1999). Multiple meta-analyses suggest that the personality trait *Antagonism* (i.e., low FFM Agreeableness) and related complex personality profiles (e.g., narcissism, psychopathy) are among the most important person-level predictors of self-reported aggression and antisocial behavior (Jones, Miller, & Lynam, 2011; Miller & Lynam, 2001; Vize, Miller, Collison, & Lynam, 2018), aggressive behavior in laboratory paradigms (Hyatt et al., 2019), and criminal recidivism (Leistico, Salekin, DeCoster, & Rogers, 2008), as well as externalizing behaviors more broadly construed (Kotov et al., 2017).

However, personality traits like *Antagonism* are but one vantage point from which to view the individual differences that make up one's personality (McAdams & Pals, 2006). Numerous contemporary theories of personality, such as Cybernetic Big 5 Theory (CB5T; DeYoung, 2015) and Whole Trait Theory (WTT; Fleeson & Jayawickreme, 2015), argue for components of personality that exist below the facet level (e.g., characteristic adaptations, social-cognitive mechanisms), but robust empirical articulations of these mechanisms are unfortunately in relatively short supply. Put differently, although it is well-established that *Antagonism* is an important person-level predictor of aggression, the social-cognitive processes that can help "fill in the gaps" or provide a mechanistic link between trait *Antagonism* and aggressive behavior are largely unknown (Hampson, 2012).

The Current Study

Empirical research on aggression has benefitted from a range of perspectives, including models that provide an explanatory framework for the social-cognitive development and enactment of aggression, as well as trait-based models of personality that aim to identify the individual difference variables that are related to an elevated propensity for aggression. In many ways, these different orientations reflect the long-standing person-situation debate in personality psychology (e.g., Mischel, 1968). Psychologists operating from a social-cognitive approach have tended to view cross-situational consistency in behavior as relatively low, and thus consider the investigation of social-cognitive mechanisms by which behavior manifests (e.g., aggression in the Social Information Processing Model) to be the preferable avenue for understanding individual differences (Fleeson, 2012; Fleeson et al., 2015). On the other hand, a trait-based approach to personality considers cross-situational consistency to be relatively high, and thus psychologists operating from this perspective tend to view relative trait-standing as the preferred

way of understanding personality. Although both approaches are beneficial, each are limited, and many theorists have recognized that in order to advance personality science, aspects of each approach must be integrated (e.g., DeYoung, 2015; Fleeson et al., 2015; McAdams & Pals, 2006).

The primary goal of the current initiative is to develop a measure of aggression expectancies in adults. As with all psychological measures, establishing construct validity around a variable that a given measure is supposed to capture is paramount (e.g., Clark & Watson, 1995, 2019). Thus, part of this validity-construction process for this new measure will involve comparing aggression expectancies scores to self-reported aggression and antisocial behavior, as well as constructs from the personality literature that are well-established as personological correlates of aggression (e.g., *Antagonism*) as indices of convergent validity. By doing so, we are able to provide a test of the hypothesis that expectancies regarding the consequences of aggressive behavior is one mechanistic process by which *Antagonism* (and related complex trait profiles) is related to aggression. In other words, individuals who are high in trait *Antagonism* tend to behave more aggressively because on average, when engaging in the actuarial decision-making process about aggressing in a given situation, they tend to expect relatively more desirable consequences of behaving aggressively, and expect relatively fewer undesirable consequences of doing so. A series of six studies was conducted in order to construct and test the psychometric properties of this new Aggression Expectancies Questionnaire, which included two studies to ensure content coverage for item generation, three iterative studies to explore the factor structure of the items and examine indices of convergent and divergent validity, and finally a confirmatory sample where we examined the factor structure using confirmatory factor analysis.

CHAPTER 2

STUDY 1

Study 1 Overview

The aim of Study 1 was to examine the aggression expectancies that laypersons endorse in order to identify major themes to inform the expert consultation in Study 2, as well as subsequent item generation.

Study 1 Methods

Participants

Two-hundred and twenty five participants were recruited from Amazon's Mechanical Turk for a study on "how people think about what may happen if they act aggressively." In order to screen for invalid responders, we used several exclusion criteria: 1) unreasonably fast responses times based on pilot efforts by undergraduates and graduate students in our lab (i.e., less than 150 seconds), 2) responses to fewer than 50% of questions, and 3) an open-ended questions that required participants to "Name one way to score points in American football." Responses such as "touchdown," "field goal," and "extra point" were considered valid, while "score" and other vague or incorrect responses were considered invalid. This excluded 33 participants, for a final sample of N = 192 participants (55.2% male; mean [SD] age = 38.1 years [12.3]; 76.1% White or Caucasian, 8.3% Black or African-American, 7.3% Asian, 4.2% reporting more than one racial identity, and 4.1% Hispanic or Latino). All participants were compensated $0.75.

Materials

After completing a brief demographic measure, participants were presented with two separate introductory prompts and subsequent sets of open-ended questions, one set for positive aggression expectancies and the other for negative aggression expectancies. We randomized the order of presentation of these prompts across participants. The positive expectancy prompt read as follows: "*Sometimes people behave in certain ways because they expect that something positive/good will happen if they do so. For example, you may drink a cup of coffee because you expect that drinking coffee will make you feel more alert and awake. Or, you may give a gift to a friend because you expect that doing so will make the friend feel appreciated and cared for*." Next, participants were presented with a series of instructions to assess their expectancies associated with physical, verbal, and relational aggression. For example, the following prompt was used for positive expectancies of physical aggression: "*Imagine that you are physically aggressive (e.g., hitting, pushing) toward another person. What are some of the positive outcomes that you might expect? Please list 3 potential positive outcomes that are associated with being physically aggressive. These outcomes could include positive thoughts, emotions, and/or behaviors. For example, some outcomes may include 'feeling self-confident', or 'impressing the people around me.'*" Complete information about the content of these prompts and instructions are available at https://osf.io/fuz6h. Overall, each participant was asked to provide 18 total expectancies: nine positive expectancies, including three each for physical, verbal, and relational aggression, and nine analogous negative expectancies.

Coding

All valid participant responses are available at https://osf.io/fuz6h/. Each participant response was reviewed by the first author twice and qualitatively coded to identify common

themes. After each response was reviewed and common themes were identified, each response was re-reviewed to ensure that it was represented, to some capacity, in the major themes. Some responses that were more niche or idiosyncratic (e.g., "get an amusing reaction out of those that I acted aggressively toward"; "scaring a small child") did not fall neatly into the generated categories.

Study 1 Results

Positive Expectancies

Based on this informal content coding, positive expectancies tended to be characterized by one or more of 10 (non-exclusive) themes, including: 1) *counter/thwart attack* (e.g., "I will defend myself"), 2) *gain social capital* (e.g., "I will be respected by others"), 3) *establish dominance* (e.g., "I will intimidate others around me"), 4) *ending conflict* (e.g., "The other person will back down"), 5) *emotion/tension release* (e.g., "I will feel better after venting my anger"), 6) *positive feelings* (e.g., I will get a self-esteem boost"), 7) *achieving a goal* (e.g., "people will bend to my will"), 8) *justice* (e.g., "teach the other person a lesson"), 9) *demonstrate efficacy* (e.g., "I will learn what I am really capable of"), and 10) *hurt others* (e.g., the person will be hurt badly").

Negative Expectancies

Negative expectancies tended to be characterized by one or more of nine (non-exclusive) themes: 1) *formal punishment* (e.g., "I will get arrested"), 2) *physical harm to self* (e.g., "I may get beaten up"), 3) *physical harm to others* (e.g., "I may hurt them seriously"), 4) *emotional harm to self* (e.g., "I would lose respect for myself"), 5) *emotional harm to others* (e.g., "I will make them cry"), 6) *loss of social capital* (e.g., "people will think poorly of me"), 7) *damage to relationships* (e.g., "I will alienate friends and family"), 8) *increased likelihood of future harm*

(e.g., “I will gain an enemy”), and 9) *immediate retaliation* (e.g., “they will come right back at me”).

Study 1 Discussion

Study 1 took a crowd-sourcing approach to distilling the major aggression expectancies held by laypersons. Review of responses suggested that individuals hold a wide range of positive and negative expectancies that may follow acts of physical, verbal, and relational aggression, and numerous expectancies were reported for both the aggressor (e.g., *gain social capital*) and the victim (e.g., *emotional harm to other*). Positive and negative response expectancies were identified (e.g., *emotion/tension release*, *emotional harm to self*), as well as positive and negative outcome expectancies (e.g., *justice*, *damage to relationships*). This suggests that participants identified aggression expectancies that were immediate and more subjective in nature (e.g., proximal positive feelings), as well as more distal, external consequences (e.g., impact on one’s social milieu).

CHAPTER 3

STUDY 2

Study 2 Overview

The aim of Study 2 was to maximize content validity by consulting experts in the study of aggression working from different backgrounds (i.e., developmental, cognitive, social) and requesting that they provide feedback on the Study 1 aggression expectancy themes. This was viewed as an essential step for establishing content validity of the AEQ (Haynes, Richard, & Kubany, 1995), given that various aggression experts are likely to have unique perspectives and insights on this topic.

Study 2 Methods

Participants

Based on information available as of July 17, 2019, we collected the names and email addresses for the 30 members of the editorial board of *Aggressive Behavior* and the 80 members of the editorial board of *Psychology of Violence*. Each individual was sent a brief email requesting that they complete a very brief survey about aggression expectancies. Out of these 110 experts who were contacted, N = 15 responses were received.

Measures

Participants were directed to a survey where they were able to see a list of all of the positive or negative expectancies distilled in Study 1. Positive or negative expectancies were presented in random order. Participants were then asked to respond to an open-ended question:

"In your view, are there any positive (or negative) expectancies of aggressive behavior that are not listed above? If so, please describe them, and be specific in your response."

Study 2 Results

Full, anonymous data from the expert responses are available at https://osf.io/fuz6h/. Review of the expert responses generally suggested that the themes generated in Study 1 were sufficient in capturing the majority of aggression expectancies. 11/15 (73%) of participants reported at least one additional suggestion for expectancies to consider adding to our list. Several of the most notable suggestions for positive aggression expectancies included "sticking up for someone else," "connecting with others," "the person we helped now owes us a favor," "taking something from someone else," and "sexual arousal." Notable suggestions for negative aggression expectancies included "getting grounded," "expecting no change," "going against my religious beliefs," "increased hostile world view," and "loss of authority."

Study 2 Discussion

The goal of Study 2 was to increase content validity of the forthcoming measure of aggression expectancies by consulting with aggression and violence researchers about the expectancies that may be omitted, overlooked, or that were not represented in the Study 1 responses. Overall, this expert consultation did not result in the addition of any major new positive or negative aggression expectancy themes; virtually all of the specific suggestions provided by these experts can be captured (to various degrees) by the themes identified in Study 1. For example, the positive expectancy "sticking up for someone else" is an example of an expectancy in line with the larger theme of *justice*. The positive expectancy "taking something from someone else" maps onto the larger theme of *achieving a goal.* Analogously, the negative expectancy "getting grounded" is in line with the larger theme of *formal punishment*. The

negative expectancy "going against my religious beliefs" is an example of an item that would likely be captured by the themes *emotional harm to self* and *damage to relationships*. Thus, this study was a small but important component of the current project, given the positive feedback from multiple, independent experts (Haynes et al., 1995).

Importantly, many of themes identified are directly in line with those reported in the previously work from the child/adolescent literature on aggression expectancies. For example, the outcome expectancy questionnaire for elementary school children developed by Perry and colleagues (1986) included the themes self-reward (paralleled by the current themes of *positive feelings* and *emotion/tension release*), tangible reward (paralleled by *achieving a goal*), adult and peer approval (paralleled by *gain social capital*), and victim suffering (paralleled by *physical harm to others* and *emotional harm to others*). Similarly, in the study by Hall and colleagues (1998), the expectancies themes included feeling bad after aggressing (paralleled by *emotional harm to self*) and being punished (paralleled *formal punishment*). This convergence between previously identified expectancies from the child literature and those endorsed by laypersons (and confirmed by experts) provided a base from which to construct items.

Table 1

Condensed Themes from Studies 1 and 2

Positive Expectancies	Sample Item	Negative Expectancies	Sample Item
Counter/prevent attack	"People will learn that there are costs to messing around with me."	*Formal punishment*	"I will get in trouble with an authority figure."
Gain social capital	"I will gain the respect of those around me."	*Physical harm to self*	"I will get beaten up."

Emotion/tension release	"I will feel better after releasing my anger."	*Physical harm to others*	"I may seriously hurt the other person."
Positive feelings	"It will make me feel good about myself."	*Emotional harm to self*	"I would feel like a jerk afterwards."
Achieving a goal	"I'll get what I want from others."	*Emotional harm to others*	"The other person will feel emotionally upset."
Justice	"I will be punishing someone who deserves to be punished."	*Damage to relationships*	"People close to me (e.g., my friends and family) would be disappointed in me."
Demonstrate efficacy	"Others will see what I'm capable of doing."	*Increased likelihood of future harm*	"I will have to watch my back in the future."

CHAPTER 4

STUDY 3

Study 3 Overview

The aims of Study 3 were threefold: 1) to generate a comprehensive, but manageable (i.e., not unwieldy or redundant) set of items that captured the essence of the aggression expectancy themes identified in Study 1 and 2, 2) administer them to a sample of community adults to explore the factor structure exhibited by these items, and 3) begin to establish construct validity by examining the convergent and divergent relations that these factors exhibit with important variables in the aggression expectancy nomological net, including self-reported aggression and FFM personality traits. The first aim was accomplished by four of the current authors (CSH, DSC, DRL, JDM) generating 3-5 items for each of the 19 themes. In doing so, we pared down a number of themes that, upon review, were conceptually very similar and resulted in highly similar item content (e.g., *gain social capital* and *establish dominance*; *physical harm to others* and *emotional harm to others*). After paring down the number of themes into seven positive and seven negative themes, six items were included for each for a total of 84 items. Table 1 contains a list of these themes, as well as a sample item for each. Items were selected based on several considerations, including brevity, clarity, breadth of content, and the degree to which the items captured the essence of the themes.

Study 3 Methods

Participants

Study 3 participants were recruited from Amazon's Mechanical Turk. A screening procedure was used in order to maximize the validity of our sample, as well as to oversample for individuals with a history of physical and verbal aggression. First, 600 individuals were recruited to take part in a very brief (i.e., < two minutes) screening questionnaire to determine eligibility for a subsequent, longer battery. Participants responded to a series of four, unrelated questions (e.g., How many days per week do you exercise? Have you voted in a political election in the past four years?), one of which was "Since the age of 18, have you gotten into a physical or verbal altercation with another person?" Finally, participants answered an open-ended question about "the steps they use to take a shower." In order to be eligible for participation in the full study, participants needed to respond affirmatively to the question about aggression since 18, and they needed to provide a valid response to the open-ended question about taking a shower. The first and senior author read each response to the shower, and coded them as either valid or invalid. Responses had to be deemed valid by both authors in order for the participant to be invited back for the full battery.

Based on these initial screens, 375 of these participants were invited to complete the full assessment battery. Of note, in the informed consent for this longer battery, participants were informed that their data would be considered invalid if they failed more than one of several embedded attention checks (e.g., "Select the second response option"). There were N = 335 valid responses, which constituted the final Study 3 sample (57.9% female; mean [SD] age = 37.3 [12.0]; 72.6% White or Caucasian, 14.3% Black or African-American, 5.5% Asian, 4.3%

Hispanic or Latino, 2.4% reporting more than one racial identity, and less than 1% each of American Indian and Pacific Islander).

Aggression Expectancy Questionnaire

The first iteration of the Aggression Expectancy Questionnaire comprised the 84 items generated to capture the 14 themes from Study 1 and 2. Participants were presented with the following prompt: "*People often behave in certain ways because they expect a certain consequence to occur. For example, you may eat a snack because you expect the consequence will be that you feel less hungry afterwards. Below is a series of statements about* ***possible consequences of being aggressive****. Please respond on a 1-5 scale to indicate how likely you think each consequence is for you.*" Participants were presented with a Likert scale with the following labels: 1 = Very Unlikely (i.e., it is very unlikely that this will happen if I behave aggressively), 2 = Somewhat Unlikely, 3 = Neither Likely nor Unlikely, 4 = Somewhat Likely, and 5 = Very Likely (i.e., it is very likely that this will happen if I behave aggressively). We included the running statement "*If I am aggressive toward other people, then I expect that…*" above each cluster of 15 items.

Reactive Proactive Aggression Questionnaire (RPAQ)

The RPAQ is a 23-item self-report measure of aggression. It includes a proactive aggression ($\alpha = .94$) and reactive aggression ($\alpha = .86$) subscale (Raine et al., 2006).

Abbreviated Crime and Analogous Behavior Scale

Participants completed an abbreviated version of the Crime and Analogous Behavior Scale (Miller & Lynam, 2003) that only included items that life history of aggressive and violent behavior. This abbreviated measure comprised 12 self-report items that measured endorsement of past physical fights, intimate partner violence, and violent and non-violent crime (e.g.,

attacking someone with a gun or knife; breaking into a house), which were summed to create an antisocial behavior composite.

FFM Domain Rating Form

The FFM Domain Rating Form is a 10-item self-report measure designed to capture individual standing the low and high end of each FFM domain (i.e., two items per domain). For example, for Agreeableness, participant reported on a scale of 1-5 (Strongly Disagree to Strongly Agree) for how much they "tend to be sympathetic to others, willing to help others in need, and believe that others are generally honest and well-intentioned. They are also sincere, humble, and compliant with authority" for the high end of Agreeableness; then they reported how much they "tend to be more arrogant, self-centered, and distrustful of others. They also view others through a competitive rather than cooperative lens in social situations. They also tend to be more aggressive, are willing to manipulate others for their benefit, and believe that they are superior to others" for the low end of Agreeableness. The low-end responses were reverse-coded to create a mean score for each domain. This measure of the FFM demonstrates large correlations to longer measures (Sleep, Lynam, & Miller, in press) and exhibited adequate internal consistency given the small number of items per scale (correlations between scale by scale: Neuroticism = .34; Extraversion = .41 Openness = .18; Agreeableness = .33; Conscientiousness = .36).

Analyses

A series of analyses were conducted with the goals of identifying the best items, culling unnecessary and underperforming items in order to make a briefer measure, and exploring the factor structure of these items. First, descriptive statistics were examined for each item, with the goal of eliminating any items that did not receive responses across the full range of response options (i.e., did not have a range of 1-5). Second, bivariate relations were computed between all

items, pairs of items that were correlated at $r \geq .70$ were identified in order to identify items that were redundant with one another. In these cases, items were eliminated that exhibited the largest number of these high correlations (i.e., the most redundancies). For example, if item X and item Y were correlated at $r = .70$, but item X was also correlated with three additional items $\geq .70$ while item Y only exhibited this high correlation with item X, then item Y was retained. In situations where the items demonstrated an equal number of redundancies (e.g., item X and item Y were correlated $\geq .70$ but neither item exhibited a correlation of this magnitude with any other items), the first and senior authors consulted and made a decision about which item to eliminate based on considerations such as item parsimony and content representation (i.e., ensuring a breadth of content within a given theme).

Lastly, to explore the factor structure of the reduced set of items, we conducted a series of exploratory factor analyses (EFA) using Principal Axis Factoring and Direct Oblimin Rotation. Parallel Analyses and Velicer's MAP test were used in order to inform the interpretation of the various factor solutions. Subsequently, we conducted EFA with maximum likelihood estimation in order to provide additional empirical evidence regarding the fit of various solutions.

Study 3 Results

Item Culling

Examination of the item-level descriptive statistics suggested that all of the aggression expectancy items administered had a range 1-5. Thus, no items were eliminated based on this criterion. In terms of redundant items (i.e., pairs of items correlated at $r \geq .70$), 56 pairs of items were observed and 18 were eliminated using the aforementioned process. The reason for the discrepancy between the number of pairs identified (i.e., 56) and the smaller number of items

deleted (i.e., 18) is that there were a number of "highly redundant" items that, once eliminated, substantially reduced the number of redundant pairs.

Exploratory Factor Analyses of Remaining Items

The remaining 66 items were entered into an EFA with no restrictions on the number of factors to extract. The scree plot of this initial analysis strongly suggested a two-factor solution. Velicer's MAP test suggested a seven-factor solution, and Parallel Analysis also suggested a seven-factor solution (first seven eigenvalues [cumulative variance]: 17.6 [26.7%], 13.5 [47.1%], 2.6 [51.0%], 2.4 [54.6%], 1.9 [57.5%], 1.5 [59.8%], 1.3 [61.7%]). Given this range of possible solutions, we conducted EFA with maximum likelihood estimation to examine the fit of solutions from one-factor through seven-factor. The fit of the four-factor through seven-factor solutions were deemed acceptable (i.e., TLI >.799), with the five- and six-factor solutions exhibiting the best fit according to BIC (five-factor BIC = -7169.78; six-factor BIC = -7163.05). Thus, the four- through seven-factor solutions were examined for content.

At the four-factor level, a "general positive" factor (sample item with high loading: "Later on, I will be happy that I did so") emerged. Three negative expectancy items emerged that were deemed "harm to self" (e.g., "I will get beaten up"), "harm to victim" (e.g., "I may seriously hurt the other person"), and "damage to self-image/reputation" factor (e.g., "I would be disappointed in myself for doing so"). At the five-factor level, the "general positive," "harm to self," and "damage to self-image/reputation" factors were virtually unchanged from the four-factor level. The "harm to victim" factor was partially retained, but it split into an additional factor deemed "trauma to self/other" that was characterized by a mix of self- and other-directed harm that was severe in nature (e.g., "The other person will never be the same again").

At the six-factor level, the "general positive" factor fractured into two subordinate factors: the first was deemed "dominance/intimidation" (e.g., "Others will know not to mess with me in the future"), and the other was deemed "sweetness/justice" (largest item loadings: "I will be doing good for others"). The "harm to self," "harm to victim," "damage to self-image/reputation," and "trauma to self/others" factors were virtually unchanged from the five-factor level. Finally, at the seven-factor level, the "sweetness/justice" factor split into a smaller "justice" factor (e.g., "I will be acting in the name of justice") and a single-item "sweet" factor ("It will feel sweet"). The remainder of the factors from the six-factor level were virtually unchanged.

Factor Relations to External Criteria

Table 2 reports the correlations between the aggression expectancy factor scores at the four- through seven-factor level with relevant external criteria. In general, the positive expectancy factors at each level (i.e., "general positive", "dominance", "sweetness/justice") scores exhibited large (i.e., $r = .30$) to very large (i.e., $r \geq .40$) positive relations with self-reported proactive and reactive aggression, as well as smaller relations with an antisocial behavior composite that were medium (i.e., $r = .20$) to large in magnitude. Analogously, these positive expectancy factors exhibited large to very large, negative relations with Agreeableness, and medium to large, negative relations with Conscientiousness.

Overall, the negative expectancy factors (i.e., "harm to self," "harm to victim," "damage to self-image/reputation," "trauma to self/other") exhibited smaller and less consistent relations with the external criteria variables, and only a small number of relations met the threshold of a medium effect size. The "damage to self-image" negative expectancy factors exhibited small-to-medium, negative relations with proactive aggression and positively relations with

Agreeableness. The "harm to victim" negative expectancy factor bore medium-to-large, positive relations to proactive and reactive aggression, and small-to-medium, negative relations with Agreeableness.

Study 3 Discussion

In Study 3, a large item pool of aggression expectancies was administered to a sample of adults who endorsed a history or physical or verbal aggression. Redundant items were then removed and the factor structure of the remaining items was explored, as well as the factors' relations with relevant criteria variables (e.g., antisocial behavior, FFM domains). Several major themes emerged. First, items from the majority of the themes generated in Study 1 exhibited high loadings onto the factors that emerged (i.e., as opposed to items from only one or two of the themes being consistently the highest loading). This was encouraging, as the aim was to capture a broad array of positive and negative expectancies rather than a few narrow ones. At the same time, items from several themes tended to consistently load onto the same factors, such as the positive expectancy themes *emotion/tension release* and *positive feelings* (i.e., the "sweetness/justice" factor); this is sensible given that both of these capture positively-valenced affective responses following the enactment of an aggressive behavior. As another positive expectancy example, items from the themes *counter/thwart attack* and *gain social capital* tended to load on highly onto the same factor (i.e., "dominance"). A similar pattern was observed for the negative expectancy themes. Items from the *physical harm to others* and *emotional harm to others* themes tended to load onto similar factors (i.e., "harm to victim"); items from the *physical harm to self* and *increased likelihood of future harm* tended to load on the same factors (i.e., "harm to self").

Second, the factors unfolded such that there tended to be a split between intrapersonal expectancies (e.g., "positive feelings"; "damage to self-image") vs. interpersonal expectancies (e.g., "interpersonal dominance"; "harm to victim"). This is consistent with the major distinction in the expectancy literature made between response expectancy (more intrapersonal in nature) and outcome expectancy (more interpersonal in nature), and was viewed as a major guiding heuristic for subsequent iterations of the measure. Third, the factors that emerged at the most complex levels (i.e., six-, seven-factor) were often difficult to interpret; for example, based on the content, it would be sensible for the items that load onto the "trauma to self/other" factor (e.g., "the other person might be permanently injured," "I will not be able to live with myself") to have loaded onto the factors "harm to victim" and "damage to self-image/reputation," respectively. Another example is the fracture of the "sweetness/justice" factor at the seven-factor level into separate "justice" and "sweetness" factors, which was viewed relatively less elegant than the combined factor at the six-factor level.

Fourth, in general, expected relations emerged between the positive expectancy factors and the external criteria, supporting convergent and divergent validity of these factors. Participants who reported higher levels of proactive and reactive aggression also tended to endorse expecting that there would be positive consequences to aggressing, and the same was true for antisocial behavior. Moreover, Agreeableness exhibited negative relations with the positive expectancies, such that individuals who are generally antagonistic tended to report that positive consequences were relatively likely following aggression. The same was true to a lesser degree for Conscientiousness – individuals who are generally irresponsible/disinhibited tended to report a higher likelihood of positive consequences following aggression. Alternately, the factors tended to demonstrate null-to-small relations with Neuroticism, Extraversion, and Openness.

Lastly, there was slimmer evidence for convergent and discriminant validity of the negative expectancy factors. For example, the "damage to self-image" factor exhibited expected relations: this factor was negatively related to proactive aggression (i.e., individuals who expect aggressing to be damaging to their self-image report lower levels of proactive aggression) and positively related to Agreeableness (i.e., more agreeable individuals tend to endorse that aggressing will damage their self-image). The "harm to victim" factor also exhibited expected relations: this factor was positively related to proactive and reactive aggression (i.e., self-reported aggression is related to expectations that you will harm the other person if you are aggressive), and negatively related to Agreeableness (i.e., more agreeable individuals report lower levels of expectations that that will harm someone if aggressive). Beyond these relations, the negative expectancy factors tended to display null-to-small relations with the external criteria.

Factor Identification

Based on these initial findings, we anticipated that the structure of the aggression expectancy measure would be divided by positive and negative expectancies, as well as several lower-order factors that may be categorized into intrapersonal and interpersonal factors. The precise number and nature of these factors was viewed as one of the primary goals of subsequent studies. The guiding dialectic for determining how many factors to retain for was to strike a balance between 1) reducing the number of items in hopes of ultimately yielding a relatively concise measure (i.e., less than ~25 items) with minimal redundancy, and 2) capturing the breadth of the major themes identified in the initial theme construction phase in Studies 1 and 2. In other words, the major consideration was to distill the basic conceptual threads that bind similar themes (e.g., intrapersonal vs. interpersonal), such that we avoid redundancy and

unnecessary complexity, while also maximizing coverage by avoid creation of "bloated specific" factors that capture overly narrow constructs.

The results of Study 3 suggest that many of the items from different themes load onto similar factors, and thus items from fourteen different factors would not be necessary. After reviewing the positive expectancy item content and factor solution at various levels, a two-factor solution was chosen as most appropriate, consisting of an "intrapersonal" and an "interpersonal" factor. These represented updated names for the "sweetness/justice" and "dominance" factors, respectively, as the former comprised almost exclusively intrapersonal items, and the latter comprised predominantly interpersonal items. After reviewing the negative expectancy items, we were initially compelled to parallel the positive expectancy scale and move forward with an anticipated two-factor "intrapersonal" vs. "interpersonal" solution. For example, the "harm to self" and "harm to victim" factors were clearly intrapersonal and interpersonal factors, respectively. However, the "damage to self-image/reputation" factor was consistently extracted at the three- through seven-level solution, and it was predominantly composed of items from the *emotional harm to self* and *damage to relationships* themes, representing a blend of intra-/interpersonal expectancies. As another example, the "trauma to self/other" factor was a blend of intra-/interpersonal items, but this factor was even more puzzling, given that the content seemed to be redundant to the "harm to self" and "harm to victim" factors. Given these considerations, we erred on the side of inclusivity and moved forward with an anticipated three-factor solution for the negative expectancy items modeled after the factors identified at the four-factor level. These factors were titled "Harm to self," "Damage to self-image/reputation," and "Harm to victim."

Table 2

Correlations Between Aggression Expectancy Factors and External Criteria in Study 3

Aggression Expectancy Factors	PA	RA	ASB	*N*	*E*	*O*	*A*	*C*
General Positive (4.1)	**.56**	**.47**	**.28**	-.01	.09	-.10	**-.41**	**-.27**
Harm to Self (4.2)	.16	.14	.11	.01	-.07	<.01	.06	-.05
Damage to Self-Image/Reputation (4.3)	**-.22**	-.05	-.12	<.01	-.13	.10	**.30**	.19
Harm to Victim (4.4)	.09	.01	.03	-.04	.06	.11	-.10	-.10
General Positive (5.1)	**.56**	**.47**	**.28**	-.01	.09	-.10	**-.40**	**-.26**
Harm to Self (5.2)	.18	.17	.13	.02	-.08	<.01	.03	-.07
Damage to Self-Image/Reputation (5.3)	**-.22**	-.06	-.12	.01	-.13	.09	**.32**	.19
Harm to Victim (5.4)	.06	.03	.01	-.03	.01	.13	-.09	-.09
Trauma to Self/Other (5.5)	.09	-.03	.05	-.03	.09	-.01	.01	-.03
Dominance (6.1)	**.49**	**.44**	**.25**	-.01	.07	-.08	**-.35**	**-.21**
Harm to Self (6.2)	.14	.15	.11	.02	-.08	<.01	.06	-.04
Damage to Self-Image/Reputation (6.3)	-.12	-.01	-.07	.01	-.13	.08	**.27**	.13
Harm to Victim (6.4)	.07	.02	.02	-.02	.01	.13	-.10	-.11
Trauma to Self/Other (6.5)	.15	.03	.07	-.04	.09	-.03	-.03	-.06
Sweetness/Justice (6.6)	**.61**	**.43**	**.31**	<-.01	.10	-.11	**-.44**	**-.33**
Justice (7.1)	**.52**	**.38**	**.28**	-.03	.09	-.10	**-.38**	**-.28**
Harm to Self (7.2)	.12	.14	.10	.02	-.09	<.01	.07	-.03
Damage to Self-Image/Reputation (7.3)	-.10	.01	-.05	.01	-.13	.07	**.25**	.12
Harm to Victim (7.4)	.15	.08	.06	-.02	.02	.11	-.15	-.15
Trauma to Self/Other (7.5)	**.21**	.08	.11	-.04	.10	-.03	-.06	-.09
Dominance (7.6)	**.44**	**.42**	**.22**	-.01	.05	-.07	**-.31**	-.18
Sweet (7.7)	**.58**	**.44**	**.27**	.07	.08	-.13	**-.41**	**-.27**

Note: PA = proactive aggression; RA = reactive aggression; ASB = antisocial behavior; N = neuroticism; E = extraversion; O = openness to experience; A = agreeableness; C = conscientiousness; relations ≥ |.20| are **bolded**.

CHAPTER 5

STUDY 4

Study 4 Overview

The goal of Study 4 was to refine the aggression expectancy measure by administering the reduced set of items from Study 3 to another sample of community adults, and examining the factor structure and criteria relations exhibited by these new factor scores. The analytic strategy was largely identical to Study 3, except for an expanded set of criteria variables examined. Specifically, in addition to self-reported aggression and antisocial behavior, participants also reported on the 30 facets of the FFM, which contain more specific personality-related content than the broad domains. Additionally, fit statistics were generated for different factor solutions using exploratory factor analyses with maximum likelihood (EFA-ML) fitting procedure with promax rotation in order to assist in adjudicating between different factor solutions.

Study 4 Methods

Participants

Study 4 participants were recruited from Amazon's Mechanical Turk.
As in Study 3, a screening procedure was implemented in order to maximize the validity of the sample. Five hundred individuals were recruited to take part in a very brief (i.e., < two minutes) screening questionnaire to determine eligibility for a subsequent, longer battery. Participants were asked to respond to an open-ended prompt about "what you eat for a typical breakfast." The first and senior author read each response and coded them as either valid or invalid. Responses

had to be deemed valid by both authors in order for the participant to be invited back for the full battery. 448 participants were deemed valid responders.

Subsequently, 350 of these participants were invited to complete the full assessment battery. In the informed consent for this longer battery, participants were informed that their data would be considered invalid if they failed more than one of several embedded attention checks (e.g., "Select the second response option"; N = 6) or validity scales (N = 22). Ultimately, there were N = 322 valid responses, which constituted the final Study 4 sample (51.1% male; mean [SD] age = 38.3 years [11.5]; 76.5% White or Caucasian, 6.5% Black or African-American, 5.1% reporting more than one racial identity, 4.0% Hispanic or Latino, 2% Asian, and less than 1% American Indian).

Aggression Expectancy Questionnaire Item Selection

To determine which items to include in the next iteration of this measure, the highest loading items on each of the five factors ("intrapersonal" and "interpersonal" positive expectancies; "harm to self," "damage to self-image/reputation," and "harm to victim" negative expectancies) from Study 3 were identified. In an effort to create an overly inclusive pool of items, eight items per factor were selected that met criteria of being sufficiently high loading (i.e., loading >.40) and also were believed to capture the essence of the factors while also including sufficient breadth of content. As one guiding heuristic to increase item heterogeneity, we elected not to include more than three items from any of the major thematic categories identified in Studies 1 and 2. As one example of these decisions, the item "I will be punishing someone who deserves to be punished" loaded onto the "intrapersonal" factor, but was not selected because there were several other item options from the *justice* theme that were more direct and concise (i.e., "I will be doing good for others"). A final issue is that only four items

loaded >.40 onto the "harm to victim" factor at the three-factor level. Each of these four items were included, and the first and senior authors wrote and selected four new items aimed at capturing this theme based on content coverage.

Aggression Expectancy Questionnaire

The second iteration of the Aggression Expectancy Questionnaire comprised 40 items, 36 of which were administered in Study 3 and the additional four items newly generated for Study 4. Participants were presented with the following prompt: "*People often behave in certain ways because they expect a certain consequence to occur. For example, you may eat a snack because you expect the consequence will be that you feel less hungry afterwards. Below is a series of statements about* ***possible consequences of being aggressive****. Please respond on a 1-5 scale to indicate how likely you think each consequence is for you.*" Participants were presented with a Likert scale with the following labels: 1 = Very Unlikely (i.e., it is very unlikely that this will happen if I behave aggressively), 2 = Somewhat Unlikely, 3 = Neither Likely nor Unlikely, 4 = Somewhat Likely, and 5 = Very Likely (i.e., it is very likely that this will happen if I behave aggressively). We included the running statement "*If I am aggressive toward other people, then I expect that…*" above each cluster of 15 items.

Reactive Proactive Aggression Questionnaire

The Reactive Proactive Aggression Questionnaire is a 23-item self-report measure of aggression that includes a proactive aggression and reactive aggression subscale (Raine et al., 2006). Both subscales displayed adequate reliability estimates (Proactive Aggression $\alpha = .87$; Reactive Aggression $\alpha = .84$).

Crime and Analogous Behavior Scale

Participants completed an abbreviated version of the Crime and Analogous Behavior Scale (Miller & Lynam, 2003) that only included items that life history of aggressive and violent behavior. This abbreviated measure comprised 12 self-report items that measured endorsement of past physical fights, intimate partner violence, and violent and non-violent crime (e.g,. attacking someone with a gun or knife; breaking into a house), which were summed to create an antisocial behavior composite.

IPIP-NEO-60

Finally, participants completed the 60-item self-report IPIP-NEO-60 (Maples-Keller et al., 2019) in order to measure 30 facets of the FFM. Of note, this measure was supplemented to include the 12 Agreeableness items from the IPIP-NEO-120 (Maples, Guan, Carter, & Miller, 2014) that are not included in the shorter version of the measure. In total, 72 items were used to measure the FFM: four items per Agreeableness facet and two items per facet for the other four domains. The domain scores exhibited adequate internal consistency (Neuroticism $\alpha = .91$; Extraversion $\alpha = .78$; Openness $\alpha = .75$; Agreeableness $\alpha = .89$; Conscientiousness $\alpha = .90$), and facet reliability estimates ranged from $\alpha = .50$ to .94 (median = .82).

Validity Scales

The two validity scales from the Elemental Psychopathy Assessment (Lynam et al., 2011) were interspersed throughout the self-report measures above. These scales assess overly virtuous responses (e.g., “I have never told a lie to anyone”) and endorsement of highly infrequent behavior (e.g., “I often forget my middle name”). Data from participants who scored above endorsed more than two or three, respectively, of these statements were not used (N = 22).

Analyses

A series of analyses were conducted with the goals of culling items in order to make a briefer measure and exploring the factor structure of these items. First, items were eliminated that did not receive responses across the full range of response options (i.e., did not have a range of 1-5). Second, bivariate relations were computed between all items, and pairs of items correlated at $r \geq .70$ were identified and deemed redundant. In these cases, the items that exhibited the largest number of these high correlations (i.e., the most redundancies) were eliminated. To explore the factor structure of the reduced set of items, a series of exploratory factor analyses (EFA) using principal axis factoring and direct oblimin rotation were conducted. Parallel Analyses and Velicer's MAP test were also used in order to inform the interpretation of the various factor solutions. In order to assist in adjudicating between various solutions, we conducted EFA with maximum likelihood fitting procedure.

Finally, to examine the relative and absolute similarity of the lower-order factors, bivariate relations were computed between all factors and the relevant external criteria variables (i.e., proactive and reactive aggression, antisocial behavior, FFM domains and facets). Double-entry intraclass correlations (McCrae, 2008) were used to compare the absolute similarity of the factors that emerged from these solutions in terms of their relations to these external criteria.

Study 4 Results

Item Culling

All items exhibited the full range of endorsement. 31 pairs of redundant items (i.e., correlated at $r \geq .70$) were observed. Notably, 18 of these pairs were between items from the factor "harm to victim" that were newly generated for this study. In this case, the four items from this factor with the lowest number of redundancies were retained. For the other 13 pairs, we

eliminated the member of the pair with the higher number of redundancies, and thus 13 items were deleted on this basis.

Exploratory Factor Analyses of Remaining Items

The remaining 27 items were entered into an EFA with no restrictions on the number of factors to extract. The scree plot of this initial analyses suggested a three or five-factor solution, Velicer's MAP test suggested a four-factor solution, and Parallel Analysis suggested a seven-factor solution. Given this range of findings, we conducted EFA-ML with promax rotation in order to examine fit at the three- through seven levels of factor solution.

The three factor solution (RMSR = .07, TLI = .730, RMSEA [90% C.I.] = .110 [.102 to .113], BIC = -287.62) displayed notably worse fit than either the four-factor solution (RMSR = .04, TLI = .828, RMSEA [90% C.I.] = .088 [.079 to .092], BIC = -600.36) or the five-factor solution (RMSR = .03, TLI = .891, RMSEA [90% C.I.] = .071 [.061 to .075], BIC = -741.04), which both exhibited marginal fit (see Brown, 2014; Sellbom & Tellegen, 2019). The six-factor solution (RMSR = .02, TLI = .917, RMSEA [90% C.I.] = .062 [.052 to .067], BIC = -743.77) and seven-factor solution (RMSR = .02, TLI = .944 RMSEA [90% C.I.] = .051 [.040 to .057], BIC = -734.79) exhibited good model fit. Thus, we present results regarding the content of the factors at the four- through seven-factor factor solutions.

At the four-factor level, a "positive intrapersonal" and "positive interpersonal" factor emerged, the latter of which included several negatively-loading negative expectancy items that loaded onto the "damage to self-image/reputation" factor in the Study 3 (e.g., "I would be disappointed in myself for doing so"). "Harm to self" and "harm to victim" factors also emerged at this level. At the five-factor level, factors emerged that were highly consistent with the five factors that we anticipated based on those identified in Study 3: "positive intrapersonal,"

"positive interpersonal," "harm to self," "harm to victim," and "damage to self-image/reputation." Almost across the board, the items that loaded >.40 onto these five factors were items that we included with the intention of capturing these respective factors.

At the six-factor level, each of the previous factors was retained, and an additional, smaller "harm to self" factor emerged that comprised two items related to getting into trouble and getting hurt. Finally, at the seven-factor level, each of the factors from the six-factor level emerged, and the seventh factor was very small, as only one item demonstrated a factor loading > .40 and this item also strongly loaded onto the "positive intrapersonal" factor. Thus, despite displaying satisfactory fit indices, the six- and seven-factor solutions were viewed as relatively unsatisfying, given that each had a very small number of items with meaningful factor loadings and the content was incoherent. As such, we decided to move forward examining the four- and five-factor solutions.

Factor Interrelations

At the four and five-factor level, the "positive intrapersonal" and "positive interpersonal" factors exhibited large, positive relations with one another (*r*s = .47, .50, respectively). Similarly, at the four- and five-factor level, the "harm to self" and "harm to victim" factors also exhibited large, positive relations (*r*s = .49, .45, respectively). The "damage to self-image/reputation" factor that emerged at the five-factor level exhibited large, positive relations to "harm to self" and "harm to victim factors (*r*s = .48 and .41, respectively), as well as large, negative relations to the "positive intrapersonal" factor at the four- and five-factor levels (*r*s = -.66 and -.47 respectively).

Factor Relations to External Criteria

In Table 3, we report the correlations between the positive and negative expectancy factor scores at the four- and five-factor level with relevant external criteria. The four- and five-factor level positive expectancy factors exhibited medium-to-very large relations (i.e., *r*s .20 to >.40) with proactive and reactive aggression, and small-to-medium (i.e., *r*s .10 to .20) relations with the antisocial behavior composite. These positive expectancy factors also exhibited very large, negative relations to Agreeableness, as well as small-to-medium, negative relations to Conscientiousness and Openness. Similarly, the four- and five-factor level negative expectancy factors displayed small-to-medium, positive relations to Agreeableness, as well as small, positive relations to Conscientiousness and Openness. These negative expectancy factors bore null-to-small relations to proactive aggression, reactive aggression, and antisocial behavior that were directionally inconsistent.

In terms of FFM facet relations, the positive expectancy factors bore large-to-very large, negative relations with *Morality* (A) and *Cooperation* (A), and medium-to-large, negative relations to *Altruism* (A), *Modesty* (A), *Sympathy* (A), *Dutifulness* (C), *Cautiousness* (C), *Artistic Interests* (O), and *Intellect* (O). Alternately, the positive expectancy factors bore medium-to-large, positive relations to *Anger* (N). The relations between the negative expectancy factors and FFM facets were smaller, but were in the expected directions. The negative expectancy factors bore medium-to-large, positive relations with *Altruism* (A) and *Sympathy* (A), as well as small-to-medium, positive relations with *Morality* (A), *Dutifulness* (C), *Imagination* (O), *Emotions* (O), *Artistic Interests* (O), *Anxiety* (N), and *Self-Consciousness* (N).

Lastly, we conducted intraclass correlations between the factors in order to quantify the degree of absolute similarity of these factors in relations to these 33 criterion variables (i.e.,

proactive aggression, reactive aggression, antisocial behavior, 30 FFM facets) external criteria. The "positive intrapersonal" and "positive interpersonal" factors bore highly similar criteria relations to one another at the four- and five-factor levels (i.e., $r_{ICCs} \geq .85$). The lower-level negative expectancy factors also bore similar criteria relations with one another (i.e., r_{ICCs} range = .31 to >.99). The "damage to self-image/reputation" factor also bore highly dissimilar criteria relations to the four- and five-level lower-order positive expectancy factors (i.e., $r_{ICCs} \leq -.84$).

Study 4 Discussion

In Study 4, we advanced the Aggression Expectancies by administering a refined pool of items to a sample of community adults, further culling redundant items, exploring the factor structure of all remaining items, and examining indices of convergent and discriminant validity (i.e., relations to self-report aggression, antisocial behavior, FFM domains and facets). As in Study 3, the factor structure of these items unfolded such that the expected positive and negative intrapersonal and interpersonal factors emerged at the four-factor level, with a "damage to self-image/reputation" factor emerging at the five-factor level. Thus, with this reduced set of items, these data provide evidence for an internal structure with factors that capture the specific themes identified in Study 1 and 2.

The positive expectancy factors exhibited strong convergent validity, such that individuals who expect positive consequences of behaving aggressively report higher levels of proactive and reactive aggression. In terms of personality domains and facets, individuals who hold positive aggression expectancies tend to be low in Agreeableness and its facets, low in *Dutifulness* (C) and *Orderliness* (C), as well as high in *Anger* (N). This pattern is consistent with existing meta-analytic work examining the relations between FFM facets and self-report and laboratory aggression (Hyatt et al., 2019, 2020; Jones et al., 2011), which suggests that (low)

Agreeableness and its facets tend to be the largest predictors of aggression, but that facets like *Anger* from other domains also bear medium-to-large relations with aggression.

Third, the convergent relations for the negative expectancy factors were somewhat smaller than those for the positive expectancy factors, but were still generally in the expected directions. Although the negative expectancy factors exhibited generally null relations to aggression, they generally exhibited positive relations to facets of Agreeableness, especially *Altruism* and *Sympathy*, as well as Neuroticism, especially *Anxiety* and *Self-Consciousness*. In other words, individuals with relatively high standing on these facets tend to expect that behaving aggressively will result in more negative consequences for themselves and the target of their aggression.

An ongoing issue at this stage was adjudicating between the various levels of factor solutions, especially between the four- and five-factor level solutions. We observed evidence that a hierarchical structure was sensible, such that the higher-order "general positive" and "general negative" factors break into more specific lower-order factors (akin to domains and facets of the FFM), but making a formal decision about the optimal lower-order structure proved challenging. Throughout this process, we attempted to strike a balance between using empirical indicators (e.g., model fit) and conceptual considerations (e.g., parsimony, content breadth) in mind. On one hand, the four-factor structure is symmetrical and intuitive, as the two higher-order factors each break into two parallel "intrapersonal" and "interpersonal" lower-order factors. The MAP test also supported a four-factor structure, but the five-factor solution displayed superior fit indices. However, the largest consideration was a blend of conceptual and empirical concerns: namely, it is unclear to what extent the "damage to self-image/reputation" is distinct from the "positive intrapersonal" factor.

Empirically, “damage to self-image/reputation” displayed very large bivariate relations to “positive intrapersonal” factors at four- and five-factor level at r = -.66 and -.47, respectively. Although these values alone are not necessarily so large to suggest redundancy, the intraclass correlations suggest that these factors have extremely similar criteria relations (but in the opposite direction): the “damage to self-image/reputation” exhibited intraclass correlations with the “positive intrapersonal” factors at four- and five-factor level at r_{ICC} = -.93 and -.91, respectively. In fairness, at the four- and five-factor levels, the two positive expectancy factors also bore highly similar bivariate relations (rs = .47 and .50, respectively) and criteria relations (r_{ICC} = .88 and .91, respectively). However, we viewed this overlap between the positive expectancy factors as less problematic, given that they are subsumed under the same “general positive” factor, while the “damage to self-image/reputation” appears to be a component of a different domain.

Conceptually, it is difficult to intuit how the “damage to self-image/reputation” factor is distinct from essentially the opposite of the content contained in the “positive intrapersonal” and “positive interpersonal” factors. For example, it is unclear what additional content that the “damage to self-image/reputation” item “I would be disappointed in myself for doing so” is capturing that is not captured (inversely) by the “positive intrapersonal” item “It will make me feel good about myself”, apart from a pedantic distinction about the degree to which feeling “disappointed in myself” and “good about myself” are opposites. Moreover, the item “I would be disappointed in myself for doing so”, which was viewed as a face valid indicator of the conceptual construct “damage to self-image”, loaded onto the “positive intrapersonal” factor >.40 at the four-factor level. As a final point, as we contemplated additional items that could more focally, differentially capture either positive intrapersonal aggression expectancies or

expectancies about disappointing oneself, it became increasingly clear that these were largely two sides of the same coin, so to speak.

In sum, a major initiative of subsequent data efforts was to continue to investigate the optimal structure for the aggression expectancy measure. Specifically, the primary concern was the extent to which a four- vs. five-factor solution was most appropriate for our data, which would ultimately inform the structure of the final version of the measure.

Table 3

Relations Between Expectancy Factors and External Criteria in Study 4

	PA	RA	ASB	N	E	O	A	C
Positive intrapersonal (4.1)	**.43**	**.20**	.19	.04	.09	**-.26**	**-.50**	**-.24**
Positive interpersonal (4.2)	**.28**	**.36**	.14	.19	.05	-.13	**-.42**	**-.22**
Harm to victim (4.3)	-.02	.01	.03	.07	.10	.09	**.24**	.16
Harm to self (4.4)	-.05	.07	-.03	.13	.00	.10	.11	.04
Positive intrapersonal (5.1)	**.42**	**.24**	.19	.06	.09	**-.25**	**-.46**	**-.21**
Positive interpersonal (5.2)	**.27**	**.36**	.14	.19	.04	-.12	**-.41**	**-.21**
Harm to victim (5.3)	.03	.02	.06	.07	.11	.06	.15	.11
Neg. external consequences (5.4)	-.05	.07	-.03	.13	.00	.09	.10	.03
Damage to self-image/reputation (5.5)	**-.24**	-.07	-.09	.01	.02	.18	**.45**	**.25**

Note: PA = proactive aggression; RA = reactive aggression; ASB = antisocial behavior; N = neuroticism; E = extraversion; O = openness to experience; A = agreeableness; C = conscientiousness; relations ≥ |.20| are **bolded**.

CHAPTER 6

STUDY 5

Study 5 Overview

The goal of Study 5 was to refine the aggression expectancy measure by administering a reduced and amended set of items to another sample of community adults, and examining the factor structure exhibited by the reduced set of items. The analytic strategy was similar to Study 4, except that the primary focus of Study 5 was to adjudicate between different levels of the factor hierarchy, with the primary interest in comparing the two-, four-, and five-factor level solutions. Thus, criterion variables were not assessed in this study, anticipating this would continue to be an essential and much-needed part of future efforts to establish the validity of this measure.

Study 5 Methods

Participants

Study 5 participants were recruited from Amazon's Mechanical Turk. Unlike Studies 3 and 4, there was no screening procedure, but rather we implemented a four-part validity check in order to screen for valid responding, including duration, open-ended response coding (i.e., "What are your two favorite websites and why do you like visiting them? Please provide one full sentence for each website."), embedded attention checks (i.e., "Select 'Somewhat Unlikely' for this answer"), and geolocation checking. All participants were informed that if they did pass each of these validity checks, that their data would be discarded and they would not be compensated. The survey was administered to a total of 525 participants. Data from N = 73 participants were

discarded based on unrealistically rapid responding (i.e., <100 seconds). The first and senior author screened each response to the open-ended response, and deemed N = 125 of these responses invalid (e.g., unintelligible, not written in complete sentences despite clear directions). An additional N = 19 did not pass both embedded attention checks, and no remaining participants had an ineligible geolocation. Thus, the final sample for Study 5 was N = 308 (54.9% male; mean [SD] age = 36.8 years [10.8]; 81.5% White or Caucasian, 5.8% Black or African-American, 5.2% Asian, 3.9% reporting more than one racial identity, and 3.6% Hispanic or Latino).

Aggression Expectancy Measure Item Selection

Items from all five aggression expectancy factors from Study 4 were included in the interest of comparing how the items from the "damage to self-image/reputation" performed in relation to the "positive intrapersonal" items. The top four highest-loading items from each of the five factors observed in Study 4 were included for an initial pool of 20 items. In addition to these 20 items, an additional 11 items were included, of which 10 were new. These new items were included in the interest of not creating overly narrow scales, or falling prey to the "attenuation paradox" (Loevinger, 1957; Watson & Clark, 2019). Specifically, Watson and Clark (2019) note that "if the top-loading items are highly redundant with one another, including all will increase internal consistency estimates but also may create an overly narrow scale that does not represent the construct optimally" (pg. 7).

Of these 10 new items, seven were generated for the factor "harm to victim." After examining the item content of the four highest-loading items, the content was judged to be skewed toward more extreme, extensive harm than may be typical of an aggressive encounter, especially a verbal one. Specifically, the items "The other person will really suffer," "I may

seriously hurt the other person," and "The other person will never be the same afterwards" were face valid in terms of capturing the concept "harm to victim," but seem to be restricted in terms of the range of severity of harm/suffering that may be experienced by a victim. Thus, an additional seven items were written with the aim of capturing a broader range of this hypothetical harm/suffering severity spectrum, including items such as "I will hurt the other person's feelings," "It will upset the other person," and "The other person will feel embarrassed." One additional new item was generated for the "positive intrapersonal" factor: "It will feel good." This new item was viewed favorably compared to the conceptually similar, but lengthier and narrower item "It will make me feel good about myself."

The final two new items added were for the "harm to self" factor: "I will get hurt," and "I will experience formal negative consequences (e.g., get arrested, lose my job)." These new items were added in the interest in increasing the content coverage beyond what was present in the four highest-loading items. The first item was added in order to the capture an expectancy about immediate harm to oneself, which was in contrast to the four highest-loading items which were exclusively limited to the potential for future retaliation (i.e., "It will only be a matter of time before they get revenge on me"; "I'd be vulnerable to retaliation"; "The other person's friends or family will try to get back at me"; "I will have to watch my back in the future"). The second item was also added in order to increase content coverage by expanding the scope of this factor beyond content related to person-to-person level harm (either immediately or in the future as retaliation), but to also include the institution-to-person impairment (e.g., losing one's job, legal consequences) that one may experience as a result of behaving aggressively, given that virtually all individuals live under a variety of societal dictates that condemn and punish aggressive behavior. Finally, with this same consideration in mind, the item "I will get into trouble" from

previous iterations of the measure was also retained, since this was viewed as a face valid item that was ambiguous enough to capture a range of ways in which one could be negatively intrapersonally impacted by behaving aggressively.

Aggression Expectancy Questionnaire

The third iteration of the Aggression Expectancy Questionnaire comprised 31 items. Participants were presented with the following prompt: "*People often behave in certain ways because they expect a certain consequence to occur. For example, you may eat a snack because you expect the consequence will be that you feel less hungry afterwards. Below is a series of statements about* ***possible consequences of being aggressive****. Please respond on a 1-5 scale to indicate how likely you think each consequence is for you.*" Participants were presented with a Likert scale with the following labels: 1 = Very Unlikely (i.e., it is very unlikely that this will happen if I behave aggressively), 2 = Somewhat Unlikely, 3 = Neither Likely nor Unlikely, 4 = Somewhat Likely, and 5 = Very Likely (i.e., it is very likely that this will happen if I behave aggressively). We included the running statement "*If I am aggressive toward other people, then I expect that…*" above each cluster of 10 items.

Analyses

A series of analyses was conducted with the ultimate goals of culling items in order to make a briefer measure (i.e., 16 to 20 items, or four items per subscale at the four- or five-factor level) and exploring the factor structure of these items. First, items were eliminated that did not receive responses across the full range of response options (i.e., did not have a range of 1-5). Second, bivariate relations were computed between all items, and pairs of items correlated at $r \geq .70$ were identified and deemed redundant. In these cases, the items that exhibited the largest number of these high correlations (i.e., the most redundancies) were eliminated. To examine the

factor structure of the reduced set of items, a series of exploratory factor analyses (EFA) using principal axis factoring and direct oblimin rotation were conducted. Parallel Analyses and Velicer's MAP test were also used in order to inform the interpretation of the various factor solutions. Finally, in order to adjudicate between various solutions, EFA with promax rotation and maximum likelihood fitting procedure were conducted.

Study 5 Results

Initial Item Refinement

All items exhibited the full range of responses. There was only one pair of items that correlated at $r > .70$. We elected to eliminate the item "It will make me feel good about myself" rather than the item "Later on, I will be happy that I did so" because the former exhibited more explicit content overlap with another item from the "positive intrapersonal" set of items (i.e., "It will feel good") than the latter.

Exploratory Factor Analyses of Remaining Items

An iterative sequence of analyses were conducted to examine four- and five-factor solutions of the remaining 30 items. At the four-factor level, the "positive interpersonal," "harm to self," and "harm to victim" factors emerged as expected. Unexpectedly, many of the new items generated for the factor "harm to victim" loaded onto a loose "positive intrapersonal" factor that also emerged at this stage. This factor represented a blend of (positively loading) positive intrapersonal items, (negatively loading) damage to self-image/reputation items, and (negatively loading) new harm to victim items, and thus was difficult to interpret.

At the five-factor level, the "positive interpersonal" and "harm to self" factors remained extremely similar, and a clearer, more focal "positive intrapersonal" factor emerged at this level. The "harm to victim" factor emerged, but the fifth factor to emerge was also largely a "harm to

victim" factor. The only apparent difference between these two "harm to victim" factors was in severity of harm, with the former representing more irreparable harm and the latter appearing to represent somewhat more transient, emotion-specific harm.

Secondary Item Refinement

After this round of EFA, the primary interpretative issue was the peculiar break-down of the "harm to victim" items that emerged, as the item loadings at the four- and five-factor levels were puzzling. At the four-factor level, the "harm to victim" items differentially loaded onto the "positive intrapersonal" and "harm to victim" factors, and at the five-factor level, these items broke into two separate factors that appeared to only different in degree, rather than in kind of harm. Notably, most (but not all) of the items that displayed problematic loadings were newly generated for this current iteration of the measure. Given that these new items largely did not perform as expected, seven of the eleven of the items generated for the "harm to victim" category were eliminated, as the ultimate goal was to include four items per subscale. A combination of criteria was used to eliminate items. First, the three items that did not exhibit a loading >.40 onto the "harm to victim" factors at either the four-factor or five-factor level were eliminated. Next, the two items that loaded >.40 onto the "positive intrapersonal" factor at the four-factor level were eliminated. Out of these six remaining items, we elected to retain four items that exhibited factor loadings >.50 onto the "harm to victim" at the four-factor level. Importantly, item loading was not the sole criterion at this final stage – we also considered item clarity and breadth of content for this factor. For example, we selected one of the relatively lower loading items ("I will harm the other person's reputation") rather than the item "The other person will never be the same afterwards", given interest in also including items that capture relatively lower levels of harm.

Exploratory Factor analysis of Remaining Items

Following this reduction in items, the EFA were repeated on these 23 items to examine how this reduced number of "harm to victim" items would perform in concert with the other items. At the four-factor level, the results were largely consistent with the four factors previously identified at this level, but with more interpretative clarity. A clear "positive intrapersonal" factor emerged, such that all four positive intrapersonal items and all four damage to self-image/reputation items loaded onto this factor >.40. Clear "positive interpersonal," "harm to self," and "harm to victim" factors also emerged, such that all of the items included for each of these factors displayed loadings >.40.

Finally, at the five-factor level, the "positive interpersonal," "harm to self," and "harm to victim" factors were largely unchanged. The "positive intrapersonal" factor fractured into two sub-factors. The items that loaded onto the first "positive intrapersonal" factor >.40 included a blend of damage to self-image/reputation items, a positive intrapersonal item, and two harm to self items. The second "positive intrapersonal" factor, which could be deemed a "justice" only had two items with loadings >.40, both of which were justice-oriented.

Final Item Refinement

After this round of analyses, we formally decided to pursue a four-factor structure, consisting of "positive intrapersonal," "positive interpersonal," "harm to self," and "harm to victim". This decision was based on the interpretative difficulties that emerged at the five-factor level, as well as the conceptual overlap between the content of the "positive intrapersonal" and "damage to self-image/reputation" factors.

The final item refinement process involved selecting four items for each of these four factors that would ultimately comprise the final aggression expectancy measure. This was

straightforward for the "positive intrapersonal," "positive interpersonal," and "harm to victim" factors, as there were only four items remaining that were included in the current version of the measure for the positive expectancy factors, and four items were chosen for the "harm to victim" factor at the previous step of these analyses. There were seven remaining items for the "harm to self" factor at this stage. One item ("I will experience negative formal consequences [e.g., get arrested, lose my job]") that did not load onto this factor >.40 at the four-factor stage was eliminated. The two items that exhibited the highest loadings onto the "harm to self" factor at the four-factor level (i.e., "I'd be vulnerable to retaliation," "I will have to watch my back in the future") were selected. Given that the content of both of these items are more retaliation/revenge-oriented, we elected to include the two other items generated for this factor (i.e., "I will get hurt," "I will get into trouble") that capture a broader range of expectancies about the ways in which one may be harmed after behaving aggressively.

Exploratory Factor Analysis of Final Item Set

A last round of EFA was conducted on this finalized set of 16 items to ensure that the expected structure emerged, and item loadings are presented in Table 4. At the four-factor level (cumulative variance = 62.3%), we observed the expected structure, such that each of the four expected factors emerged and all items designated for those factors displayed loadings >.45 (and in 15/16 cases, >.50). Finally, we examined the fit statistics of these items with EFA with promax rotation and a maximum likelihood fitting procedure. The results suggested that the four-factor level solution (RMSR = .03, TLI = .916, RMSEA [90% C.I.]= .063 [.047 to .075], BIC = -222.30) displayed relatively good model fit. Of note, we also examined a five-factor structure. In this solution, the "positive intrapersonal" factor split into "positive feelings" and "justice" factors that were correlated with the four-factor level "positive intrapersonal" factor at r = .80 and .79,

respectively. This solution exhibited good model fit (RMSR = .02, TLI = .945, RMSEA [90% C.I] = .051 [.031 to .066], BIC = -199.24) but was inferior to the four-factor solution in terms of the BIC.

Factor Interrelations

Factor interrelations can also be found in Table 4. The "positive intrapersonal" and "positive interpersonal factors displayed a very large, positive relation (r = .47), as did the "harm to self" and "harm to victim" factors (r = .62). Interestingly, the "positive intrapersonal" factor exhibited medium, negative relations with the "harm to self" and "harm to other" factors (rs = -.30 and -.32, respectively), while the "positive interpersonal" factor exhibited small, positive relations with these lower-order negative expectancy factors (rs = .11 and .16, respectively).

Study 5 Discussion

The primary advancement in Study 5 was the establishment of a four-factor structure for the Aggression Expectancy Questionnaire. This was due to the elimination of the "damage to self-image/reputation" factor for conceptual reasons (i.e., difficult to interpret how this content is distinct from the inverse of "positive intrapersonal" items), as well as empirical concerns that began in Study 4. Although the "damage to self-image/reputation" items broke into a more coherent factor at the five-factor level in Study 4, this factor displayed very large bivariate (r = -.66) and intraclass (r_{ICC} = -.93) relations to the "positive intrapersonal" factor. Similarly, in the current study, the items from each of these factors loaded onto a "positive intrapersonal" factor at the four-factor level, and these items did not break into clear subscales at the five-factor level. Additionally, fit indices of the various solutions did not suggest the fit was notably better at the five-factor level than the four-factor level. Thus, we elected to move forward with a 16-item

version of the measure that would comprise four subscales (i.e., *Positive Intrapersonal*, *Positive Interpersonal*, *Harm to Self*, and *Harm to Victim)* with four items each.

Table 4

Study 5 Item Loadings and Factor Interrelations

Positive Intrapersonal Items	*4.1*	*4.2*	*4.3*	*4.4*
1. Later on, I will be happy that I did so.	**.57**	.07	-.18	.07
2. I will be doing it because it's the right thing to do.	**.86**	-.06	.07	-.13
3. I will feel that I acted in the name of justice.	**.69**	.08	.07	-.05
4. It will feel good.	**.46**	.20	-.16	.05
Positive Interpersonal Items	*4.1*	*4.2*	*4.3*	*4.4*
5. I will appear more dominant to others.	.03	**.72**	.05	-.09
6. Others will learn not to mess with me.	.09	**.77**	.04	<-.01
7. Others will see what I'm capable of.	.07	**.67**	-.02	<-.01
8. They will be afraid of me in the future.	-.03	**.51**	.03	.34
Harm to Self Items	*4.1*	*4.2*	*4.3*	*4.4*
9. I will be vulnerable to retaliation.	-.05	.08	**.81**	-.10
10. I will get hurt.	.01	-.10	**.57**	.18
11. I will have to watch my back in the future.	.01	.17	**.76**	.05
12. I will get into trouble.	-.06	-.12	**.54**	.19
Harm to Victim Items	*4.1*	*4.2*	*4.3*	*4.4*
13. The other person will really suffer.	-.12	.10	.04	**.58**
14. I may seriously hurt the other person.	-.13	.06	.12	**.60**
15. I will negatively affect their quality of life.	-.13	-.01	-.02	**.68**
16. I will harm the other person's reputation.	.18	-.06	.11	**.51**
Factor Interrelations	*4.1*	*4.2*	*4.3*	*4.4*
4.1 Positive Intrapersonal	-			
4.2 Positive Interpersonal	**.48**	-		
4.3 Harm to Self	-.29	.09	-	
4.4 Harm to Victim	-.28	.20	**.61**	**-**

Note: factor loadings >.40 and factor interrelations >.40 are **bolded.**

CHAPTER 7

STUDY 6

Study 6 Overview

The goal of Study 6 was to examine the proposed four-factor structure of the Aggression Expectancy Questionnaire in an independent sample using a confirmatory factor analytic approach. We conducted a series of analyses to assess the fit of a four-factor structure using maximum likelihood estimation method. An additional goal was to continue to examine the nomological network of these aggression expectancies by investigating the criteria relations with the five domains and 30 facets of the FFM. These criteria relations also permitted investigation of the absolute similarity (i.e., intraclass correlations) of these factors.

Study 6 Methods

Participants

Study 6 participants were recruited from Amazon's Mechanical Turk.
A multi-component screening procedure was used in order to maximize data validity. An iterative data collection process was implemented, and ratings of responses to an open-ended prompt (i.e., "what is your favorite weekend leisure activity, and what do you like most about it?") were the first method of determining valid responses. The first and senior author read each response and coded them as either invalid or valid; data from participants whose responses were deemed valid by both authors were used. Out of a total N = 589, N = 395 participants provided valid responses to this open-ended prompt. We used a series of three attention checks (e.g., "Select the middle option for this response") to screen for inattentive responding, and we

discarded data from participants (N = 33) that failed one or more of these checks. Finally, we discarded data from participants (N = 4) who completed the survey in an unrealistically fast response time (132 seconds). This resulted in a final sample of N = 358 valid responders (53.8% male; mean [SD] age = 37.0 years [11.4]; 70.8% White or Caucasian, 11.1% Black or African-American, 7.8% Asian, 4.5% Hispanic or Latino, 4.5% reporting more than one racial identity, 1.1% American Indian, less than 1% Native Hawaiian or Pacific Islander).

Aggression Expectancy Questionnaire

The final iteration of the aggression expectancy measure comprised the 16 items identified in Study 5. As in each previous study, participants were presented with the following prompt: "*People often behave in certain ways because they expect a certain consequence to occur. For example, you may eat a snack because you expect the consequence will be that you feel less hungry afterwards. Below is a series of statements about* ***possible consequences of being aggressive****. Please respond on a 1-5 scale to indicate how likely you think each consequence is for you.*" Participants were presented with a Likert scale with the following labels: 1 = Very Unlikely (i.e., it is very unlikely that this will happen if I behave aggressively), 2 = Somewhat Unlikely, 3 = Neither Likely nor Unlikely, 4 = Somewhat Likely, and 5 = Very Likely (i.e., it is very likely that this will happen if I behave aggressively). We included the running statement "*If I am aggressive toward other people, then I expect that…*" above each cluster of 5 items.

We calculated scores for each of the four scales by averaging the four items that comprise each subscale. These scales demonstrated good internal consistency estimates: *Positive Intrapersonal* $\alpha = .81$, mean inter-item correlation (mIIC) = .52; *Positive Interpersonal* $\alpha = .79$,

mean inter-item correlation (mIIC) = .48; *Harm to Self* α = .80, mean inter-item correlation (mIIC) =.50; and *Harm to Victim* α = .77, mean inter-item correlation (mIIC) = 45.

Five Factor Model Rating Form

The Five Factor Model Rating Form (FFM-RF) is a 30-item self-report measure that uses one item to assess each of the 30 facets of the FFM. We created domain composites by averaging the scores for facets that comprise each domain: Neuroticism (α = .82), Extraversion (α = .82), Openness (α = .72), Agreeableness (α = .78), and Conscientiousness (α = .86).

Analyses

First, using the *lavaan* package in R, we conducted confirmatory factor analyses using maximum likelihood estimation procedure to examine the model fit of these items at the four-factor level. The latent factors were not constrained to orthogonality. Next, interrelations between the aggression expectancy scales were computed, as well as with the FFM domains and facets. Finally, intraclass correlations were calculated between each aggression expectancy scale using the 30 FFM facets as criteria.

Study 6 Results

Confirmatory Factor Analysis

Examination of fit statistics (CFI = .899, TLI = .876, AIC = 16602.72, BIC = 16750.18, RMSEA = .079, SRMR = .079) suggested that the four-factor model displayed good fit by most general standards (Brown, 2014; Sellbom & Tellegen, 2019).

Factor Interrelations

Relations between factors as well as factor relations with the FFM domains can be found in Table 5. As expected, the positive and negative subscales bore very large, positive relations to one another (*Positive Intrapersonal – Positive Interpersonal r* = .40; *Harm to Self – Harm to*

Victim r = .58). The *Positive Intrapersonal* factor bore medium, negative relations to the *Harm to Self* and *Harm to Victim* subscales (*r*s = -.33), and the *Positive Interpersonal* factor bore small, positive relations with these scales (*r*s = .11 and .10, respectively).

Relations to External Criteria

The *Positive Intrapersonal* and *Positive Interpersonal* scales bore a medium-to-large, negative relations to Agreeableness (*r* = -.32, -.18, respectively), and a similar pattern was observed for the positive aggression expectancy scales and Conscientiousness and Openness (i.e., small-to-medium, negative relations), although these relations were smaller in magnitude. The *Harm to Self* and *Harm to Victim* scales bore small-to-medium, positive relation to Agreeableness (*r* = .14, 17, respectively).

Positive Intrapersonal bore medium-to-large, negative relations to *Straightforwardness* (A), *Compliance* (A), *Tendermindedness* (A), *Dutifulness* (C), *Deliberation* (C), and *Warmth* (E), as well as medium, positive relations to *Anger* (N) and *Assertiveness* (E). Overall, *Positive Interpersonal* exhibited a similar pattern of relations as *Positive* Intrapersonal, but these relations were smaller in each instance. *Harm to Self* demonstrated medium, negative relations to *Assertiveness* (E) and *Activity-Level* (E), as well as medium, positive relations to *Tendermindedness* (A) and *Self-Consciousness* (N). *Harm to Victim* also displayed a medium, positive relation to *Tendermindedness* (A).

Intraclass Correlations

The *Positive Intrapersonal* and *Positive Interpersonal* scales bore a relatively high degree of absolutely similarity (r_{ICC} = .62). The *Harm to Self* and *Harm to Victim* scales exhibited smaller absolute similarity (r_{ICC} = .34). Finally, unexpectedly, the *Positive Intrapersonal* and *Harm to Victim* subscales also exhibited a high degree of absolute dissimilarity (r_{ICC} = -.74).

Table 5

Study 6 Factor Interrelations, Intraclass correlations, and FFM Domain Relations

	4.1	4.2	4.3	4.4
4.1 *Positive Intrapersonal*	-	.62	-.43	-.74
4.2 *Positive Interpersonal*	**.40**	-	-.01	-.61
4.3 *Harm to Self*	**-.33**	.11	-	.34
4.4 *Harm to Victim*	**-.33**	.10	**.57**	-
Neuroticism	.10	.11	.14	-.02
Extraversion	-.04	-.07	-.19	.06
Openness	-.16	-.12	-.11	.02
Agreeableness	**-.32**	-.18	.14	.17
Conscientiousness	**-.26**	-.10	.01	.08

Note: values below the diagonal are bivariate relations, values above the diagonal are intraclass correlations; N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness; bivariate relations > .20 are **bolded**.

CHAPTER 8

DISCUSSION

Summary

Given the well-documented societal costs of aggression, it is imperative to understand its precipitants, such as the expectancies that individuals hold about the relatively likelihood of what may result if one is aggressive. The goal of the current manuscript was to describe the development a measure of general aggression expectancies suitable for use in adult populations. Across a series of six studies, we described the iterative scale construction process, which yielded a 16-item measure that we hereafter refer to as the Aggression Expectancy Questionnaire (AEQ).

Indices of Validity

At this stage, it is beneficial to do a brief review of the evidence for/against the construct validity of the AEQ in its final form. First, in the interest of content validity, it is worth revisiting the seven positive and seven negative aggression expectancy themes (see Table 1). By and large, the final version of the AEQ captures the majority of these themes. Table 6 lists the final AEQ items, as well as appended themes that we believe are captured by each item. Based on this informal coding, all seven of the positive expectancy themes are captured, and six of the seven negative expectancy themes are captured, with the exception being *damage to relationships* (e.g., "people will think poorly of me"). This apparent gap in content coverage is due, in part, to the decision in Study 5 to eliminate the items included for the factor "damage to self-image/reputation", which comprised items that were face valid for this theme (e.g., "People close

to me [i.e., my friends and family] will be disappointed in me for doing so"). The data from Studies 3-5 suggest these items tend to load strongly onto the "positive intrapersonal" factor. Moreover, absolute similarity analyses in Study 4 suggest that the "damage to self-image/reputation" factor bears remarkably similar relations to external criteria as the "positive intrapersonal" factor. These findings suggest that while these items appear capture a distinct type of social consideration, the data suggests that empirically, these concerns as very highly related. Furthermore, we argue that this inverse of this theme may be captured to some degree with the item "Later on, I will be happy that I did so" (reverse-scored). As such, this underlying latent disposition toward anticipating damage to one's relationships is largely captured by the current items.

In terms of convergent and discriminant validity, in Studies 3, 4, and 6, different iterations of the *Positive Intrapersonal* and *Positive Interpersonal* factors showed expected relations with self-reported aggression as well as (low) Agreeableness and (low) Conscientiousness. Furthermore, at the facet level, these positive expectancy factors bore the largest relations to the facets previously identified as the most important predictors of aggression (Hyatt, Chester, Zeichner, & Miller, 2020; Jones et al., 2011), including (low) *Sympathy* (A), (low) *Morality* (A), and *Anger* (N). In other words, individuals who tend to be aggressive and relatively antagonistic expect more positive intrapersonal and interpersonal consequences following an aggressive act. On the other hand, the negative expectancy factors generally exhibited lower convergence with the criteria used herein. Although *Harm to Self* and *Harm to Victim* subscales exhibited positive relations to Agreeableness, suggesting that more agreeable individuals view negative consequences for oneself and others as more likely following an aggressive act, these effect sizes were small in magnitude.

One potential consideration is the role of efficacy, which is also a major component of Social Learning Theory and may represent a confound or unexamined moderator of these relations. For example, consider the *Harm to Victim* item "The other person will really suffer" – endorsement of this expectancy is likely related to one's own sense of capability to inflict harm on the victim, especially in the case of physical aggression. Thus, individuals are likely considering several factors when interpreting and responding to this item, including both their own empathic sense of how much an aggressive behavior will subjectively impact the victim, as well as their own ability to bring about such harm. Although we considered this issue from the outset of this measure development project, we found it difficult to write items that separated these constructs that were not long-winded and meaningfully changed the prompt (e.g., "if I aggress against another person and I am effective at harming them, then the other person will really suffer") or counterfactual (e.g., "if I was able to aggress against another person effectively, then the other person would really suffer"). Given that efficacy is clearly an important variable to consider for certain forms of aggression, we viewed this as an important area to clarify with future research.

Theoretical Considerations

The current results can be viewed from several theoretical perspectives. First, the finding that positive aggression expectancies are related to elevated aggression is consistent with the tenets of Social Learning Theory that types of expectancies are important precipitants of social behavior. Moreover, the AEQ reflects the response vs. outcome distinction that is prevalent in expectancy research on other behaviors (e.g., substance use), and the current results suggest that both types of aggression expectancies, especially positive ones, may predict aggression. Analogous to the substance use literature, the current results suggest that positive expectancies,

rather than negative expectancies, may be more potent predictors of future behavior (Treloar et al., 2015). This is sensible given that the associated rewards (e.g., feeling good) are more proximal than the associated punishments (e.g., getting in trouble), although the potential for harm in an aggressive interaction may also be proximal (e.g., getting hurt). Future research may illuminate similarities for how different types of expectancies inform these various behaviors.

Second, the current results suggest that these expectancies may be one avenue by which antagonism and related complex personality profiles (e.g., psychopathy, narcissism) are related to aggressive behavior. In other words, although the links between these traits and aggression are well-established, the intermediary psychological links between these distal temperamental tendencies and the behavioral manifestation of this harmful outcome are largely unknown, although several have been proposed for both antagonism (e.g., situation selection, Bresin & Robinson – 2015; difficulty disengaging with antisocial stimuli – Wilkowski, Robinson, & Meier, 2006; c.f. Vize and Lynam, 2020) and complex personality profiles (e.g., attenuated startle response in components of psychopathy – Benning, Patrick, & Iacono, 2005; narcissistic rage in response to ego-threat – Krizan & Johar, 2015). Identification of such mechanistic links is a pressing need for personality science, and individual differences in positive and negative expectancies about the likely consequences of aggression may be one such socio-cognitive mechanism (Fleeson & Jayawickreme, 2015).

Future Directions

Despite some evidence of validity, the primary future direction at this stage is to continue to investigate the psychometric properties of this measure. Although there are, of course, many studies that could be conducted to this effect, we highlight several that are particularly imperative in terms of creating a solid empirical base around the AEQ. First, all of the studies

conducted with the AEQ herein used predominantly white, American adults recruited from Amazon's MTurk. While we believe there is a value in this data collection platform (if the appropriate validity measures are in place), there is a pressing need to examine the AEQ in other populations with more demographic diversity (e.g., race, age, socioeconomic status, country of origin). Additionally, given the content of the AEQ, studies on specific populations that have significant experience with behaving aggressively (e.g., incarcerated violent offenders; athletes trained in mixed martial arts) would be enlightening, as one might posit that an individual with substantial experience with aggression (e.g., physical fights) might tend to expect more positive (and less negative) consequences of being aggressive.

Second, although we have proposed a model in which individual differences in aggression expectancies predict aggression, the cross-sectional nature of the current data prohibits our ability of formally testing this directionality. In fact, consistent with the notion that expectancies are learned components of social cognition that begin early in life, it is likely that there is bidirectionality to consider (i.e., one's experiences with aggression shape one's temperament and vice versa). Thus, it would informative to track changes in aggression expectancies over the course of development, and it may be especially beneficial to examine how direct and indirect experiences with aggressive behavior (e.g., perpetration, victimhood) impact these expectancies. Analogous work has been conducted that suggests that expectancies about alcohol use predict problematic drinking behavior one year later (Christiansen, Smith, Roehling, & Goldman, 1989).

Third, it is critical to examine the various contextual factors that may lead to differential consideration of these aggression expectancies. For example, an acute affective state such as anger may lead one to place relatively more subjective value on positive rather than negative

expectancies. Indeed, although an expectancy model suggests that aggression is the outcome of a "cold" decision making process, it is much better described as a "hot" process that often occurs in an acute emotional state such as anger or fear (see Crick and Dodge, 1994 for a similar criticism about the Social Information Processing model). Future research on the ways in which aspects of a given situation influence aggression expectancies and their activation in a given social interaction is highly encouraged.

Finally, this work also has clinical relevance. If aggression expectancies are meaningful predictors of aggression, and if these expectancies can be altered through therapeutic intervention (e.g., psychoeducation, cognitive restructuring), then this could serve as a potentially useful intervention module for clinicians to use with chronically aggressive or hostile individuals. As before, there is analogous research on substance use expectancies (e.g., Cruz & Dunn, 2002; Fulton, Krank, & Stewart, 2012; Trudeau, Spoth, Lillehoj, Redmond, & Wickrama, 2003) that generally suggests that expectancies can be altered through intervention, concordant with changes in behavioral indices like initiation of substance use and escalation over time (Treloar et al., 2015). In addition to providing a framework for exploring a client's beliefs about aggression, the AEQ could ultimately prove useful as an assessment tool for tracking treatment progress. This clinically-oriented work could be complemented by research examining if exposure to certain types of stimuli (e.g., a victim's account of being assaulted) can alter one's aggression expectancies. Ultimately, we believe that even small effects may be important to consider, given that preventing one act of aggression may hold far-reaching benefits for both the potential aggressor and victim. In closing, we hope to see future research on the AEQ prove useful to psychologists working in a variety of basic and applied domains. To facilitate this

utility, the full version of this measure, as well as instructions for scoring the AEQ scales in R and SPSS is freely available at https://osf.io/fuz6h/.

Table 6

Standardized CFA Loadings in Study 6 and Relation to Study 2 Themes

Aggression Expectancy Item	λ	SE	Relevant Study 2 Themes
Positive Intrapersonal Items			
1. Later on, I will be happy that I did so.	.64	.06	*Positive feelings;* *Emotion/tension release*
2. I will be doing it because it's the right thing to do.	.79	.07	*Justice*; *Achieving a goal*
3. I will feel that I acted in the name of justice.	.80	.06	*Justice;* *Achieving a goal*
4. It will feel good.	.63	.06	*Positive feelings;* *Emotion/tension release*
Positive Interpersonal Items			
5. I will appear more dominant to others.	.63	.07	*Establish dominance;* *Gain social capital*
6. Others will learn not to mess with me.	.76	.06	*Establish dominance;* *Counter/prevent attack;*
7. Others will see what I'm capable of.	.82	.06	*Demonstrate efficacy;* *Gain social capital*
8. They will be afraid of me in the future.	.57	.06	*Demonstrate efficacy;* *Counter/prevent attack*
Harm to Self Items			
9. I will be vulnerable to retaliation.	.71	.06	*Immediate retaliation;* *Increased likelihood of future harm*
10. I will get hurt.	.70	.07	*Physical harm to self;* *Emotional harm to self*
11. I will have to watch my back in the future.	.69	.06	*Increased likelihood of future harm*
12. I will get into trouble.	.73	.06	*Formal punishment*
Harm to Victim Items			
13. The other person will really suffer.	.74	.06	*Physical harm to others;* *Emotional harm to others*
14. I may seriously hurt the other person.	.76	.06	*Physical harm to others;* *Emotional harm to others*
15. I will negatively affect their quality of life.	.50	.06	*Physical harm to others;* *Emotional harm to others*
16. I will harm the other person's reputation.	.74	.06	*Physical harm to others;* *Emotional harm to others*

REFERENCES

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology, 53*, 27-51.

Archer, J., & Haigh, A. M. (1997). Do beliefs about aggressive feelings and actions predict reported levels of aggression?. *British Journal of Social Psychology*, *36*, 83-105.

Bandura, A. (1965). Influence of models' reinforcement contingencies on the acquisition of imitative responses. *Journal of Personality and Social Psychology*, *1*, 589-595.

Bandura, A. (1969). Social-learning theory of identificatory processes. In D. A. Goslin (Ed.), *Handbook of Socialization Theory and Research*. Chicago: Rand McNally.

Bandura, A. (1973). *Aggression: A social learning analysis*. Prentice-Hall.

Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggression through imitation of aggressive models. *The Journal of Abnormal and Social Psychology*, *63*, 575-582.

Bandura, A., Ross, D., & Ross, S. A. (1963). Vicarious reinforcement and imitative learning. *The Journal of Abnormal and Social Psychology*, *67*, 601-607.

Baron, R. A., & Richardson, D. R. (2004). *Human aggression*. Springer Science & Business Media.

Barriga, A. Q., Landau, J. R., Stinson, B. L., Liau, A. K., & Gibbs, J. C. (2000). Cognitive distortion and problem behaviors in adolescents. *Criminal Justice and Behavior*, *27*, 36-56.

Baumeister, R. F., Bushman, B. J., & Campbell, W. K. (2000). Self-esteem, narcissism, and aggression: Does violence result from low self-esteem or from threatened egotism?. *Current Directions in Psychological Science*, *9*, 26-29.

Benning, S. D., Patrick, C. J., & Iacono, W. G. (2005). Psychopathy, startle blink modulation, and electrodermal reactivity in twin men. *Psychophysiology*, *42*, 753-762.

Bentley, K. M., & Li, A. K. (1996). Bully and victim problems in elementary schools and students' beliefs about aggression. *Canadian Journal of School Psychology*, *11*, 153-165.

Berkowitz, L. (1989). Frustration-aggression hypothesis: Examination and reformulation. *Psychological Bulletin*, *106*, 59-73.

Bettencourt, B., & Miller, N. (1996). Gender differences in aggression as a function of provocation: a meta-analysis. *Psychological Bulletin*, *119*, 422-447.

Bettencourt, B., Talley, A., Benjamin, A. J., & Valentine, J. (2006). Personality and aggressive behavior under provoking and neutral conditions: a meta-analytic review. *Psychological Bulletin*, *132*, 751-777.

Blanchard, D. C., & Blanchard, R. J. (2003). What can animal aggression research tell us about human aggression?. *Hormones and Behavior*, *44*, 171-177.

Blonigen, D. M. (2010). Explaining the relationship between age and crime: Contributions from the developmental literature on personality. *Clinical Psychology Review*, *30*, 89-100.

Boldizar, J. P., Perry, D. G., & Perry, L. C. (1989). Outcome values and aggression. *Child Development*, 571-579.

Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist, 16,* 681-684.

Bresin, K., & Robinson, M. D. (2015). You are what you see and choose: Agreeableness and situation selection. *Journal of Personality, 83*, 452-463.

Brown, T. A. (2014). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Burt, S. A., Hyde, L. W., Frick, P. J., Jaffee, S. R., Shaw, D. S., & Tremblay, R. (2018). Commentary: Childhood conduct problems are a public health crisis and require resources: a commentary on Rivenbark et al. (2018). *Journal of Child Psychology and Psychiatry*, *59*, 711-713.

Bushman, B. J., & Cooper, H. M. (1990). Effects of alcohol on human aggression: An intergrative research review. *Psychological Bulletin*, *107*, 341-354.

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

Campbell, A., Muncer, S., & Coyle, E. (1992). Social representation of aggression as an explanation of gender differences: A preliminary study. *Aggressive Behavior*, *18*, 95-108.

Chester, D. S. (2017). The role of positive affect in aggression. *Current Directions in Psychological Science*, *26*, 366-370.

Christiansen, B. A., Smith, G. T., Roehling, P. V., & Goldman, M. S. (1989). Using alcohol expectancies to predict adolescent drinking behavior after one year. *Journal of Consulting and Clinical Psychology, 57,* 93-99.

Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, *7*, 309-319.

Clark, L. A., & Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*.

Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin*, *115*, 74-101.

Crick, N. R., & Dodge, K. A. (1996). Social information□processing mechanisms in reactive and proactive aggression. *Child Development*, *67*, 993-1002.

Cruz, I. Y., & Dunn, M. E. (2003). Lowering risk for early alcohol use by challenging alcohol expectancies in elementary school children. *Journal of Consulting and Clinical Psychology*, *71*, 493-503.

De Castro, B. O., Veerman, J. W., Koops, W., Bosch, J. D., & Monshouwer, H. J. (2002). Hostile attribution of intent and aggressive behavior: A meta□analysis. *Child Development*, *73*, 916-934.

DeYoung, C. G. (2015). Cybernetic big five theory. *Journal of Research in Personality*, *56*, 33-58.

Dill, K. E., Anderson, C. A., Anderson, K. B., & Deuser, W. E. (1997). Effects of aggressive personality on social expectations and social perceptions. *Journal of Research in Personality*, *31*, 272-292.

Dodge, K. A. (2014). A social information processing model of social competence in children. In *Cognitive perspectives on children's social and behavioral development* (pp. 85-134). Psychology Press.

Dodge, K. A., & Crick, N. R. (1990). Social information-processing bases of aggressive behavior in children. *Personality and Social Psychology Bulletin*, *16*, 8-22.

Dodge, K. A., & Somberg, D. R. (1987). Hostile attributional biases among aggressive boys are exacerbated under conditions of threats to the self. *Child Development*, 213-224.

Dollard, J., Doob, L., Miller, N., Mowrer, O., & Sears, R. (1939). *Frustration and aggression*. New Haven, CT: Yale University Press

Donnellan, M. B., & Lucas, R. E. (2008). Age differences in the Big Five across the life span: evidence from two national samples. *Psychology and Aging*, *23*, 558-566.

Fleeson, W. (2012). Perspectives on the person: Rapid growth and opportunities for integration. In: Deaux, K, & Snyder, M. (Eds.), *The Oxford handbook of personality and social psychology* (pp. 33-63). New York, NY: Oxford University Press.

Fleeson, W., & Jayawickreme, E. (2015). Whole trait theory. *Journal of Research in Personality*, *56*, 82-92.

Fox, J. A., & Fridel, E. E. (2017). Gender differences in patterns and trends in US homicide, 1976–2015. *Violence and Gender*, *4*, 37-43.

Fulton, H. G., Krank, M. D., & Stewart, S. H. (2012). Outcome expectancy liking: A self-generated, self-coded measure predicts adolescent substance use trajectories. *Psychology of Addictive Behaviors*, *26*, 870-879.

Greenwood, J. D. (1999). Understanding the "cognitive revolution" in psychology. *Journal of the History of the Behavioral Sciences*, *35*, 1-22.

Griffin, S. A., & Samuel, D. B. (2014). A closer look at the lower-order structure of the Personality Inventory for DSM-5: Comparison with the Five-Factor Model. *Personality Disorders: Theory, Research, and Treatment*, *5*, 406-413.

Grusec, J. E. (1992). Social learning theory and developmental psychology: The legacies of Robert Sears and Albert Bandura. *Developmental Psychology*, *28*, 776-786.

Haggbloom, S. J., Warnick, R., Warnick, J. E., Jones, V. K., Yarbrough, G. L., Russell, T. M., ... & Monte, E. (2002). The 100 most eminent psychologists of the 20th century. *Review of General Psychology*, *6*, 139-152.

Hall, J. A., Herzberger, S. D., & Skowronski, K. J. (1998). Outcome expectancies and outcome values as predictors of children's aggression. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, *24*, 439-454.

Hampson, S. E. (2012). Personality processes: Mechanisms by which personality traits "get outside the skin". *Annual Review of Psychology*, *63*, 315-339.

Hare, R. D., & McPherson, L. M. (1984). Violent and aggressive behavior by criminal psychopaths. *International Journal of Law and Psychiatry, 7*, 35-50.

Haynes, S. N., Richard, D., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, *7*, 238-247.

Hull, J. G., & Bond, C. F. (1986). Social and behavioral consequences of alcohol consumption and expectancy: a meta-analysis. *Psychological Bulletin*, *99*, 347-360.

Hyatt, C. S., Chester, D. S., Zeichner, A., & Miller, J. D. (2020). Facet-level analysis of the relations between personality and laboratory aggression. *Aggressive Behavior, 3,* 266-277.

Hyatt, C. S., Zeichner, A., & Miller, J. D. (2019). Laboratory aggression and personality traits: A meta-analytic review. *Psychology of Violence*, manuscript in press.

Hyde, J. S. (1984). How large are gender differences in aggression? A developmental meta-analysis. *Developmental Psychology*, *20*, 722-736.

Jones, B. T., Corbin, W., & Fromme, K. (2001). A review of expectancy theory and alcohol consumption. *Addiction*, *96*, 57-72.

Jones, S. E., Miller, J. D., & Lynam, D. R. (2011). Personality, antisocial behavior, and aggression: A meta-analytic review. *Journal of Criminal Justice*, *39*, 329-337.

Knight, G. P., Guthrie, I. K., Page, M. C., & Fabes, R. A. (2002). Emotional arousal and gender differences in aggression: A meta□analysis. *Aggressive Behavior, 28*, 366-393.

Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., ... & Eaton, N. R. (2017). The Hierarchical Taxonomy of Psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, *126*, 454-477.

Krizan, Z., & Johar, O. (2015). Narcissistic rage revisited. *Journal of Personality and Social Psychology*, *108*, 784-801.

Leistico, A. M. R., Salekin, R. T., DeCoster, J., & Rogers, R. (2008). A large-scale meta-analysis relating the Hare measures of psychopathy to antisocial conduct. *Law and Human Behavior*, *32*, 28-45.

Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development*, *71*(1), 107-118.

Loevinger, J. (1954). The attenuation paradox in test theory. *Psychological Bulletin, 51,* 493-504.

Loewenstein, G., & Lerner, J. S. (2003). The role of affect in decision making. *Handbook of Affective Science*, 619-642.

Lynam, D. R., Gaughan, E. T., Miller, J. D., Miller, D. J., Mullins-Sweatt, S., & Widiger, T. A. (2011). Assessing the basic traits associated with psychopathy: Development and

validation of the Elemental Psychopathy Assessment. *Psychological Assessment*, *23*, 108-124.

Maples, J. L., Guan, L., Carter, N. T., & Miller, J. D. (2014). A test of the International Personality Item Pool representation of the Revised NEO Personality Inventory and development of a 120-item IPIP-based measure of the five-factor model. *Psychological Assessment*, *26*, 1070-1084.

Maples-Keller, J. L., Williamson, R. L., Sleep, C. E., Carter, N. T., Campbell, W. K., & Miller, J. D. (2019). Using item response theory to develop a 60-item representation of the NEO PI–R using the international personality item pool: Development of the IPIP–NEO–60. *Journal of Personality Assessment*, *101*, 4-15.

McAdams, D. P., & Pals, J. L. (2006). A new Big Five: fundamental principles for an integrative science of personality. *American Psychologist*, *61*, 204-217.

Meis, L. A., Murphy, C. M., & Winters, J. J. (2010). Outcome expectancies of partner abuse: Assessing perpetrators' expectancies and their associations with readiness to change, abuse, and relevant problems. *Assessment*, *17*, 30-43.

Miller, G. A. (2003). The cognitive revolution: a historical perspective. *Trends in Cognitive Sciences*, *7*, 141-144.

Miller, J. D., & Lynam, D. (2001). Structural models of personality and their relation to antisocial behavior: A meta□analytic review. *Criminology*, *39*, 765-798.

Miller, J. D., Lynam, D. R., & Jones, S. (2008). Externalizing behavior through the lens of the five-factor model: A focus on agreeableness and conscientiousness. *Journal of Personality Assessment*, *90*, 158-164.

Mischel, W. (1968). *Personality and assessment*. John Wiley & Sons Inc. Hoboken, NJ, US.

Moffitt, T. E. (2003). Life-course persistent and adolescence-limited antisocial behavior: A research review and a research agenda. In B. Lahey, T. E. Moffitt, & A. Caspi (Eds.), *The causes of conduct disorder and serious juvenile delinquency* (pp. 49 –75). New York, NY: Guilford.

Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychosocial adjustment. *JAMA*, *285*, 2094-2100.

Österman, K., Björkqvist, K., Lagerspetz, K. M., Kaukiainen, A., Huesmann, L. R., & Frączek, A. (1994). Peer and self□estimated aggression and victimization in 8□year□old children from five ethnic groups. *Aggressive Behavior*, *20*, 411-428.

Parrott, D. J., & Eckhardt, C. I. (2018). Effects of alcohol on human aggression. *Current Opinion in Psychology*, *19*, 1-5.

Perry, D. G., Perry, L. C., & Rasmussen, P. (1986). Cognitive social learning mediators of aggression. *Child Development*, 700-711.

Raine, A., Dodge, K., Loeber, R., Gatzke□Kopp, L., Lynam, D., Reynolds, C., ... & Liu, J. (2006). The reactive–proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, *32*, 159-171.

Reidy, D. E., Zeichner, A., & Seibert, L. A. (2011). Unprovoked aggression: Effects of psychopathic traits and sadism. *Journal of Personality*, *79*, 75-100.

Rivenbark, J. G., Odgers, C. L., Caspi, A., Harrington, H., Hogan, S., Houts, R. M., ... & Moffitt, T. E. (2018). The high societal costs of childhood conduct problems: evidence from

administrative records up to age 38 in a longitudinal birth cohort. *Journal of Child Psychology and Psychiatry*, *59*, 703-710.

Rotter, J. B. (1954). *Social learning and clinical psychology.* Upper Saddle River, NJ, US, NJ, US: Prentice-Hall, Inc. doi:10.1037/10788-000

Sellbom, M., & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment, 31,* 1428-1421.

Sleep, C. S., Lynam, D. R., & Miller, J. D. (2020). A comparison of the validity of very brief measures of the Big Five/Five Factor Model of personality. *Assessment*, manuscript in press.

Smithmyer, C. M., Hubbard, J. A., & Simons, R. F. (2000). Proactive and reactive aggression in delinquent adolescents: Relations to aggression outcome expectancies. *Journal of Clinical Child Psychology*, *29*, 86-93.

Tackett, J. L., Kushner, S. C., Herzhoff, K., Smack, A. J., & Reardon, K. W. (2014). Viewing relational aggression through multiple lenses: Temperament, personality, and personality pathology. *Development and Psychopathology*, *26*, 863-877.

Tackett, J. L., Waldman, I. D., & Lahey, B. B. (2009). Etiology and measurement of relational aggression: A multi-informant behavior genetic investigation. *Journal of Abnormal Psychology*, *118*, 722-733.

Taylor, S. P., Gammon, C. B., & Capasso, D. R. (1976). Aggression as a function of the interaction of alcohol and threat. *Journal of Personality and Social Psychology*, *34*, 938.

Treloar, H., Pedersen, S. L., & McCarthy, D. M. (2015). The role of expectancy in substance-abuse progression. *Addictions: A Social Psychological Perspective*, 120-147.

Trudeau, L., Spoth, R., Lillehoj, C., Redmond, C., & Wickrama, K. A. (2003). Effects of a preventive intervention on adolescent substance use initiation, expectancies, and refusal intentions. *Prevention Science*, *4*(2), 109-122.

Smith, G. T., McCarthy, D. M., & Zapolski, T. C. (2009). On the value of homogeneous constructs for construct validation, theory testing, and the description of psychopathology. *Psychological Assessment*, *21*, 272-284.

Vize, C. E., & Lynam, D. R. (2020). Merging structural and process□related approaches to the study of agreeableness: A preregistered replication and extension attempt. *European Journal of Personality*, *34*, 139-163.

Vize, C. E., Lynam, D. R., Collison, K. L., & Miller, J. D. (2018). Differences among dark triad components: A meta-analytic investigation. *Personality Disorders: Theory, Research, and Treatment*, *9*, 101.

Vogell-Sprott, M., & Fillmore, M. T. (1999). *Learning Theory and Research*. (K. E. Leonard & H. T. Blane, Eds.) (2nd ed., pp. 292–327). New York, NY: Guilford Press.

Walters, G. D. (2007). Predicting institutional adjustment with the Lifestyle Criminality Screening Form and the Antisocial Features and Aggression scales of the PAI. *Journal of Personality Assessment*, *88*, 99-105.

Wilkowski, B. M., Robinson, M. D., & Meier, B. P. (2006). Agreeableness and the prolonged spatial processing of antisocial and prosocial information. *Journal of Research in Personality, 40*, 1152–1168

Wilkowski, B. M., & Robinson, M. D. (2010). The anatomy of anger: An integrative cognitive model of trait anger and reactive aggression. *Journal of Personality*, *78*, 9-38.

Wrangham, R. W. (2018). Two types of aggression in human evolution. *Proceedings of the National Academy of Sciences*, *115*, 245-253.