Cross-Cultural Generalizability of the Comprehensive Assessment of Psychopathic Personality (CAPP) in South Korea

by

Jonghan Sea

M.A., Ajou University, 2005

B.Sc., Ajou University, 2002

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Psychology
Faculty of Arts and Social Science

SIMON FRASER UNIVERSITY

Spring 2018

Approval

Name:	**Jonghan Sea**
Degree:	**Doctor of Philosophy**
Title:	**Cross-Cultural Generalizability of the Comprehensive Assessment of Psychopathic Personality (CAPP) in South Korea**
Examining Committee:	**Chair:** **Tim Racine** Professor
	Stephen D. Hart Senior Supervisor Professor
	Kevin S. Douglas Supervisor Professor
	P. Randall Kropp Supervisor Registered Psychologist
	Eric Beauregard Internal Examiner Professor School of Criminology
	John F. Edens External Examiner Professor Department of Psychology Texas A&M University
Date Defended/Approved:	April 17, 2018

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University

or has conducted the research

c. as a co-investigator, collaborator, or research assistant in a research project approved in advance.

A copy of the approval letter has been filed with the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

Update Spring 2016

Abstract

The Comprehensive Assessment of Psychopathic Personality (CAPP; Cooke, Hart, Logan, & Michie, 2013) is a new lexically-based conceptual model of psychopathy that has potential clinical utility. The main purpose of the current research was to investigate the generalizability of the CAPP conceptual model in South Korea. In Study 1, I conducted a study in which experts and lay people in South Korea were asked to rate the prototypicality of symptoms of psychopathy using a Korean language translation of the CAPP model (K-CAPP). The results indicated that, consistent with past research in other countries, Korean experts and lay people on average rated K-CAPP symptoms as being moderately to highly prototypical of psychopathy, and also more prototypical of psychopathy on average than symptoms theoretically unrelated to psychopathy. The prototypicality ratings for K-CAPP symptoms made by Korean experts and lay people were similar to each other, as well as to those made by experts and lay people using the CAPP in other countries. In Study 2, I evaluated the reliability and concurrent validity of expert ratings of psychopathy made using a Korean translation of a CAPP-based clinical measure, CAPP-Institutional Rating Scale (K-CAPP-IRS), in a sample of correctional offenders in South Korea. Reliability analyses based on simple intraclass correlations indicated very high (> .80) interrater reliability for almost all the K-CAPP-IRS symptom, domain, and total ratings. But a more sophisticated examination using a Generalizability Theory framework, with a Persons (89 offenders) x Raters (3 experts) x Occasions (2 occasions, three-month interval) x Items (33 K-CAPP-IRS symptoms) design, revealed complex but substantial interactions involving Raters; however, the impact of these interactions was mitigated when K-CAPP-IRS ratings were made by increasing the number of Raters, as opposed to Occasions. Concurrent validity analyses that K-CAPP-IRS total scores were correlated highly with total scores on the Korean translation of the Hare Psychopathy Checklist-Revised (Cho & Lee, 2008), r = .647; and moderately with total scores on the Korean translation of the Psychopathic Personality Inventory-Revised (Park & Lee, 2013), r = .350. Overall, the results of Studies 1 and 2 indicate that the concept of psychopathy, as captured by the CAPP concept map, appears to be cross-culturally valid in South Korea.

Keywords: Psychopathy; CAPP-IRS; Cross-cultural; G-theory; Prototypicality rating

Dedication

Dedicated to my lovely wife, Moon

Her sacrifice and unconditional support make me who I am today

Acknowledgements

First of all, I'd like to thank my dissertation committee members, Tim, Steve, Kevin, Eric, Randy, and John. From the very first moment in arriving in Vancouver, my supervisor, Stephen Hart has supported me with gentle smile, warm heart, countless knowledge, and spice-like humor. Whenever I asked for his help, he was beside me like Iron man. He has encouraged me to finish all the adventure and safely arrived at the port, leading to another adventure. Thanks, my dear friend, Steve.

I would also like to thank Eric Beauregard for his cooperation, insight, and invaluable feedback. Through his comments and knowledge, I could accomplish important research and project during my graduate life. I love his French-style English, smiling, mustache, and his generous appetite, liking cabbage and radish Kimchi. Thanks, Eric.

I would also like to thank Kevin Douglas for his feedback and help. I should say he has been enthusiastic, friendly, and knowledgeable in all ways, especially teaching me. I've learned his inspiration and expertise through his lectures. I hope I see you with crispy Korean chicken and beers in Korean restaurants sooner or later. Thanks. Kevin.

Personally, I want to special thank for all of Korean correctional psychologists in Korean National Correctional Service and Ph.D. student researchers in Ajou University as well as Bang Sungil in Simon Fraser University. Specially, I want to thank Ms. Jung and Professor Kim Kyungil, my M.A. supervisor, who provided complete support and belief for the project. All things were well done by their concern and encouragement. Thanks, all of you.

Lastly, I want to thank my dear colleagues, Tara, Sarah, and Yan. They welcomed to me in Simon Fraser with warm heart and guided me to live in Vancouver. They always be my friends in my memory. Thanks, generous guys.

Table of Contents

List of Tables

List of Figures

Chapter 1.

Introduction

1.1. Personality and Personality Disorder across Cultures

1.1.1. Cultural Pathoplasticity of Personality and Personality Disorder

Zutt (1967) and Lenz (1964), fifty years ago, first proposed the term pathoplasticity in German psychiatry in order to describe the culture-sensitive part of the symptomatology of personality, related symptoms, and mental disorder. Whether and to which extent the prevalence and shape of personality and related disorders depends on culture has long been a question. More recently, in a similar context, Cooke, Michie, Hart and Clark (2005) released critical findings that culture has an important effect on personality disorder as a pathogenic/pathoplastic agent and moreover the cultural effect differs across symptoms. It was suggested that personality could have a syndromal structure that is different across cultures and the presence of specificity of effect could increase the plausibility of cultural facilitation (Susser, 1973). Alacon and Foulks (1995) also found out that cultural contextualization plays a significant role in the construction of self-concept, self-image, the egocentric/sociocentric dichotomy and the determinations of biases, playing three roles as an interpretive/explanatory tool, as a pathogenic/patholastic agent, and as a diagnostic/nosological factor on personality and related disorders.

Until today, most professionals would agree that cultural pattern may influence personality and these disorders, but still in a wide sense this is still an unsolved question. On these accounts, I tried to investigate the cultural sensitive proportion of the total variance on personality and personality disorder among different societies.

1.1.2. How Does Culture Shape Personality and Personality Disorder?

According to Ashton and Lee (2007, 2009), people differently act, think, feel and experience in a different way as a function of their cultural position (Nisbett, Peng et al, 2001). Thus the flexible model and multi-dimensional explanation to personalities should be adjusted like the HEXACO model [PI-R version; Ashton & Lee, 2009] (e.g., Honesty-

Humility, Emotionality, eXtraversion, Agreeableness versus Anger, Conscientiousness, and Openness to Experience; Ashton & Lee, 2007) or Hwabyung syndrome (a unique Korean culture syndrome, meaning 'anger syndrome'; Kim, 2007; Kim, Hwang, & Park, 2017), which is more focusing on interpersonal relations and a different perspective of the self (Cheung, Cheung, & Jainxin, 2004; Lin, 2017). For instance, studies related to the HEXACO model revealed significant associations among dark triad such as psychopathy, narcissism and Machiavellianism based on German, Chinese, Indian, etc. with culturally perspectives (Ashton & Lee, 2007; Witt, Donnellan & Blonigen, 2009; Miller, Gaughan et al, 2009; Lee et al., 2013; Ruchensky & Donnellan, 2017). This also means the study of Personality Disorder might need to be framed (Lynam & Widiger, 2007) to reflect distinctively cultural indigenes.

As culture itself could engender a culturally valued trait to the extent that there may be greater levels of the trait in the population, the culture might be one of causative factors for a personality disorder, which could be deduced from the gaps in the rates of violent crimes across cultures and also the differences in the rates of personality disorder (Harrendorf, Heiskanen, & Malby, 2010). Furthermore, according to Five-Factor Model (FFM; Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism; McCrae & Costa, 1997; McCrae et al., 1998; McCrae et al., 2005), European and North American cultures were higher in extraversion and openness to experience, and lower in agreeableness than Asian and African cultures (Allik & McCrare, 2004; Schmitt et al., 2007). This could indicate that the two Western cultures had much greater variabilities, suggesting a greater heterogeneity than Asian and African cultures (Costa et al., 2001). Also, some research compared the culture of the Filipions with that of Western countries such as Americans, New Zealanders and Canadians, showing that the Eastern country scored relatively low on Neuroticism on average compared to other cultures measured while scoring in the middle of the scale of Extraversion. However, Western people scored higher on Extraversion while scoring moderately on Neuroticism.

In the regard, it may be no wonder that this could suggest cultures with similar backgrounds or history shared more analogous personality profiles (Schmitt et al., 2007) and the words that cultures can have an impact on the prevalence of certain personality traits could be supported. Suggested by Ryder, Sun, Dere, & Fung. (2014), as the level of [functional] impairment associated with the personality disorder could differently be acknowledged in cultural variations as well as personal and social reaction to the extremity

of traits could or could not result in functional impairment, culture may have an impact on whether the personality would be considered disordered or not.

1.1.3. Representation of PPD across Cultures

Hare (1998) came to acknowledge that the behavioral representation of the personality disorder relied on both the societal and cultural context and he further admitted that “it is more difficult to determine how psychopaths express themselves in societies that are highly structured and in which there some strong traditions to conform to group standard” (p. 105). In line with aforementioned statement, there would be cross-cultural gaps in the way PPD manifested. For example, even some of PPD traits would be cardinal elements in Western societies; however, it would be secondary or marginal in Eastern societies. In this context, differences in interpersonal symptoms were found when comparing samples from South Korea (referred to simply as Korea hereinafter) to samples from the United Kingdom and the differences could be because of Korean cultural background, Korean Confucianism (Sea, Lee, & Cooke, 2017). Another study also revealed that the differences were greatest for an arrogant and deceptive interpersonal style and least for both deficient affective experience and impulsive irresponsible behavioral style, comparing Korean samples to Canadian samples (Sea, Hart, Lee, 2017 in preparation). In addition, the differences in interpersonal and affective items were noted when comparing Iranian samples to North American samples, which were thought to be due to an Iranian cultural practice known as ta’arof (see Shariat et al., 2010). Ta’arof was described as “the great national trait of exaggerated politesse, modesty, and self-deprecation that Iranians seem to be born with” (p. 819).

Specifically, Shariat et al. (2010) indicated that being too much superficial and charming in North America may be related to some problematic behaviors for individual but in Iran, culturally, those traits could not be pathological and no impairments because of *taarof* which is socially accepted and common. Shariet et al. (2010) also pointed out that lack of empathy and remorse could be greatly discriminant even at the low levels because of the cultural variations to the traits. This could be due to the fact that regarding the interpersonal interaction value collectivistic societies may be more sensitive to self-centered emotions and expressions, which could be recognized to be deviant in Iranian society and others. For example, a study by Caldwell-Harris and Aycicegi (2006) suggested that idiocentric personality styles (e.g., egocentric-independent attitude and

lifestyle) were correlated with poorer social outcomes among Turkish students, whereas among American students they are associated with better psychological adjustment.

1.1.4. Individualism versus Collectivism on PPD

In terms of more general cross-cultural contexts, Markus and Kitayama (1991, 1994) suggested that a number of major cultural differences between the US and non-Western countries result from cultural differences in independence vs. interdependence or collectivism vs. individualism (Cooke, Hart, & Michie, 2004; Hofstede, Hofstede, & Minkov, 1991). Simply stated, while US society stresses the independence of the self, Japanese and Korean societies emphasize the interdependence of individuals (Markus & Kitayama, 1994; Fiske, Kitayama, Marksu, & Nisbett., 1998; Smith & Bond, 2003). In Korea, the word for conformity has a positive connotation, meaning maturity and inner strength, which is also similar with harmony in China (Uno, 1991). Based on this perspective, impulsive anti-sociality, egocentricity, irresponsible, parasitic and carefree lifestyle in Korea is perceived in an essentially or qualitatively different way from the U.K. or U.S.A. That is, differences in independence and interdependence in USA and Japan could potentially influence the metric as well as structure equivalence of the various psychopathic models in conceptions and measures (Yokota, 2012).

Cultural differences are often evoked in response to aspects of the social and ecological environments that differ across different cultures (Tooby & Cosmides, 1992; Yokota, 2012). As such, cultural differences could be used to explain why differences in psychological presentation occur. In other words, individual's behavior can be thought of as an ecologically rational behavior, adapted to different types of cultural backgrounds. In the case of the facets or factors of psychopathy and their relationship, it is likely that differences will be created in response to behavioral differences in the social environment of Korea, Japan and China versus Western societies, particularly domains associated with interpersonal relationships and personal lifestyles. For example, in Japan, China and Korea, it is crucial what they are specifically supposed to do at any given age and to fulfill their social duty and moral values. East Asians feel satisfaction when they are doing what they should be doing according to social norms (Suh, Diener, Oishi, & Triandis, 1998). However, in case of showing antisocial behaviors (criminal acts such as raping others and so on), there is much more risks or anxiety of being alienated from social bondages (Yamagishi, Hashimoto, & Schug, 2008). Therefore, most people in East Asian cultural

contexts are actively and openly supposed to follow norms. On the other hand, in the USA, where uniqueness and assertiveness are regarded as important and desirable, self-assertion of one's superiority is an adaptive strategy to dominate or control others (Wierzbicka, 1997), which might be related to conning or manipulative. Displays of anger which lead to aggressiveness associated with poor behavior controls and impulsivity are naturally likely to be suppressed in Korea and Japan (Matsumoto, Kudoh, Scherer, & Wallbott, 1988). Maintaining harmony with others requires not only sympathy but also the concealment of aggressiveness (Markus, Kitayama, & Heiman, 1996). Also, seeking and receiving social resources and intimate bonds are more essential in Asian societies (Kitayama & Uchida, 2005). Therefore, emotional traits such as callousness, lack of empathy and lack of remorse or guilt could lead to possibly avoidance in Asian cultures.

1.1.5. Geert Hofstede's Dimensions on PPD

Hofstede's cultural dimensions theory is a well-known and robust framework for cross-cultural communication, developed by Geert Hofstede. It describes the effects of a society's culture on the values of its members and how these values related to behaviors. He developed his original model with culturally diverse perspectives and proposed six dimensions: power distance, individualism, masculinity, uncertainty avoidance, long-term orientation and indulgence (see details in Hofstede, 1991, 2001, 2010, 2011; Hofstede, Hofstede, & Minkov, 1991). First, Power Distance index (PDI) is defined as 'the extent to which the less powerful members of institutions within a country accept and expect that power is distributed unequally.' In this dimension, Korea is slightly more hierarchical society than UK and North America. This could suggest that the Korean naturally accepts a hierarchical structure in which everybody has a place, reflecting inherent inequalities and acceptable centralization. In addition, Korea is considered to be one of the strong collectivistic societies, indicating social relationships where everyone takes responsibility for members in their own group are important.

Due to these cultural perspectives, Korean society tends to be more deceptive because such cultural phenomena could encourage glibness and grandiose as easy or/and effective ways in order to form social connections as well as desirable responsibilities. Overall, it might be regarded as a culturally acceptable characteristic that would be similar to Iranian Taarof (Fathi, 2004; The Art of Taarof, 2004; Shariat et al., 2010) as well as Asian Confucianism (Uno, 1991; Littlejohn, 2010). This characteristic is

a form of discretion and compliment (similar with glibness) with which Eastern people pro-socially and verbally try to show their respect, honor and hospitality in a colorful way. In line with Iranian and Chinese societies (i.e., collectivistic) suggested by Shariat et al (2010) and Assadi et al. (2006) as well as Billioud and Joel (2015), individuals in Korea may be somewhat deceitful in this specific context; such behavior cannot be considered as pathological as it does not markedly deviate in collectivism-society. It should not be used to rate deceitfulness on this basis.

Second, Masculinity vs. femininity index (MAS) is defined as 'a preference in society for achievement, heroism, assertiveness and material rewards for success vs. a preference for cooperation, modesty, caring for the weak and quality of life with others.' Regarding this dimension, Korea is considered feminine to care for others and their quality life and for consensus, modesty, solidarity and quality in their working lives. Thus, decision making in people's lives is achieving through involvement such as compromise and negotiation. In that context, social characteristics including self-centered, egocentric, selfish and self-absorbed could be discouraged.

Uncertainty avoidance (UA) is defined as the extent to which the members of a culture feel threatened by ambiguous situations and have created beliefs and institutions that try to avoid obscurities. According to the Hofstede's cultural model (Hofstede, 2011), Korea is one of the most uncertainty avoiding countries in the world which stick to inflexible or rigid codes of belief and behavior and are intolerant of unorthodox behavior and ideas while in UK or Canada as a nation doesn't generally have too many rules and accepts more uncertainty. Conversely, in Korea, social security is an important issue in individual motivation and behavior and moreover, in terms of long-term orientation (i.e., how every society has to maintain links with its own past while dealing with the challenges of the present and the future), Korea is also one of the most pragmatic, long-term oriented societies. Koreans tend to live their lives guided by virtues and practical good examples. Overall, these phenomena could contribute to social suppression [functioning as inhibitors or stabilizers; Douglas et al., 2013] on the following symptoms such as impulsive, reckless, restless, unreliable, lacks perseverance, disruptive and so forth.

As for indulgence which is defined as the extent to which people try to control their desires and impulses, Korean society has been shown to be one of the restraint countries which have a strong tendency to criticize others' behaviors. Unlike UK, an indulgent

society, Korea as a restrained society does not put much emphasis on independence or freewheelingness but control the gratification of their desires and stimulation in a community. Thus, Koreans feel that indulging without orientation is thought to be somewhat wrong, accepting that their actions are restrained by social norms. In this situation, people have a tendency to orient themselves to the surroundings and be faithful to social regulation whereas symptoms such as impulsive, sense of being extraordinary, exceptional, special, sense of being deserving, sense of being invincible and so on were more likely to be suppressed.

1.1.6. Measuring PPD across Cultures

Most studies trying to validate the cross-culture of PPD have used the Psychopathy Checklist-Revised (PCL-R; Hare, 1991) (Cooke et al., 2005a; Olver et al., 2013; Shariat et al., 2010; Wilson et al., 2014) for use in decision making (Cooke et al., 2004b). Numerous studies showed potential difference and its presentation in different cultures around the world (McCord & McCord, 1964, p. 87; Tutuncu et al., 2015; Morana et al., 2015; Wilson et al., 2014), echoing "the behaviors expressions of psychopathy, as well as the degree to which they stand out from the behavior of others, are influenced by societal and cultural structures and norms" (Hare, 2006, p. 106). In Western societies, Cooke, Michie, and other colleagues have found out the significant difference in prevalence and presentation using the PCL-R (e.g., Cooke & Michie, 1999; Cooke et al., 2004; Cooke et al., 2005; Wilson et al., 2014; Zukauskiene et al., 2010) by comparisons between European samples (e.g., Scotland, UK, Bulgaria, Lithuania) and North American samples.

There has also been dramatically increasing interest in the construct of psychopathy among East Asian researchers. The PCL-R has been translated into Chinese (Liu, Huang, & Lv, 2010; Shi, Lv, Zhang, & Han, 2013), Japanese (Yuki, 2008) and Korean (Joe & Lee, 2008). Shi et al. (2013) demonstrated that violent schizophrenia patients had significantly higher scores in the PCL-R affective and antisocial facets than both the healthy control and the schizophrenia without violence. Lui et al. (2010) found that the second factor of PCL-R which describes social deviance or a chronically unstable and antisocial lifestyle, was significantly associated with antisocial personality disorder in Chinese forensic samples. Zhou et al. (2016) also reviewed that PCL-R is good/fair to excellent in the reliability and validity through meta-analysis studies in Chinese sample

(Liu et al., 2010). In Korea, Lee et al. (2015) study also compared two-, three- and four-factor PCL-R models with samples of 1092 adult criminal offenders. The three-factor model (Cooke & Michie, 2001) provided the better fit to the most current data, followed by the four-factor and two-factor model, which is in accordance with previous literatures (Sohn & Lee, 2014; Sohn, Webb, & Lee, 2010), but still there is a debating issue as a substantial number of literatures have approved the four-factor model (Hare, 2003; Hill, Neumann, & Rogers, 2004; Hare & Neumann, 2005; Neumann, Hare, & Newman, 2007; Vitacco, Neumann, & Jackson, 2005; Lee, Ko, & Kim, 2009).

IRT is a set of mathematical models that are used for psychometric analysis and scoring and has been increasingly popular in areas of measurement of clinical and personality constructs (Cooke & Michie, 1997; Embreston & Reise, 2013; Morizot, Anisworth, & Reise, 2009; Thomas & Locke, 2010; Walton, Roberts, Krueger, Blonigen, & Hicks, 2008; Shou, Sellbom, & Xu, 2017). IRT (Item Response Theory; Embretson & Reise, 2013) analysis indicated that offenders with equivalent standing on the underlying latent trait of psychopathy in North America received total scores on the PCL-R that were two or three points higher than their counterparts from Europe. The cross-cultural psychometric differences were largest for 'interpersonal and behavioral symptoms' of psychopathy whereas smallest for 'affective symptoms'. The 'defective affective experience' component of psychopathy may relatively be the more stable aspect of the disorder across cultures. Shariat et al. (2010) also examined that the factor structure of the PCL: SV (Hart, Cox, & Hare, 1995) in the Iranian sample (e.g., 351 prison inmates in Qasr prison); this fitted well with the three-factor hierarchical model. IRT analysis showed that significant differential item function (DIF) exists between the Iranian and the North American sample. The most significant differences were related to two of the three factors: Arrogant and Deceitful Interpersonal Style and Deficient Emotional Experience. All three items of factor 1 had less discriminatory power in the Iranian sample. These findings were theorized to be associated with the cultural characteristics in the Iranian society. As aforementioned, Maghbouleh (2013) suggested this could be related to a Persian cultural term known as ta'arof representing a compliment style to give their respect toward others in their own way (Hashemain, Roohani, & Karami, 2016).

Also, through the examination of measurement bias via DIF analysis, Shou, Sellbom, and Xu (2017) recently suggested a number of items of the Triarchic Psychopathy Measure (TriPM; Patrick, 2014) could attribute to non-equivalence across

the Chinese and the U.S. samples. Some modifications and adaptations of items should be considered in TriPM for Chinese participants, especially indicating cross-cultural differences (e.g., Disinhibition and Meanness items had smaller discrimination power for Chinese sample than for U.S. sample). Importantly, this is possible because parents in Asia are usually more controlling and critical and have much lower tolerance of children's lack of control (Ng, Pomerantz, & Deng, 2014). Furthermore, Boldness scale items had DIFs. This may be due to the fact that, as Yokota (2012) suggested, dominance is effective in controlling others depends on social environments; thus in East Asia culture, asserting one's superiority is not the most adaptive strategy for controlling others, as it is not conforming to social rules, such as Renqing orientation (the degree to which a person abides to rules of reciprocity regarding favor exchanges; Huang, 2011; Shou, Sellbom, & Xu, 2017).

Other researchers have recently examined cross-cultural equivalence of self-reporting psychopathy measurement in Asian cultures. This research has included Levenson's Self-Report Psychopathy Scale (LSRP; Levenson, Kiehl, & Fitzpatrick, 1995; Japanese version, Osumi, Kanayam, Sugiura, & Ohira, 2007; Chinese version, Shou, Sellbom, & Han, 2016a), the Psychopathic Personality Inventory–Revised (PPI-R; Lilienfeld & Widows, 2005; Japanese version, Yokota, 2012; Korean version, Lee & Park, 2007) and the Triarchic Psychopathic Measure (TriPM; Patrick, 2014; Chinese Version, Shou, Sellbom, & Han, 2016b; no Korean version). However, the evidence for the cross-cultural validity of these measures is still inconclusive (Shou, Sellbom, Xu, Chen, and Sui, 2016). Although the external correlates of various LSRP scores, PPI-R and SRP-III scores seem promising in Korean, Japanese and Chinese samples (Zhang, Zou, Wang, & Finy, 2015; Smithson & Shou, 2015; Osumi, Kanayama, Sugiura, & Ohira, 2007; Masui, Fujiwara, & Ura, 2013; Lee & Park, 2007), the internal structure of most measurements seems to be problematic in East Asian samples (Williams, Paulhus, & Hare, 2007; Neal & Sellbom, 2012; Neumann, Schmitt, Carter, Embley, & Hare, 2012; Yokota, 2012; Lee, Gong, & Park, 2007). Specifically, findings relating to the latent factor structure (e.g., Hare's two-factor model and the PPI-R three factor model) in these studies were commonly inconsistent with previous studies that used Western samples. There is also a lack of integration of East Asian indigenous perspectives and the Western theorizing and operationalization of psychopathy. For this reason, current knowledge about the construct of psychopathy and its manifestation among East Asians is still limited.

To date, the major finding from previous psychometric studies is the evidence of small but robust cross-cultural differences in the assessment of psychopathy (Cooke, Hart, & Michie, 2004; Cooke, Kosson, & Michie, 2001; Cooke et al., 2005b; Skeem, Edens, Camp, & Cowell, 2004; Ullrich, Paelecke, Kahle, & Marneros, 2003; Shariat et al., 2010; Boduszek & Debowska, 2016). On these accounts, research on the various theories and measurements for use to assess PPD should be evaluated and further whether the core features of psychopathy are consistent across different cultures be also explored. More fundamentally, Cooke et al. (2004b) have argued that "The PCL-R may be regarded as having caused construct drift–drift from the traditional conceptualization of the disorder….it may be time to correct course before running into the danger that the measure becomes the construct" (p. 350). This raises a request to develop a measure including encompassing a full core psychopathy components applied to a variety of settings across cultures, based on the literatures.

1.2. A New Approach to Studying PPD across Cultures

The PCL-R is widely considered the "gold standard" for assessing PPD in adult male correctional offenders and forensic psychiatric patients (Hare et al., 2000; Singh et al., 2004). The PCL-R is not without limitations, however. First, concern has been expressed concerning the extent to which the PCL-R adequately evaluates the full range of PPD symptomatology. For example, some have argued that the way some features of PPD are defined in the PCL-R over-focuses on or is unduly influenced by official criminality (Hoff et al., 2014), and others have argued that "the PCL-R excludes or provides insufficient coverage of some core symptoms of PPD, such as anxiety" (Skeem & Cooke, 2010, p. 436).

Second, the PCL-R measures the lifetime presence of features of PPD. Thus, it implicitly assumes that the PPD is highly stable over time, making it impossible to assess the current severity or changes in the severity of PPD symptoms (Hoff et al., 2014).

Third, the PCL-R was developed on the basis of research using adult male correctional offenders and forensic psychiatric patients in North America, and subsequently has been used for both research and clinical purposes in diverse populations without research confirming it is invariant across age, gender, or culture. But some have argued that the way features of PPD are defined and measured in the PCL-R

makes it unlikely that the test can capture diversity across groups in the presentation of PPD (Hoff et al., 2014; Boduszek & Debowska, 2016; Skeem & Cooke, 2010; Skeem, Edens, Camp, & Cowell, 2004; Ullrich, Paelecke, Kahle, & Marneros, 2003).

Fourth, and perhaps most importantly, the PCL-R was based on a conceptualization of PPD that is either unclear, vague, or idiosyncratic. Hare has discussed that his original development of the test was influenced by the classic clinical description of Cleckley (Cleckly, 1988, 1976, 1941). Some have argued that the PCL-R incompletely or even inaccurately reflects Cleckley's conceptualization of PPD (Patrick, 2005; Patrick & Drislane, 2015; Patrick, Fowles, & Krueger, 2009). In response, Hare has noted that he was also influenced by other clinical descriptions of PPD, as well as by his own experiences conducting research with offenders (Hare & Neumann, 2010). But in the absence of a clear statement by Hare concerning the theoretical model on which the PCL-R was based, it is impossible to critically evaluate either the theoretical model itself or the extent to which the PCL-R is faithful to it.

In response to the perceived limitations of the PCL-R, a variety of theoretical models of PPD have been proposed in recent years, such as the Five-Factor Model (FFM; Lynam & Widger, 1998, 2007), the Triarchic Model of Psychopathy (TriMP; Patrick & Drislane, 2015), and the Comprehensive Assessment of Psychopathic Personality (CAPP; Cooke, Hart, Logan, & Michie, 2004, 2012), as well as a number of measures of PPD based on those models, such as the FFM (Lynam & Widger, 1998, 2007), the Triarchic Psychopathy Measure (TriPM; Patrick, Fowles, & Krueger, 2009), and the CAPP Institutional Rating Scale (CAPP-IRS; Cooke et al., 2012). Below, I discuss the CAPP model of PPD and the CAPP-IRS measure of PPD in some detail, as they formed the basis for the research described herein.

1.2.1. The CAPP Model

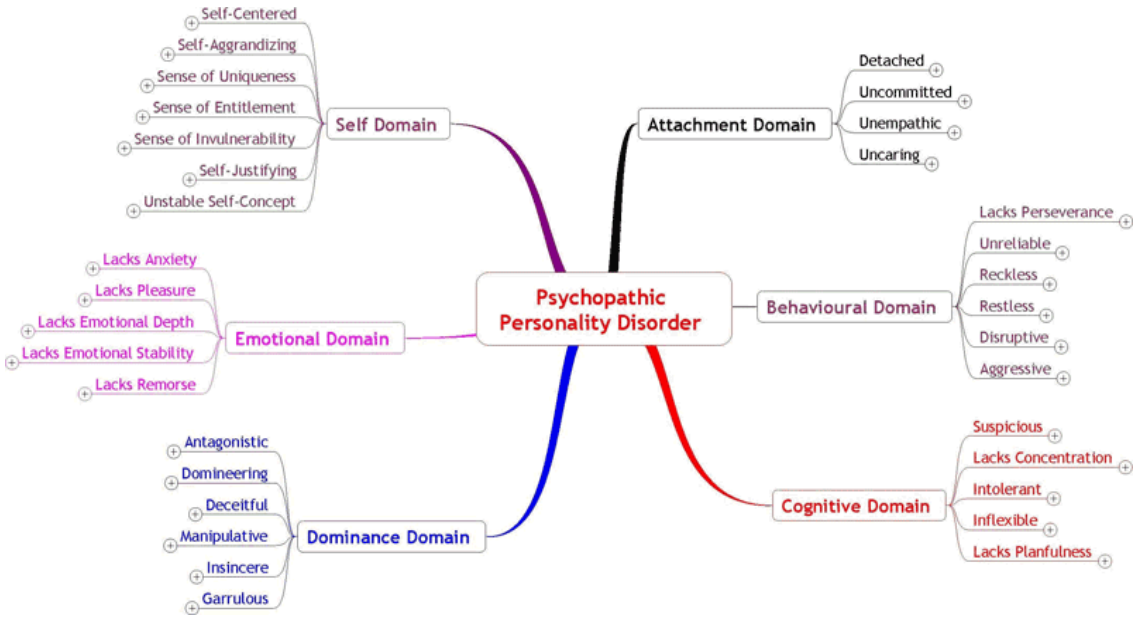

Figure 1. The CAPP Model
Note. Retrieved from http://www.gcu.ac.uk/capp/model/index.html)

The CAPP was developed as a theoretical model of PPD, or what also may be referred to as a descriptive model or "concept map" of the disorder. Based on an extensive literature review and consultation with international subject matter experts (SMEs), Cooke and colleagues identified a large set of putative features of PPD, erring on the side of overinclusivenss by including features concerning which there was lack of unanimity in the literature or among SMEs. Based on lexical approach to understanding personality and personality disorder, which assumes that important features of personality should be encoded in lay language (e.g., Goldberg, 1993; Saucier & Goldberg, 2001), Cooke and colleagues translated the features into trait-descriptive adjectives or brief adjectival phrases, then grouped cognate terms into clusters using lexical tools such as dictionaries and thesauruses. Each cluster was then reduced to four trait-descriptive adjectives or brief adjectival phrases: one selected to represent the symptom (i.e., specific trait), and three others selected to define that symptom. The end result was a set of 33 symptoms. To help organize and further clarify the meaning of the symptoms, they were then divided rationally according to the six domains of personality functioning they reflected: the Attachment,

Behavioural, Cognitive, Dominance, Emotional, and Self domains. The full CAPP model is presented in Table 1 and Figure 1.

Table 1. Description of the CAPP Model

Domain	Symptom		Definition
Attachment	A1	Detached	Remote, Cold, Distant
	A2	Uncommitted	Unfaithful, Undevoted, Disloyal
	A3	Unempathic	Uncompassionate, Cruel, Callous
	A4	Uncaring	Inconsiderate, Thoughtless, Neglectful
Behavioural	B1	Lacks Perseverance	Idle, Undisciplined, Unconscientious
	B2	Unreliable	Undependable, Untrustworthy, Irresponsible
	B3	Reckless	Rash, Impetuous, Risk-Taking
	B4	Restless	Overactive, Fidgety, Energetic
	B5	Disruptive	Disobedient, Unruly, Unmanageable
	B6	Aggressive	Threatening, Violent, Bullying
Cognitive	C1	Suspicious	Distrustful, Guarded, Hypervigilant
	C2	Lacks Concentration	Distractible, Inattentive, Unfocused
	C3	Intolerant	Narrow-minded, Bigoted, Hypercritical
	C4	Inflexible	Stubborn, Rigid, Uncompromising
	C5	Lacks Planfulness	Aimless, Unsystematic, Disorganized
Dominance	D1	Antagonistic	Hostile, Disagreeable, Contemptuous
	D2	Domineering	Arrogant, Overbearing, Controlling
	D3	Deceitful	Dishonest, Deceptive, Duplicitous
	D4	Manipulative	Devious, Exploitative, Calculating
	D5	Insincere	Superficial, Slick, Evasive
	D6	Garrulous	Glib, Verbose, Pretentious
Emotional	E1	Lacks Anxiety	Unconcerned, Unworried, Fearless
	E2	Lacks Pleasure	Pessimistic, Gloomy, Unenthusiastic
	E3	Lacks Emotional Depth	Unemotional, Indifferent, Inexpressive
	E4	Lacks Emotional Stability	Temperamental, Moody, Irritable
	E5	Lack Remorse	Unrepentant, Unapologetic, Unashamed
Self	S1	Self-Centred	Egocentric, Selfish, Self-Absorbed
	S2	Self-Aggrandizing	Self-Important, Conceited, Unenthusiastic
	S3	Sense of Uniqueness	Sense of Being Extraordinary, Exceptional, or Special
	S4	Sense of Entitlement	Demanding, Persistent, Sense of Being Deserving
	S5	Sense of Invulnerability	Sense of Being Invincible, Indestructible, or Unbeatable
	S6	Self-Justifying	Minimizing, Denying, Blaming
	S7	Unstable Self-Concept	Labile, Incomplete, or Chaotic Sense of Self

Note: After Cooke et al. (2012).

The CAPP is a simple and efficient way of mapping the concept of PPD. It specifies what is and what is not PPD according to the literature reviewed and the SMEs interviewed in simple lexical terms, focusing on consensus rather than unanimity (i.e., including symptoms that were controversial, but excluding those that were idiosyncratic). Reliance on the lexical approach avoids the need to specify a set of behaviors that define (extensionally) each symptom, and in particular avoids defining symptoms in terms of official recorded criminality. It also avoids the need to specify the course or timeframe required to define a symptom. Thus, the CAPP model is well-suited to evaluating the conceptual invariance of PPD across groups. In particular, since it avoids technical language (jargon), it should be readily translated into languages other than the original English for use in cross-cultural research.

It is important to emphasize that the CAPP is a theoretical model or concept map of PPD. The CAPP identifies and defines symptoms of PPD, but does not specify operations for determining the presence or severity of PPD symptoms in a particular case. Thus, the CAPP can facilitate research on the concept of PPD. Also, it can be (and has been) used as the basis for developing measures of PPD.

1.2.2. Empirical Support for the CAPP Model

In the following sections, I will focus on content studies related to the CAPP model. Specifically, how the use of CAPP model could promote professionals to understand the representation of PPD through prototypicality analysis will be summarized.

1.2.2.1 Prototypicality Studies

One way to use the CAPP to study the concept of psychopathy is to examine the extent to which people consider various CAPP symptoms to be prototypical of the disorder. As discussed by Kreis (2008), this approach was borrowed from prototype theory (e.g., Broughton, 1990). Thus, researchers can use prototypicality analysis to identify the core symptoms of PPD (as defined using the CAPP), as well as symptoms considered to be “at the boundary, or indeed, outside the boundary of the construct” (Kreis et al., 2012, p. 403). Kreis (2008) developed a set of materials to facilitate prototypicality research using the CAPP. Similar research using the PCL-R was conducted in the past, but was hampered by the fact prototypicality ratings were based only on the names of test

items, as it was not possible to provide participants with copies of the test manual so they could see the full definition of the items (e.g., Rogers et al., 1994).

In the first such study, Kreis, Cooke, Michie, Hoff, and Logan (2012) carried out a prototypicality study using the original English version of the CAPP model. They recruited a sample of 132 mental health professionals and asked them to rate the prototypicality of the 33 CAPP symptoms. To aid interpretation of those ratings, the participants were also asked to rate the prototypicality of 9 symptoms considered to be theoretically irrelevant to PPD, which they called foil symptoms (e.g., shy, considerate, conscientious). Prototypicality ratings were made on a 7-point scale (1 = low, 7 = high). Nearly all the CAPP symptoms were rated as highly prototypical of psychopathy (i.e., average prototypicality rating of 5 or higher out of 7); only three CAPP symptoms were rated as being of medium or low prototypicality (i.e., average ratings lower than 5). In comparison, all the foil symptoms were rated as having low prototypicality (i.e., average ratings lower than 4). A number of subsequent protypicality studies conducted using the English language version of the CAPP with sample of “experts” (i.e., mental health professionals) have found similar results (e.g., Murray, 2014)

Other studies have been conducted to examine whether prototypicality ratings of CAPP symptoms vary across experts versus lay people, or according to the gender of the person with PPD (e.g., Kreis & Cooke, 2012; Kreis et al, 2012; Kreis & Cooke, 2011; Smith et al., 2014). According to Kreis and Cooke (2011), most CAPP symptoms and domains rated as highly representative of psychopathy but Unstable Self-Concept, Lacks Pleasure and Lacks Concentration received the lowest in prototypicality ratings and the Cognitive domain was rated as least prototypical of psychopathy. A study by Smith et al. (2014) also suggested that symptoms contained the Behavioral and Cognitive domains (Lacks Planfulness, Lacks Concentration and Lacks Perseverance) were not considered prototypical of psychopathy with ratings in same range of the foils. Research on gender difference also showed good content validity with both genders. Core symptoms and the construct of psychopathy were similarly exhibited. One of the findings suggested that women with PPD may be more emotionally unstable and manipulative, but less grandiose and physically aggressive and domineering than men with PPD (Kreis & Cooke, 2011; Verona & Vitale, 2006; Forouzan & Cooke, 2005).

1.2.2.2 Cross-Cultural Studies based on The CAPP Model

There have been a number of studies evaluating the validity of the CAPP as a concept map (see Cooke et al., 2012). One common approach has been to attempt to translate the CAPP into languages other than the original English. There are more than 18 translations completed (e.g., Norwegian, French, Lithuanian, Russian, Persian, Sesotho, Hungarian, Hebrew, and so on) or in progress. So far, translators have found comparable networks of the relevant trait descriptive adjectives within the languages considered. Hence, numerous prototypicality studies have been done and compared in several countries using translated versions of CAPP to investigate differences in the expression of psychopathic traits across cultures (Cooke & Logan, 2015; Heinzen, Fittkau, Kreis, & Huchzermeier, 2011; Hoff et al., 2012; Hoff et al., 2014; Sörman et al., 2014; Stoll, Florez, et al., 2014). Nearly identical patterns of prototypicality ratings have been observed in studies (e.g., Flórez, Casas, Kreis, Forti, Martínez, & Fernández, 2014; Hoff, Rypdal, Mykletun, & Cooke, 2012; Stoll, Heinzen, Kohler, & Huchzermeier, 2011).

All the findings among cross-cultural studies have similarly converged on the results that almost all CAPP symptoms were as either highly or very highly prototypical and only a small number of symptoms revealed as being moderately or low in prototypicality. On the other hand, the foil items were rated as being the least prototypical, quoting “the translators have found comparable networks of the relevant trait descriptive adjectives within the languages considered” (Cooke & Logan, 2015, p. 267).

1.2.3. A Measure of PPD Based on the CAPP Model: The CAPP-IRS

In this section, I will describe how the CAPP-IRS was developed based on the CAPP, as well as experience to date using the CAPP-IRS in research.

1.2.3.1 Transition from Model to Measures

The CAPP-IRS is an expert observer ratings scale of PPD symptoms based on the CAPP model. The presence of each symptom is rated with respect to both trait extremity and associated functional impairment (0 = low, 4 = high), as well as with respect to global severity (1 = not present, 7 = extremely severe) on the basis of all available clinical data, including interview, observation, reports from collateral informants, and records. To assist administration, a semi-structured interview schedule and a staff rating form (both optional)

have been developed (see Cooke & Logan, 2015). The timeframe for ratings is flexible and can be set at lifetime or can be shortened to, for example, the past 3, 6, 12, or 24 months.

1.2.3.2 CAPP-IRS: Interrater Reliability and Predictive Validity

To date, research on the CAPP-IRS has been limited but yielded promising results. For example, the CAPP-IRS has been used to make lifetime ratings in a sample of Danish adult forensic psychiatric patients (Pedersen et al, 2010), lifetime ratings in sample of Canadian serious and violent young offenders (e.g., Dawson, McCuish, Hart, & Corrado, 2012; McCormick, Corrado, Hart, & Cohen, 2008), 6-month ratings in a sample of adult forensic psychiatric patients and prisoners in the United Kingdom (Kreis & Cooke, 2011; Cooke et al., 2012), and lifetime ratings in a sample of Norwegian adult correctional offenders (Sandvik et al., 2012). In these studies, the interrater reliability of CAPP-IRS total or domain scores generally was in good to excellent (i.e., ICC1 > .80), although the interrater reliability of individual symptom ratings was typically fair (i.e., ICC1 about .50 to .60).

In all the studies, CAPP-IRS ratings had good concurrent validity vis-à-vis established measures of (lifetime) psychopathic symptoms, such as the Hare Psychopathy Checklist-Revised and the Screening Version or Youth Version of the PCL-R (PCL:SV and PCL:YV, respectively; Hart, Cox, & Hare, 1995; Forth, Kosson, & Hare, 2003). Finally, Pedersen et al. (2010) found that with respect to violent and non-violent recidivism for forensic psychiatric patients, the predictive validity of CAPP-IRS total scores was comparable to that of ratings made using the PCL:SV. The predictive validity of the domain scores was generally similar, although scores on the Cognitive and Self domains were not statistically significant.

1.2.4. CAPP-IRS: Clinical Utility

As Cooke and Logan (2015) pointed out, the CAPP encourages and facilitates an idiographic approach to assessment of PPD. The CAPP model and the CAPP-IRS assist the clinical process of diagnostic and violence risk formulation by providing a clear, detailed, and systematic way to understanding the nature of a client's past, current, and future problems related to PPD symptomatology (see Cooke et al., 2012).

In the Korean justice system, almost all of the validated PPD assessments and related structured forensic measures were actually designed based on nomothetic approach, such as the PPI-R, PCL-R, and their derivatives. While there is no doubt that these instruments have hitherto assisted clinicians to understand individuals with PPD in terms of norm-based reference group, it has also a limitation in making specific management plans and intervention strategies for the violence. For example, peoples who scored over 27 on the PCL-R are, on average, more likely to be violent than those who scored under 7 with respect to the mean level of psychopathy. But, this is also true that this could not explain at the individual level (Cooke & Michie, 2010). Currently, Korean psychological treatment centers in the Correctional Service were mainly established for the intensive treatment programs targeted at chronically sexual offenders and criminal psychopaths by the criminal law. So, beyond just typical process of the classification of psychopathic offenders in prisons, this may lead to more attention on comprehensive management and guidance at individuals with PPD. With this respect, the CAPP-IRS could be a promising clinical instrument in light of comprehensiveness, emphasis on symptom-change and clinically management plan.

1.3. The Current Study

To date, cross-cultural research on PPD in Korea has relied heavily on the PCL-R. But recently, I translated the CAPP model and CAPP-IRS measure into Korean—referred to below as the K-CAPP and K-CAPP-IRS—and evaluated their cross-cultural validity in Korea in two separate studies.

In Study 1, I conducted a prototypicality study using the K-CAPP to determine whether Koreans conceptualized PPD similarly to people from other cultures. I examined prototypicality ratings by samples of experts and lay people from Korea, comparing their ratings for K-CAPP symptoms to those for foil symptoms; evaluating the similarity of ratings for K-CAPP symptoms and foils made by experts versus lay people; and comparing the ratings for K-CAPP symptoms and foils made by experts and lay people in Korea to those by experts and lay people in other countries using the original English CAPP or translations of the CAPP into other languages. I expected that the prototypicality ratings made using the K-CAPP would be higher for symptoms of PPD than for foil symptoms, both in experts and in lay people. Furthermore, I expected that the prototypicality ratings of PPD and foil symptoms made by experts versus lay people using

the K-CAPP would be very similar to each other, and also very similar to those made by lay people and experts using the original English language CAPP or translations of it into other languages.

In Study 2, I evaluated the interrater reliability of the K-CAPP-IRS in a sample of adult male correctional offenders in Korea, each of whom was evaluated by a team of three experts on two different occasions, about 3 months apart. I examined the interrater reliability of total, domain, and symptom ratings using intraclass correlation coefficients, and also within the framework of Generalizability (G) theory. I expected that the interrater reliability of K-CAPP-IRS total and domain ratings would be good to excellent on average, whereas the interrater reliability of symptom ratings would be fair to good on average.

In Study 2, I also evaluated the concurrent validity of K-CAPP-IRS total and domain ratings with respect to the Korean translations of the PCL-R (K-PCL-R; Joe & Lee, 2008) and the Psychopathic Personality Inventory-Revised (K-PPI-R; Park & Lee, 2009). I expected to find a very large correlation ($r >> .50$) between K-CAPP-IRS and K-PCL-R total scores. I also expected large correlations ($r > .50$) between corresponding K-CAPP-IRS domains scores and K-PCL-R facet scores, specifically between: K-CAPP-IRS Dominance and Self domains and K-PCL-R Facet 1; K-CAPP-IRS Attachment and Emotional domains and K-PCL-R Facet 2; and K-CAPP-IRS Behavioural and Cognitive scores and K-PCL-R Facet 3. Finally, I expected to find a moderate to large correlation ($.30 < r < .50$) between K-CAPP-IRS and K-PPI-R total scores, and moderate correlations ($r > .30$) between K-CAPP-IRS domains and K-PPI-R factors.

Chapter 2. Study 1.

2.1. Method

2.1.1. Ethics and Protection of Human Subjects

Initial ethics approval for Study 1 was obtained from KNCS and Ajou University, who collaborated on the collection of data in Korea under the direction of SJH. Once the data were collected, ethics approval was sought from Simon Fraser University to undertake secondary analysis of the data as part of SJH's dissertation research.

2.1.2. Participants

Participants comprised two samples: lay people (i.e., community residents) and experts (e.g., clinical forensic mental health professionals) on the basis of their experience and knowledge in dealing with PPD.

The lay people were 618 community residents of various cities in Korea (e.g., Suwon, Seoul, Kyungnam, Daegu) who had no specialized education, training, or experience related to PPD. They were recruited from three universities (n=470) and five lifelong education centers (n=148) in Korea through web based on-line survey. They were excluded from participation if they indicated that they worked, or had worked, in mental health or criminal justice settings or were enrolled in mental health or criminal justice training programs. Only small proportion of the lay people who were approached refused to participate in the current study (i.e., refusal rate of 12%).

The experts were 88 employees of the Korean National Forensic Service, Police Agency, Forensic Mental Hospital, and Corrections Service (the latter referred to as KNCS hereinafter) recruited at their workplaces, or members of the Korean Psychological Association (referred to as KPA hereinafter) recruited at professional conferences. Experts received information about the study through their workplaces or at professional conferences and volunteered to participate. All experts had a high level of training (e.g., PCL-R, PCL:SV, ICD, DSM, actuarial instruments) and experience working with offenders suffering from personality disorders in clinical-forensic settings. Only 12 experts who were

approached refused to participate (i.e., refusal rate of 12%). Table 2 provides a summary of the demographic characteristics of participants in the two groups.

Table 2. Demographic Characteristics of Experts and Lay People

	Experts (*N* = 88)	Lay People (*N* = *618)*
Gender, *n* (%) male	42 (47.7%)	246 (39.8%)
Age, *M* (*SD*) years	40.15 (7.86)	39.89 (8.91)
Education, *n* (%) Masters/Doctoral	78 (88.6%)	36 (5.8%)
Marital Status, *n* (%) never married	43 (48.8%)	456 (73.7%)
Background of Experts		
Occupation, *n* (%) psychologist	68 (77.2%)	--
Work experience, *M* (*SD*) months	70.35 (57.39)	--
Contact with PPD, *n* (%) at least weekly	68 (77.2%)	--
Background of Lay People	--	
Occupation, *n* (%) students	--	470 (76%)

2.1.3. Procedure

Participants were directed to a link to a secure, online survey administered anonymously via a Korean equivalent of SurveyMonkey. The survey was anonymous.

2.1.4. Instruments

2.1.4.1 Short Demographic Questionnaire

Participants were asked to provide information concerning their age, gender, marital status, and level of education; as well as their occupation, work experience, and frequency of contact with offenders or patients with PPD at work.

2.1.4.2 Prototypicality Ratings

Next, participants made prototypicality ratings of the 33 CAPP symptoms and 9 "foil" symptoms based on a Korean translation of the CAPP, or K-CAPP. The K-CAPP was developed using a translation procedure similar to that described by Hoff et al. (2014, p. 358) and Sellbom et al. (2016, p. 439):

1. Consultation with the CAPP-IRS developers. I (SJH) met with one of the CAPP developers (SDH), who explained the CAPP model in detail and provided additional

explanation of the words and phrases selected to represent symptoms, as well as the adjectival descriptors of the symptoms.

2. Meeting with the translation team. SJH convened a meeting of the full translation team—comprising two forensic psychologists, two criminal behavior analysts, two professional translators, and three lay people bilingual in English and Korean (undergraduate students with no experience working in forensic settings and no knowledge of the CAPP)—at which the team generated the first draft of the K-CAPP symptoms and adjectival descriptors. This meeting helped to identify areas of potential discrepancy between the English and Korean versions.

3. Consultation with other psychology experts. SJH consulted with one social psychologist and one linguistic psychologist at Korean universities, neither of whom had forensic training or experience. The consultants reviewed the first draft of the translation and provided feedback.

4. Expert linguistic advice. A university-based philologist (Dr. Chung) provided linguistic feedback concerning the first draft of the translation, focusing on the equivalence of the original English and corresponding Korean terms as well as the familiarity and clarity of the terms in Korean.

5. Translation team consensus. Taking into account the feedback from Steps 3 and 4, the translation team prepared a second draft of the translation.

6. Back-translation. The second draft of the translation was back-translated by three professional translators, working independently. The back-translation was reviewed by SJH and SDH, who concluded that two symptoms (Sense of Invulnerability and Sense of Entitlement) were not back-translated consistently, and therefore made a first round of minor revisions. This process was repeated for a second round of revisions, with back-translation by three different professional translators, at which time SJH and SDH were satisfied that the back-translations were consistent with each other and accurately reflected the intended denotative and connotative meaning of the various symptoms and adjectival descriptors.

The end result of the translation process, the K-CAPP, was used as the basis for developing a prototypicality rating form following the Universal Protocol for Conducting

Prototypicality Studies with the CAPP (see Kreis, 2008; Kreis et al., 2012). Participants reviewed 42 symptoms—the 33 K-CAPP symptoms and 9 foil symptoms, each defined by 3 adjectival descriptors—and rated the extent to which they considered each to be prototypical of people with PPD using a 7-point Likert-type scale (1 = low, 7 = high).

2.1.5. Data Analysis

Data analyses were performed using SPSS version 20.0 version for Windows. First, I calculated descriptive statistics (means and standard deviations) for ratings made by lay people versus experts. Following past research (e.g., Hoff et al., 2012; Florez et al., 2015; Rogers et al., 1992), mean ratings between 4 and 5 were considered moderately prototypical, and mean ratings of 5 or greater were considered highly prototypical. Second, I compared the mean prototypicality ratings for 33 K-CAPP symptoms to those for foil symptoms for lay people versus experts, using t-tests. Finally, I examined the correlation the mean prototypicality ratings made by lay people and experts in Korea versus those made by lay people and experts in other countries, using the original English version of the CAPP or translations of it into other languages.

2.2. Results

2.2.1. Prototypicality Ratings of the K-CAPP

The CAPP-IRS domains and total score descriptive are shown in Table 3. With the exception of the Behavioural domain, significantly positively skewed, all CAPP–IRS domain scores were distributed evenly. The CAPP total scores ranged from a minimum of 2 to a maximum of 158 with a mean value of 78 (SD=34.12). The total scores were distributed evenly in the sample.

Table 3. Descriptive of the K-CAPP Domain Scores

Domains	*M*	*Mdn*	*SD*	*Skew (SE)*	*Kurtosis (SE)*	*Min.*	*Max.*
Attachment	10.46	9	6.11	-.064 (.26)	-1.099 (.52)	.00	22.00
Behavioural	14.58	14	7.84	-.153 (.26)	-.595 (.52)	.00	30.00
Cognitive	13.61	12	6.87	.100 (.26)	-.829 (.52)	.00	28.00
Dominance	12.34	12	6.81	.119 (.26)	-.478 (.52)	.00	30.00
Emotional	12.96	13	6.34	.010 (.26)	-.450 (.52)	.00	27.00
Self	14.90	13	7.96	.114 (.26)	-.865 (.52)	.00	33.00
Total	78.87	76	34.12	-.191 (.27)	-.184 (.54)	2.00	158.00

2.2.2. Prototypicality Ratings for K-CAPP and Foil Symptoms

There was considerable variability in the mean prototypicality ratings for K-CAPP and Foil symptoms (see Table 4). In both experts and lay people, the majority of the 33 K-CAPP symptoms were rated as moderately or highly prototypical of psychopathy, which is in line with previous prototypicality literatures (Florez et al., 2015; Hoff, Rypdal, Mykletun, & Cooke, 2012; Kreis, 2008; Kreis & Cooke, 2011; Kreis, Cooke, Michie, Hoff, & Logan, 2012; Robinson, 2017; Sörman et al., 2014). For the Attachment domain, all four symptoms (Undetached, Unempathic, Uncommitted and Uncaring) were considered as highly prototypical by both groups. For the Behavioral domain, two symptoms (Lacks of Perseverance and Restless) were rated as low prototypicality by both groups. Four Behavioural symptoms (Lacks Perseverance, Reckless, Restless, and Unreliable) were rated as significantly more prototypical by experts than by lay people. For the Cognitive domain, three symptoms (Intolerance, Suspicious, and Inflexible) were rated at least moderately prototypical by both groups, and two symptoms (Lacks Concentration and Lacks Planfulness) were rated as low prototypicality by both groups. Three symptoms (Lacks Concentration, Lacks Planfulness, and Intolerant) were rated as significantly more prototypical by experts than by lay people. For the Dominance domain, five symptoms (e.g., Antagonistic, Domineering, Deceitful, Manipulative, and Insincere) were rated as highly prototypical by both groups. One symptom (Garrulous) was rated as very low in prototypicality (i.e., less than 3) by lay people, but as moderately prototypical by experts. Three symptoms (Antagonistic, Domineering, and Garrulous) were rated as significantly more prototypical by experts than by lay people. In the Emotional domain, two symptoms (Lacks Emotional Depth and Lacks Remorse) were rated as highly prototypical by both groups, and another symptom (Lacks Anxiety) was rated as highly prototypical by lay people but only moderately prototypical by experts. Finally, four Self domain symptoms (Self-centered, Self-justifying, Sense of Uniqueness, and Sense of Entitlement) were rated

as highly prototypical by both groups; one symptom (Sense of Invulnerability) was rated as highly prototypical by experts but only moderately prototypical by lay people; and the remaining two symptoms (Self-aggrandizing and Unstable Self-concept) were rated as moderately prototypical by both groups.

Of the foil symptoms, most (seven of nine) were rated as very low in prototypicality by both groups (see Table 4). The exceptions were Perfectionistic and Strange, which were rated as moderately prototypical by both groups.

2.2.2.1 Experts versus Lay People in Kor.

Table 4. Distribution (M, SD) of CAPP symptoms, foils, and domains in experts and lay people

(K-)CAPP	Experts		Lay People	
Symptoms	*M*	*SD*	*M*	*SD*
Unempathic	6.17	1.53	6.19	1.40
Self-Centered***	6.10	1.20	5.63	1.52
Lacks Emotional Depth	6.07	1.54	5.87	1.54
Lack Remorse	6.04	1.58	6.12	1.52
Self-Justifying***	5.82	1.38	5.08	1.77
Uncaring	5.70	1.50	5.58	1.53
Antagonistic***	5.69	1.41	5.00	1.69
Manipulative	5.69	1.52	5.33	1.66
Domineering***	5.65	1.36	5.05	1.68
Deceitful	5.60	1.57	5.26	1.68
Detached***	5.54	1.37	4.76	1.68
Aggressive	5.51	1.21	5.32	1.57
Sense of Uniqueness	5.50	1.44	5.09	1.70
Insincere***	5.42	1.67	4.92	1.56
Sense of Entitlement	5.38	1.43	5.09	1.74
Unreliable*	5.26	1.58	4.56	1.77
Sense of Invulnerability***	5.22	1.44	4.50	1.82
Uncommitted	5.14	1.64	5.00	1.63
Suspicious	5.00	1.34	4.72	1.58
Disruptive	4.93	1.77	4.62	1.82
Lacks Anxiety	4.93	1.65	5.00	1.78
Self-Aggrandizing***	4.77	1.62	4.00	1.76
Intolerant**	4.62	1.48	4.02	1.71
Lacks Emotional Stability	4.61	1.66	4.82	1.82
Reckless***	4.55	1.77	3.73	1.83
Unstable Self-Concept	4.52	1.62	4.55	1.80
Garrulous***	4.51	1.59	2.80	1.53
Inflexible	4.28	1.61	4.23	1.64
*Perfectionist****	*4.22*	*1.60*	*4.87*	*1.58*
*Strange***	*4.13*	*1.74*	*4.80*	*1.67*
Lacks Pleasure	4.01	1.79	4.05	1.76
Lacks Perseverance***	3.96	1.75	3.11	1.60

Restless*	3.70	1.73	3.28	1.58
Restrained	*3.62*	*1.90*	*3.84*	*1.97*
Lacks Concentration**	3.54	1.56	3.03	1.55
Lacks Planfulness***	3.54	1.79	2.48	1.54
*Dependent***	*3.38*	*1.60*	*2.76*	*1.65*
*Cautious***	*3.34*	*1.74*	*3.89*	*1.79*
Shy	*2.93*	*1.65*	*3.09*	*1.61*
*Self-Conscious***	*2.75*	*1.51*	*3.09*	*1.61*
Considerate	*2.03*	*1.51*	*2.28*	*1.60*
Conscientious	*1.93*	*1.64*	*2.10*	*1.51*

Note *p < .05; **p < .01; ***p < .001; italics for nine foils.

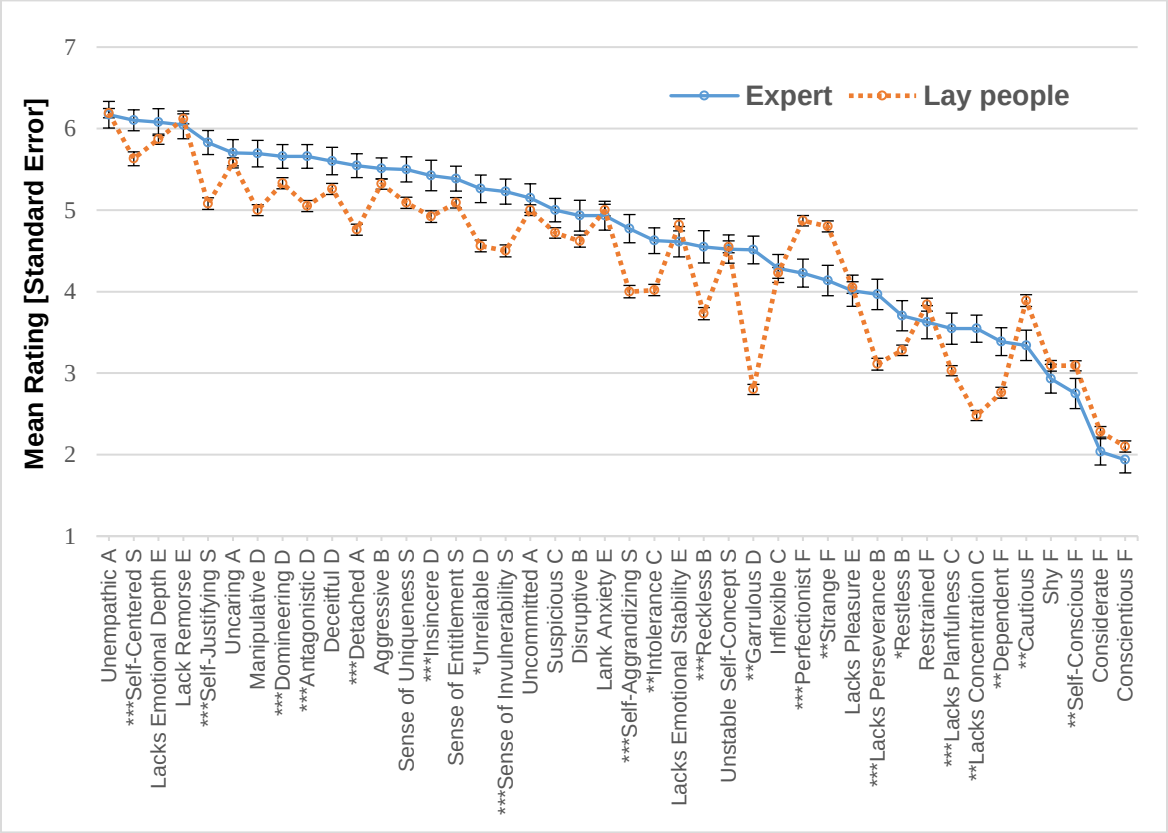

Figure 2. Prototypical evaluations of K-CAPP symptoms and foils between expert group and lay people

Note. *p < .05. **p < .001. ***p < .000.

A series of 42 independent t-tests was conducted to identify significant differences in the prototypicality ratings for symptoms made by the two groups (see Figure 2). A significance level of .05 was used to reject the null hypothesis and conclude there were

significant differences in rating between the two groups. The findings indicated significant differences in sixteen items for two CAPP symptoms were found.

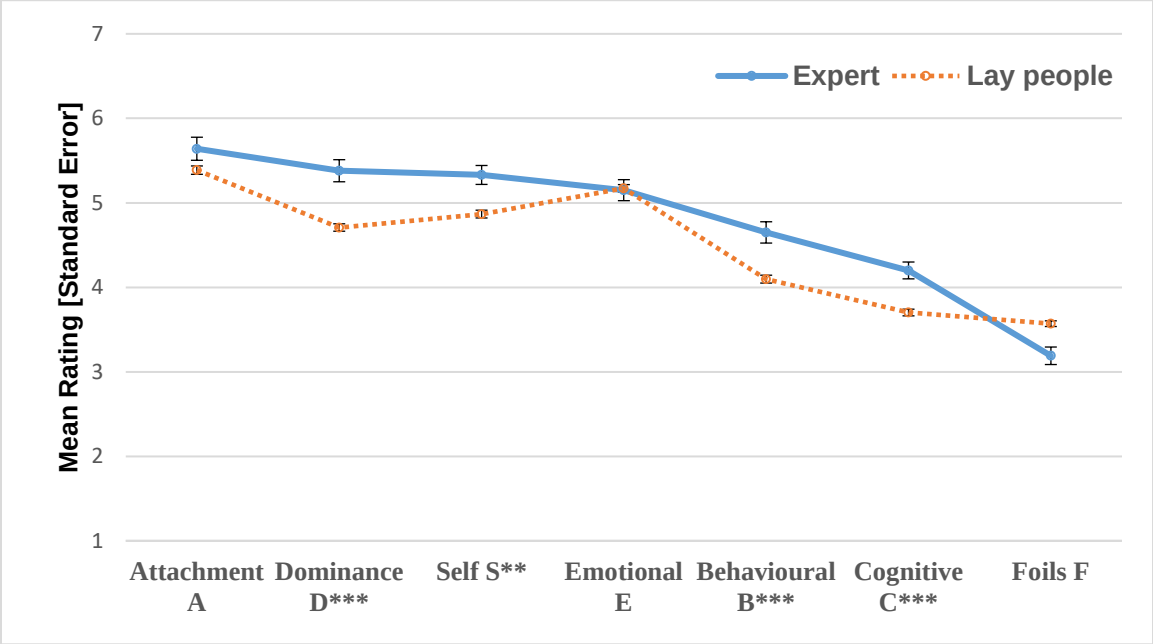

Figure 3. Prototypical evaluations of (K-CAPP) domains and foils between expert group and lay people

Note. *p < .05. **p < .001. ***p < .000.

Variability was found in the ratings of all the CAPP domains after analyzing the means and standard deviations (see Figure 3). Four domains (e.g., Attachment, Dominance, Self, and Emotional) in expert and two domains (Attachment, Emotional) in lay people were rated the highly prototypical of psychopathy. Two domains (e.g., Behavioural and Cognitive) in expert and three domains (Dominance, Self, Behavioural) were rated as moderately prototypical among two groups. Finally, the foils were rated as least prototypical of psychopathy among the two groups. Independent t-tests were conducted to identify if significant differences exist in the ratings of the domains between the two groups. The results indicate that significant differences did exist between the ratings for four domains (Dominance, Self, Behavioral, Cognitive).

2.2.2.2 Lay People in Kor. versus Lay People in Europe

Table 5. Mean and Standard Deviation of CAPP symptoms, foils, and domains among lay people in three countries

Domains (*n*)		*Mean*			*S.D.*		
	Symptoms	Korea	Spain [a]	Norway [b]	Korea	Spain	Norway

Attachment (4)	5.38	5.43	5.25	1.22	1.11	1.13
Detached	4.76	5.54	4.98	1.68	1.55	1.68
Uncommitted	5.00	5.14	4.92	1.63	1.64	1.58
Unempathic	6.10	5.84	6.06	1.40	1.53	1.38
Uncaring	5.58	5.19	5.07	1.53	1.64	1.61
Behavioural (6)	4.09	4.24	4.00	1.11	0.74	1.04
Lacks Perseverance	3.11	3.35	2.71	1.60	1.90	1.60
Unreliable	4.56	4.57	4.62	1.77	1.83	1.76
Reckless	3.73	3.49	3.49	1.83	1.84	1.77
Restless	3.28	4.31	3.81	1.58	1.79	1.75
Disruptive	4.62	4.44	4.03	1.88	1.84	1.69
Aggressive	5.32	5.29	5.33	1.57	1.58	1.49
Cognitive (5)	3.70	4.49	4.36	0.97	0.90	0.91
Suspicious	4.72	5.76	5.39	1.58	1.27	1.47
Lacks Concentration	3.03	3.14	3.21	1.55	1.87	1.74
Intolerance	4.02	5.25	5.17	1.71	1.48	1.50
Inflexible	4.23	5.22	5.27	1.64	1.44	1.47
Lacks Planfulness	2.48	3.09	2.75	1.54	1.96	1.67
Dominance (6)	4.70	5.02	5.40	1.07	0.93	0.88
Antagonistic	4.99	4.65	4.78	1.69	1.68	1.59
Domineering	5.05	5.50	5.95	1.68	1.59	1.33
Deceitful	5.26	5.54	5.58	1.68	1.58	1.41
Manipulative	5.33	5.94	6.37	1.66	1.36	1.11
Insincere	4.92	5.15	4.99	1.74	1.53	1.67
Garrulous	2.80	3.35	4.74	1.53	1.99	1.74
Emotional (5)	5.17	4.96	4.99	1.07	0.94	0.88
Lacks Anxiety	5.00	3.89	4.46	1.78	1.90	1.77
Lacks Pleasure	4.05	4.20	3.31	1.76	1.98	1.73
Lacks Emotional Depth	5.87	5.37	5.43	1.54	1.76	1.72
Lacks Emotional Stability	4.82	5.52	5.73	1.82	1.65	1.38
Lack Remorse	6.12	5.84	6.02	1.52	1.67	1.37
Self (7)	4.86	5.10	5.38	1.13	0.91	0.87
Self-Centered[4]	5.63	5.80	6.07	1.43	1.35	1.25
Self-Aggrandizing	3.96	4.54	5.22	1.76	1.75	1.59
Sense of Uniqueness	5.09	5.51	5.55	1.70	1.60	1.55
Sense of Entitlement	5.25	5.40	5.34	1.58	1.47	1.53
Sense of Invulnerability	4.50	5.31	5.19	1.80	1.68	1.57
Self-Justifying	5.08	5.28	5.66	1.77	1.60	1.52
Unstable Self-Concept	4.55	3.88	4.66	1.80	1.80	1.79
Foils (9)						
Dependent	2.74	3.28	2.81	1.65	1.95	1.76
Perfectionist	4.87	4.84	4.92	1.58	1.86	1.64
Conscientious	2.10	3.81	2.31	1.51	2.00	1.55
Considerate	2.28	2.28	2.31	1.60	1.70	1.54
Strange	4.80	4.59	4.04	1.67	1.90	1.84
Restrained	3.84	3.70	3.69	1.97	1.86	1.82
Shy	3.09	4.50	2.69	1.61	1.83	1.64
Cautious	3.89	3.31	2.32	1.79	2.32	1.52
Self-Conscious	3.09	3.30	2.73	1.68	1.81	1.66

Note. a The same data (lay people in Spain) from Florez et al. (2012) for the comparison; b the same data (lay people in Norway) from Hoff et al. (2012) data for direct comparison.

Shown in Figure 4, on the whole, the CAPP symptoms across three countries were typically rated above 4 and the foils were generally rated below 4. Among the three countries, the four symptoms rated around 6 as very highly prototypical were Unempathic, Lacks remorse, Self-centered, and Manipulative, whereas three symptoms (Lacks concentration, Lacks planfulness and Lacks perseverance) were rated below 4 as moderately prototypical. Compared to the CAPP symptoms, five foils (Cautious, Dependent, Self-conscious, Considerate and Conscientious) across three countries were rated below 4, whereas two symptoms (Perfectionist and Strange) were rated as moderately prototypical. In Korean sample, Lacks emotional depth was rated as very highly prototypical above 6. Generally, in terms of lay people rating on the CAPP symptoms, there is strongly similar prototypical pattern across the three countries (regardless of the location of the compared countries in either the West or the East).

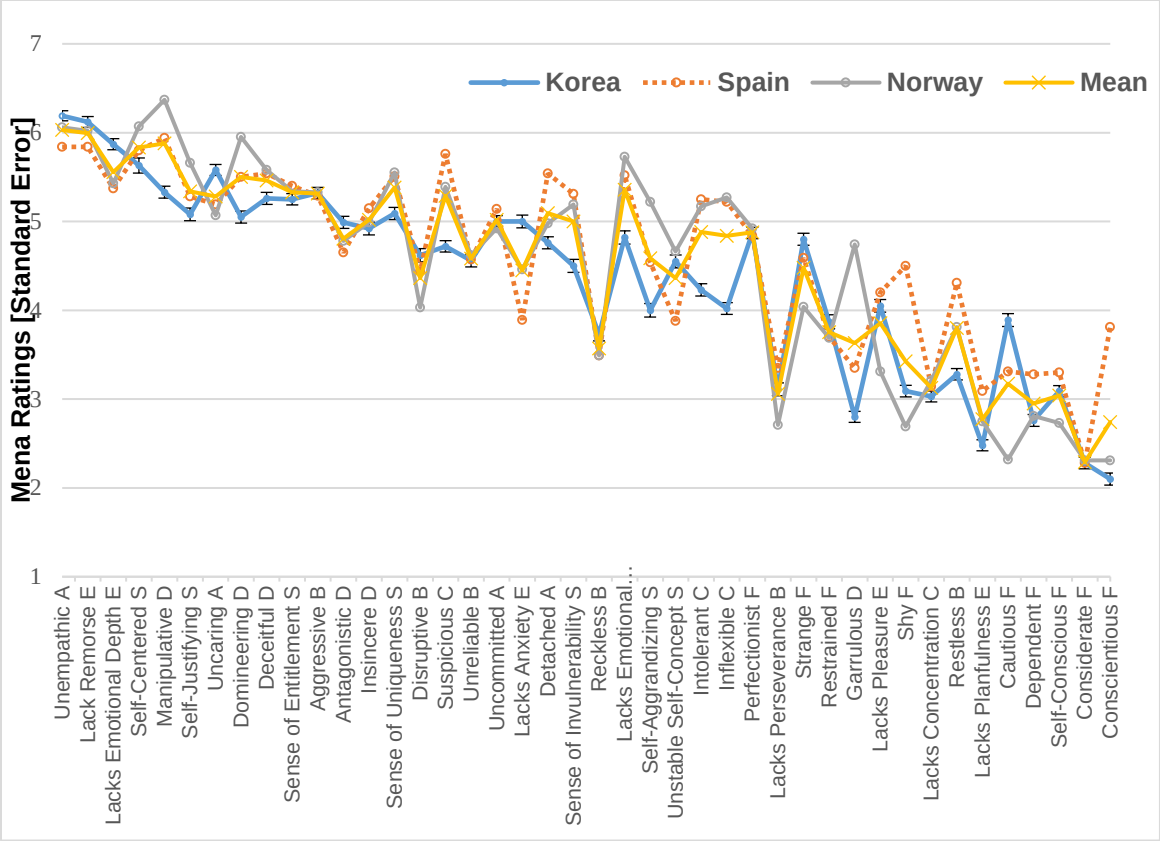

Figure 4. Prototypical evaluations of CAPP symptoms and foils among lay people in three countries.

Note. The same data (lay people in Spain) from Florez et al. (2012) for the comparison; b the same data (lay people in Norway) from Hoff et al. (2012) data for direct comparison under the permission of the data by the first-authors

Table 6. Pearson's r Correlations between Lay People's Prototypicality Studies

	Korea	Spain	Norway
Spain	0.83	-	
Norway	0.84	0.88	-
Total	0.93	0.95	0.96

Note. All correlations sig. at p<.001 level (2-tailed); The same data (lay people in Spain) from Florez et al. (2012) for the comparison; the same data (lay people in Norway) from Hoff et al. (2012) data for direct comparison under the permission of the data by the first-authors.

I investigated the correlations between the mean prototypicality ratings of the CAPP and foils made by lay people in the present study to those made by lay people in the previous studies. In summary, the correlations were all strongly large, r ranging from .93 to .96 in total level (all, $p < .001$) (see Table 6). In summary, these results clearly

showed that lay people in the current study conceptualized PPD in almost same way as that of lay people in previous research using the CAPP.

2.2.2.3 Kor. Experts versus European Experts

Table 7. Mean and Standard Deviation of CAPP Symptoms, Foils, and Domains among Experts in the three countries

Domains / Symptoms	*Mean (SD)*			
	Korea	Sweden[a]	Spain[b]	Norway[c]
Attachment (4)	5.77 (0.99)	5.32 (1.01)	5.74 (0.85)	5.75 (0.98)
Detached	5.22 (1.39)	4.97 (1.65)	5.47 (1.43)	5.25 (1.58)
Uncommitted	5.25 (1.57)	5.27 (1.40)	5.78 (1.21)	5.87 (1.23)
Unempathic	6.37 (1.04)	6.08 (1.43)	5.86 (1.50)	6.27 (1.07)
Uncaring	5.88 (1.22)	4.96 (1.56)	5.85 (1.19)	5.59 (1.44)
Behavioural (6)	4.94 (0.96)	4.82 (0.90)	4.74 (0.87)	4.97 (0.95)
Lacks Perseverance	4.45 (1.53)	4.22 (1.55)	4.27 (1.66)	4.20 (1.59)
Unreliable	5.29 (1.57)	6.01 (1.34)	5.21 (1.53)	5.57 (1.52)
Reckless	5.13 (1.53)	5.74 (1.39)	4.50 (1.77).	4.54 (1.69)
Restless	3.62 (1.56)	4.10 (1.80)	4.18 (1.64)	4.55 (1.59)
Disruptive	5.45 (1.47)	4.23 (1.59)	4.98 (1.64)	5.33 (1.44)
Aggressive	5.75 (1.17)	5.23 (1.35)	5.52 (1.23)	5.65 (1.30)
Cognitive (5)	4.38 (0.93)	4.10 (0.93)	4.43 (0.87)	5.01 (0.95)
Suspicious	5.34 (1.30)	4.36 (1.61)	4.83 (1.46)	5.50 (1.45)
Lacks Concentration	3.70 (1.60)	3.29 (1.67)	3.17 (1.49)	4.04 (1.67)
Intolerant	4.72 (1.56)	4.78 (1.47)	5.20 (1.42)	5.59 (1.44)
Inflexible	4.60 (1.65)	4.51 (1.61)	5.11 (1.48)	5.53 (1.43)
Lacks Planfulness	3.56 (1.80)	3.48 (1.82)	3.83 (1.79)	4.39 (0.84)
Dominance (6)	5.47 (0.93)	5.73 (0.64)	5.50 (0.80)	5.71 (0.84)
Antagonistic	5.73 (1.19)	3.92 (1.42)	5.26 (1.28)	5.32 (1.50)
Domineering	5.88 (1.15)	5.80 (0.94)	5.89 (0.99)	5.99 (1.24)
Deceitful	5.83 (1.32)	6.26 (0.88)	5.70 (1.35)	5.62 (1.45)
Manipulative	6.13 (1.15)	6.63 (0.71)	6.21 (0.95)	6.34 (0.89)
Insincere	5.50 (1.42)	6.07 (1.15)	5.51 (1.30)	5.57 (1.39)
Garrulous	3.94 (1.60)	5.68 (1.36)	4.45 (1.62)	5.42 (1.53)
Emotional (5)	5.35 (0.73)	5.10 (0.65)	5.03 (0.74)	5.27 (0.79)
Lacks Anxiety	5.24 (1.64)	5.53 (1.32)	4.44 (1.59)	4.75 (1.60)
Lacks Pleasure	3.87 (1.68)	2.43 (1.59)	3.98 (1.87)	4.00 (1.74)
Lacks Emotional Depth	6.26 (1.00)	5.87 (1.44)	5.55 (1.41)	5.45 (1.63)
Lacks Emotional Stability	5.12 (1.58)	5.19 (1.36)	4.80 (1.63)	5.85 (1.30)
Lack Remorse	6.32 (1.11)	6.42 (0.94).	6.41 (0.85)	6.19 (1.14)
Self (7)	5.52 (0.77)	5.67 (0.66)	5.32 (0.71)	5.72 (0.77)
Self-Centered	6.22 (0.93)	6.17 (0.93)	6.14 (0.91)	6.36 (0.88)
Self-Aggrandizing	5.00 (1.54)	6.14 (1.05)	5.24 (1.21)	5.82 (1.37)
Sense of Uniqueness	5.50 (1.32)	6.03 (1.12)	5.88 (1.18)	6.39 (1.32)
Sense of Entitlement	5.81 (1.02)	5.83 (0.93)	5.89 (1.06)	5.98 (1.15)
Sense of Invulnerability	5.19 (1.44)	5.99 (0.99)	5.25 (1.35)	5.47 (1.49)
Self-Justifying	6.05 (1.13)	6.13 (1.23)	5.88 (1.27)	6.39 (0.90)
Unstable Self-Concept	4.83 (1.59)	3.37 (1.66)	3.14 (1.69)	4.07 (1.93)
Foils (9)				

Dependent	3.44 (1.79)	2.55 (1.73)	2.43 (1.72)	3.47 (1.86)
Perfectionist	4.56 (1.56)	3.56 (1.81)	3.27 (1.57)	4.30 (1.88)
Conscientious	1.84 (1.41)	2.40 (1.57)	2.90 (1.62)	2.10 (1.33)
Considerate	1.89 (1.43)	2.04 (1.44)	2.13 (1.47)	2.37 (1.57)
Strange	4.29 (1.73)	2.90 (1.73)	2.91 (1.53)	4.22 (1.76)
Restrained	4.02 (1.86)	2.50 (1.60)	2.50 (1.54)	3.19 (1.70)
Shy	3.81 (1.76)	1.64 (1.13)	2.50 (1.56)	2.12 (1.38)
Cautious	3.55 (1.78)	1.78 (1.20)	2.54 (1.58)	2.26 (1.43)
Self-Conscious	3.24 (1.63)	1.70 (1.06)	2.09 (1.32)	2.60 (1.69)

Note. Comparison of mean prototypicality ratings across CAPP studies. a The Sörman et al. (2014) data are for the forensic evaluator sample (n = 90); b the Florez et al. (2012) data are for the comparison is expert group (n = 187); c the Hoff et al. (2012) data are for the forensic psychiatric sample (n = 211)

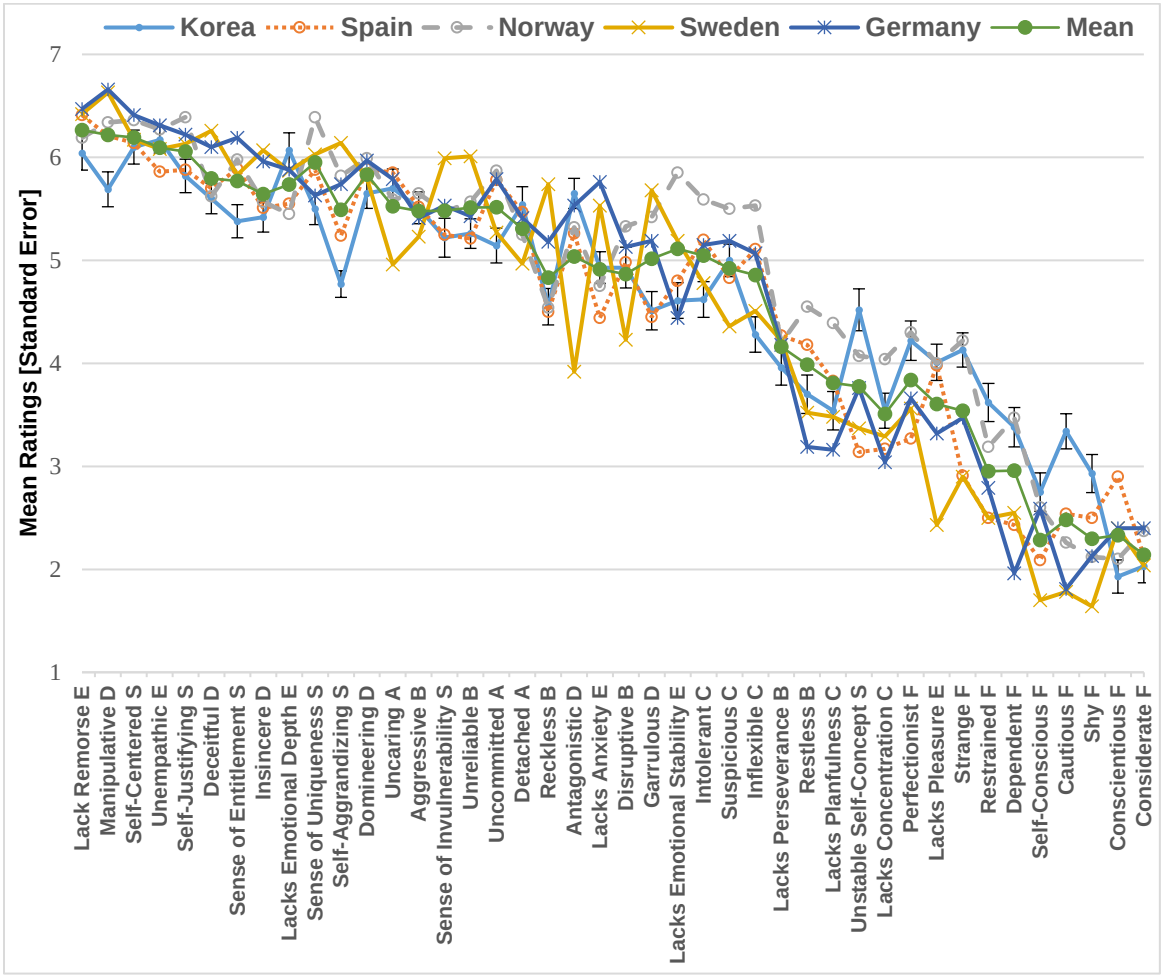

Figure 5. Prototypicality of CAPP Symptoms and Foils Among Experts in European Countries (Spain, Norway, Sweden, Germany)

Note. Comparison of mean prototypicality ratings across CAPP studies. a The Sörman et al. (2014) data are for the forensic evaluator sample (n = 90); b the Florez et al. (2012) data are for the comparison is expert group (n = 187); c the Hoff et al. (2012) data are for the forensic psychiatric sample (n = 211); Hanna (2018, no published) has German forensic experts.

Among the three European countries showed in Figure 5, five CAPP symptoms (Unempathic, Lacks Remorse, Self-centered, Manipulative and Self-justifying) were commonly rated as very highly prototypical, whereas Lacks concentration, Lacks perseverance, and Unstable self-conception were rated as least prototypical. Even the four foils (Shy, Cautious, Dependent, Self-conscious, Considerate, and Conscientious) were rated below 4 as least prototypical and one symptom (Perfectionist) was somewhat mixed (Korea and Norway rated above 4 versus Spain and Sweden rated below 4). One symptom (Unstable self-concept) was rated somewhat higher than in other countries while three symptoms (Intolerant, Lacks Planfulness, and Inflexible) in Cognitive domain, one symptom (Garrulous) in Dominance and two symptoms (Restless) in Behavioural domain were rated somewhat lower than in the other countries. Also, compared to the Western countries, Lacks Emotional Depth was rated as very highly prototypical. For foils, one foil (e.g., Cautious) was rated more prototypical in Korea than in Sweden, Spain and Norway but all of the symptoms were rated below 4 as least prototypical.

2.2.2.4 Kor. Experts versus Experts in Internationals

Table 8. Mean and Stand Deviation of CAPP symptoms, foils, and domains among experts in International samples

Domains / Symptoms	Prototypicality Rating Studies			
	Korea	Lim et al. (2014) [a]	Murray (2014) [b]	Kreis et al. (2012) [c]
Attachment (4)				
Detached	5.22 (1.39)	5.62	4.74	5.64 (1.32)
Uncommitted	5.25 (1.57)	5.23	5.04	5.37 (1.35)
Unempathic	6.37 (1.04)	6.29	6.37	6.49 (0.85)
Uncaring	5.88 (1.22)	5.90	5.46	5.93 (0.98)
Behavioural (6)				
Lacks Perseverance	4.45 (1.53)	4.17	3.08	4.83 (1.39)
Unreliable	5.29 (1.57)	5.62	4.57	5.99 (5.99)
Reckless	5.13 (1.53)	5.77	5.13	5.75 (1.07)
Restless	3.62 (1.56)	3.90	2.87	5.16 (1.42)
Disruptive	5.45 (1.47)	5.14	4.50	5.58 (1.04)
Aggressive	5.75 (1.17)	5.81	4.99	5.64 (1.10)
Cognitive (5)				
Suspicious	5.34 (1.30)	4.74	3.94	4.83 (1.33)
Lacks Concentration	3.70 (1.60)	3.78	2.08	3.84 (1.48)
Intolerant	4.72 (1.56)	4.38	3.66	5.40 (1.18)
Inflexible	4.60 (1.65)	4.53	3.99	4.60 (1.49)
Lacks Planfulness	3.56 (1.80)	4.05	2.61	5.00 (1.46)
Dominance (6)				
Antagonistic	5.73 (1.19)	5.26	5.02	5.19 (1.35)
Domineering	5.88 (1.15)	5.45	5.04	5.72 (1.13)

Deceitful	5.83 (1.32)	6.34	6.44	6.27 (0.73)
Manipulative	6.13 (1.15)	6.38	6.35	6.44 (0.74)
Insincere	5.50 (1.42)	6.30	6.28	6.26 (0.83)
Garrulous	3.94 (1.60)	5.44	4.85	4.50 (1.51)
Emotional (5)				
Lacks Anxiety	5.24 (1.64)	5.22	5.27	5.11 (1.55)
Lacks Pleasure	3.87 (1.68)	3.07	1.31	3.22 (1.59)
Lacks Emotional Depth	6.26 (1.00)	5.81	5.74	6.40 (0.82)
Lacks Emotional Stability	5.12 (1.58)	4.59	3.97	4.74 (1.72)
Lack Remorse	6.32 (1.11)	6.65	6.49	6.56 (0.72)
Self (7)				
Self-Centered	6.22 (0.93)	6.19	6.21	6.48 (0.69)
Self-Aggrandizing	5.00 (1.54)	6.03	5.60	6.14 (0.84)
Sense of Uniqueness	5.50 (1.31)	5.61	5.25	5.58 (1.18)
Sense of Entitlement	5.81 (1.02)	6.05	6.07	6.03 (1.29)
Sense of Invulnerability	5.19 (1.44)	5.64	5.20	5.61 (1.08)
Self-Justifying	6.05 (1.13)	6.01	6.07	6.03 (1.03)
Unstable Self-Concept	4.83 (1.59)	3.81	2.31	3.96 (1.69)
Foils (9)				
Dependent	3.44 (1.79)	2.29	0.00	2.41 (1.41)
Perfectionist	4.56 (1.56)	3.23	1.49	2.63 (1.50)
Conscientious	1.84 (1.14)	2.27	0.65	1.96 (1.52)
Considerate	1.89 (1.43)	1.55	0.33	1.67 (1.25)
Strange	4.29 (1.73)	2.95	0.86	3.03 (1.47)
Restrained	4.02 (1.86)	2.29	1.26	2.46 (1.48)
Shy	3.81 (1.76)	2.21	0.51	1.66 (1.13)
Cautious	3.55 (1.78)	1.89	0.51	2.12 (1.31)
Self-Conscious	3.24 (1.63)	1.93	0.56	3.14 (2.01)

Note. Comparison of mean prototypicality ratings across CAPP studies. a The Lim et al. (2014) data for the comparison is the forensic psychiatric sample; b the 132 MHPs' ratings from Murray (2014, dissertation paper) were made on a 4-point scale (0-3), and so here was prorated to a 7-point scale; c the Kreis et al. (2012) for the comparison is 132 mental health professionals.

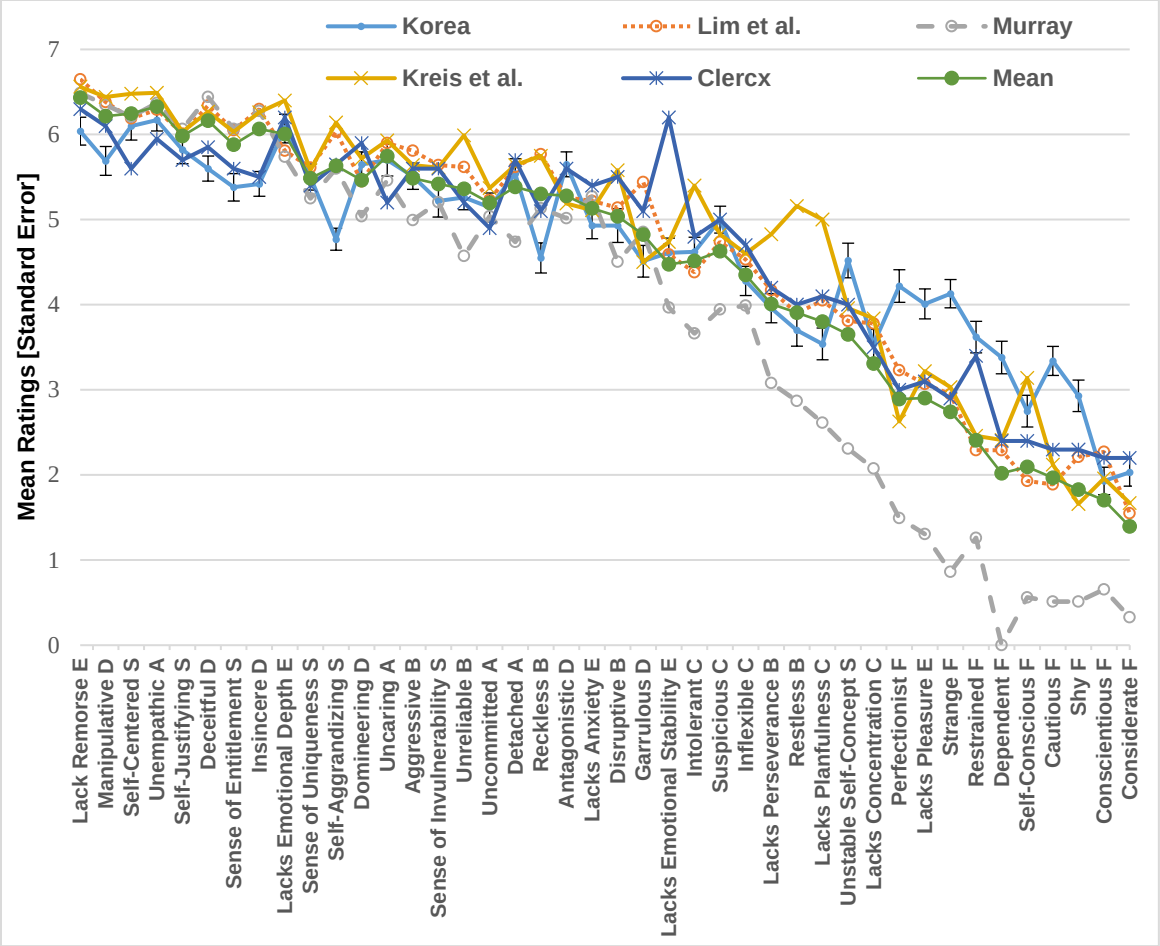

Figure 6. Prototypical evaluations of CAPP symptoms and foils among experts in the International samples (in English version)

Note. Comparison of mean prototypicality ratings across CAPP studies. The Lim et al. (2014) data the comparison is the forensic psychiatric sample; the 132 MHP ratings from Murray (2014, dissertation paper) were first made on a 4-point scale (0-3), and prorated here to a 7-point scale; the Kreis et al. (2012) for the comparison is 132 mental health professionals; Clercx et al. (2018, thesis paper, not published) has 172 professionals in juvenile forensic mental hospital.

As can be seen in Table 8 and Figure 6, Korean sample and three International samples showed similar pattern on 33 CAPP symptoms and nine foils. Representatively, five symptoms (Unempathic, Lacks Remorse, Self-centered, Manipulative and Self-justifying) were commonly rated as highly prototypical (ratings above 5). As would be expected from past research, a few CAPP symptoms (e.g., Lacks Concentration, Lacks Perseverance, and Unstable self-conception) were rated as least prototypical (ratings below 4). In Korean expert, the two symptoms (Unstable Self-concept and Lacks Planfulness) were somewhat rated higher than in other groups (4.83, 3.87 respectively). Even though the four foils (Shy, Cautious, Dependent, Self-conscious, Considerate, and Conscientious) were rated below 4 at least prototypical, two symptoms (Perfectionist and

Strange) were rated as slightly over 4 only in Korean experts. Compared to the mean of the all groups, ratings in Korea appeared to be somewhat lower across CAPP symptoms and higher across several foils (e.g., Shy, Restrained, Cautious, Dependent). Nonetheless overall, across the ratings, experts rated CAPP symptoms as being much more prototypical of PPD than were the Foils symptoms (see Figure 6).

2.2.2.5 Summary in prototypicality analysis

Table 9. Pearson's r Correlations Between CAPP Prototypicality Studies

	Asia		European					International (English)			
	Sea	Vahid	Hanna	Florez.	Hoff	Sörman	Pauli	Clercx	Lim	Murray	Kreis
Vahid	.80										
Hanna	.92	.81									
Hoff	.90	.78	.95								
Florez	.91	.85	.93	.94							
Sorman	.86	.80	.95	.91	.91						
Pauli et al.	.85	.75	.91	.88	.90	.90					
Clercx	.91	.80	.94	.93	.92	.93	.86				
Lim et al.	.91	.79	.96	.95	.92	.96	.87	.95			
Murray	.90	.79	.97	.92	.91	.95	.87	.96	.99		
Kreis et al.	.88	.79	.92	.93	.91	.92	.80	.94	.96	.95	-
Total	.94	.85	.98	.97	.96	.97	.92	.97	.98	.98	.96

Note. All correlations sig. at p<.001 level (2-tailed); the present study has 88 forensic experts; the Sörman et al. (2014) data has 90 forensic evaluators; the Florez et al. (2012) data is 187 expert group; the Hoff et al. (2012) data is 211 forensic psychiatric professionals; the Lim et al. (2014) data for the comparison has forensic psychiatric sample; the 132 MHPs' ratings from Murray (2014, dissertation paper); the Pauli et al. (2018) has 87 correctional staffs; the Kreis et al. (2012) has 132 mental health professionals, e.g., twenty-one different nationalities and six world regions; Clercx et al. (2018, thesis paper, not published) data has 172 professionals in juvenile forensic mental hospital; Pauli et al. (2017) has 87 correctional officers in two institutes; Hanna (2018, not published) has German forensic experts; Vahid (2018) has Iran forensic experts.

I investigated the correlations between the mean prototypicality ratings of the CAPP and foils made by experts in the present study to those made by experts in the previous studies. The mean prototypicality ratings in present study had generally very large correlations with other countries' ratings (range = .80 - .92). Relatively, Valid (2108)'s study has somewhat lower correlations (range = .75 - .85) with other countries possibly due to experts sampling or translation procedure. Nonetheless, these are also large correlations (>.70). Overall, the correlations were all strongly large, r ranging from .71 to .99 in total level (all, $p < .000$) (see Table 9). In summary, these results clearly showed that experts in the current study conceptualized PPD in almost same way as that of experts in previous research using the CAPP.

Table 10. Mean and Stand Deviation of CAPP symptoms, foils, and domains among experts in the eleven prototypicality studies

	Domain	Number	Symptom	*Mean*	*SD*	95%CI LB	95%CI UB
1	Emotional	5	Lacks remorse	6.32	0.24	6.18	6.46
2	Dominance	4	Manipulative	6.30	0.21	6.18	6.42
3	Self	1	Self-Centered	6.18	0.25	6.03	6.33
4	Attachment	3	Unempathic	6.15	0.24	6.00	6.29
5	Dominance	3	Deceitful	5.99	0.31	5.80	6.17
6	Self	6	Self-Justifying	5.99	0.20	5.87	6.10
7	Self	4	Sense of Entitlement	5.90	0.16	5.81	6.00
8	Dominance	5	Insincere	5.81	0.37	5.59	6.02
9	Emotional	3	Lacks emotional Depth	5.78	0.41	5.54	6.03
10	Dominance	2	Domineering	5.77	0.35	5.57	5.97
11	Self	3	Sense of Uniqueness	5.71	0.47	5.43	5.99
12	Self	2	Self-Aggrandizing	5.68	0.35	5.48	5.89
13	Attachment	4	Uncaring	5.55	0.37	5.33	5.77
14	Self	5	Sense of Invulnerability	5.51	0.28	5.34	5.67
15	Behavioural	6	Aggressive	5.45	0.26	5.30	5.61
16	Attachment	2	Uncommitted	5.40	0.41	5.16	5.64
17	Behavioural	2	Unreliable	5.39	0.48	5.11	5.67
18	Dominance	1	Antagonistic	5.20	0.47	4.92	5.48
19	Attachment	1	Detached	5.18	0.65	4.80	5.56
20	Emotional	1	Lacks anxiety	5.17	0.40	4.94	5.41
21	Behavioural	3	Reckless	5.09	0.53	4.78	5.41
22	Behavioural	5	Disruptive	5.08	0.52	4.77	5.39
23	Dominance	6	Garrulous	5.03	0.49	4.74	5.32
24	Emotional	4	Lacks emotional Stability	5.01	0.68	4.61	5.41
25	Cognitive	3	Intolerance	4.91	0.56	4.57	5.24
26	Cognitive	4	Inflexible	4.86	0.55	4.54	5.18
27	Cognitive	1	Suspicious	4.81	0.44	4.55	5.07
28	Behavioural	1	Lacks perseverance	4.10	0.66	3.71	4.49
29	Behavioural	4	Restless	3.99	0.69	3.58	4.39
30	Cognitive	5	Lacks planfulness	3.78	0.74	3.34	4.21
31	Self	7	Unstable Self-Concept	3.76	0.70	3.34	4.17
32	Foils	2	Perfectionist	3.61	1.14	2.94	4.29
33	Cognitive	2	Lacks concentration	3.35	0.61	2.99	3.71
34	Foils	5	Strange	3.33	1.10	2.68	3.98
35	Emotional	2	Lacks pleasure	3.17	0.77	2.72	3.63
36	Foils	6	Restrained	2.90	0.96	2.33	3.46
37	Foils	1	Dependent	2.41	0.93	1.86	2.96
38	Foils	3	Conscientious	2.23	0.75	1.79	2.68
39	Foils	8	Cautious	2.18	0.73	1.75	2.61
40	Foils	9	Self-Conscious	2.15	0.69	1.74	2.56
41	Foils	7	Shy	2.01	0.56	1.68	2.34
42	Foils	4	Considerate	1.99	0.71	1.56	2.41

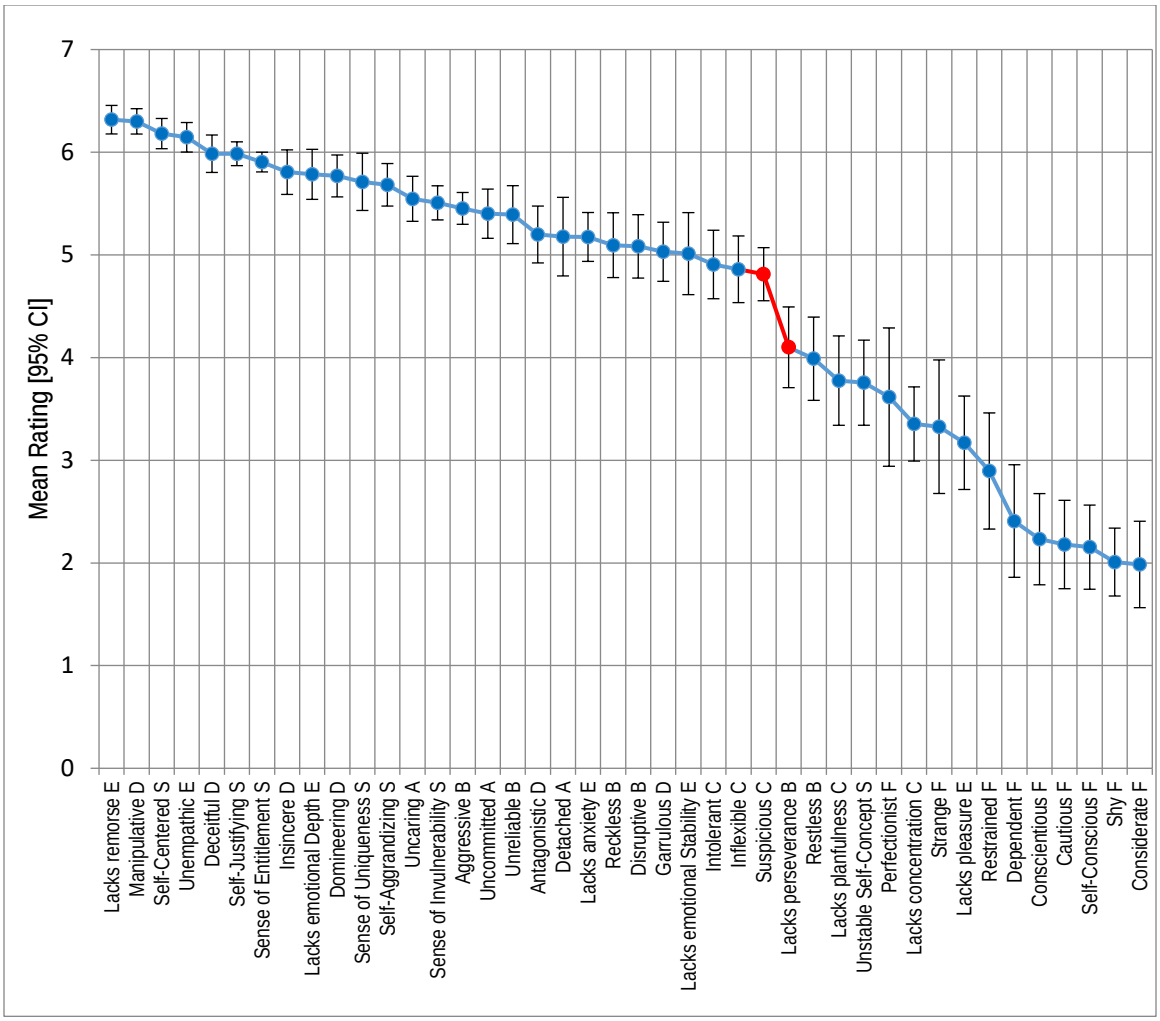

Figure 7. Prototypical evaluations of CAPP symptoms and foils among experts in the eleven prototypicality studies

As seen in Table 10 and Figure 7, the 33 CAPP symptoms and 9 foils by total mean based on the 11 studies were sorted in descending order and then stock-graphed. As a result, there is a big "scree" (see red line) showing discontinuity at a rating value of about 4.5; seen in Figure 7, it could be possibly suggested that "Good" CAPP symptoms have mean values and 95% CI LB values > 4.5; "bad" CAPP symptoms, and all the foils, have mean values and 95% CI UB values < 4.5.

In other words, six symptoms such as Lacks perseverance, Restless, Lacks planfulness, Unstable self-concept, Lacks concentration, and Lacks pleasure were commonly outsiders and less prototypical of PPD.

Chapter 3. Study 2.

3.1. Method

3.1.1. Ethics and Protection of Human Subjects

Initial ethics approval for Study 2 was obtained from KNCS and Ajou University, who collaborated on the collection of data in Korea under the direction of SJH. Once the data were collected, ethics approval was sought from Simon Fraser University to undertake secondary analysis of the data as part of SJH's dissertation research.

3.1.2. Participants

The offenders were 100 inmates—88 men and 12 women—recruited from six correctional facilities of the KNCS: Seoul, Uijeongbu, Gunsan, Chungju, Chungju (women), and Pohang. Most offenders in these prisons have served long-term sentences for the charges such as murder, sexual assault, arson, fraud, assault, robbery, and theft. Offenders at the institutions were approached by prison psychologists on behalf the KNCS and invited to participate in the present study in return for a payment equivalent to about CAD $50. All offenders who read and consented to the study procedures and volunteered to participate were recruited, except those identified by prison psychologists as suffering from acute or severe mental health symptoms that might impair their ability to provide informed consent or complete the assessment procedures. Recruitment continued until the sample size reached the initial target of 100 offenders (see Table 11 for a short overview). A total of 11 offenders withdrew from the study prior to completing the second assessment, resulting in a total sample of 89 (79 male and 10 female). The mean age at the time of assessment was 40.09 (SD = 12.61) and range of 21 years to 69 years old. Within the sample, most participants reported being married (27%), divorced/separated (16%), or not married/single (56%). Many offenders were charged with more than one index offense and most commonly serious violence.

Table 11. Demographic Information of Serious and Violent Incarcerated offenders

Variables	*N (SD)*
Gender, *n* (%) male	79 (88.7%)
Age, *M* (*SD*)	40.09 (12.61)
Education, *n* (%)	
≤ Grade 6 (elementary level)	20 (22.4%)
≤ Grade 9 (middle level)	13 (14.6%)
≤Grade 12 (high level)	32 (35.9%)
> Grade 12 (college level)	24 (26.9%)
Marital status	
Single/Never married/	50 (56.1%)
Married/Common-law	24 (26.9%)
Divorced/Separated	14 (15.7%)
Widowed	1 (1.1%)
Total convictions, *M (SD)*	8.69 (8.70)
Months of imprisonment, *M (SD)*	71.22 (86.95)
Disciplinary infractions, *M (SD)*	4.12 (8.56)
Main crime, *n* (%)	
Homicide	15 (16.8%)
Sexual assault	30 (33.7%)
Assault causing bodily Harm/ Infliction of injury	23 (25.8%)
Robbery	7 (7.8%)
Theft	5 (5.6%)
Arson	3 (3.3%)
Fraud	5 (5.6%)
Drugs related	1 (1.1%)

3.1.3. Procedure

Offenders who volunteered to participate were provided with a full version of description of the study and clarification of its requirements and purpose, and made aware of their own human rights as study participants, any potential risks associated with the administration of the study, and the potential benefits of the study for the field of psychology. Then some limits of confidentiality were introduced. Offenders were advised that their participation, and the information they provided during participation, would not affect their status, condition, prison level, and treatment process within the KNCS in any way.

Participants who provided informed consent took part in two assessment sessions, 3 months apart, each of which lasted approximately four to five hours. Each participant was assessed by the same team of four experts on both occasions. The experts were 24 correctional psychologists from six correctional institutions, that is, team from each

institution. The use of teams to conduct assessments is consistent with usual custom in Korea, where the law requires at least three correctional staff members to be present during assessment and treatment activities for safety reasons.

At the first assessment occasion, the experts independently reviewed the case file information and staff observations of the participant, then jointly interviewed the participant. One expert was selected to conduct both the sociodemographic interview and the semi-structured interview for the K-CAPP-IRS, a second to conduct the semi-structured interview for the K-PCL-R, and the third and fourth to observe the interviews. Following the joint interviews, the K-CAPP-IRS interviewer and the two observers independently made ratings on the K-CAPP-IRS, the K-PCL-R interviewer made K-PCL-R ratings, and the participants last completed K-PPI-R. The independent ratings were based on all available information from records, staff observations, K-CAPP-Staff form (i.e., informant form) and the interviews. K-CAPP-Staff form was not validated in Korea and thus used as a side-reference data for the ratings. At the second assessment occasion, the same team of experts repeated the entire process. Each expert played the same role (K-CAPP-IRS interviewer, K-PCL-R interviewer, or observer) at the first and second assessments of the same participant at the same place interviewed.

3.1.4. Instruments

3.1.4.1 K-PCL-R

The K-PCL-R (Cho & Lee, 2008) is the Korean translation of the widely used PCL-R, an expert observer ratings scale of psychopathy. Based on their review of file information concerning and interviews with offenders, experts rated the lifetime presence of 20 features of psychopathy on a 3-point scale (0 = item does not apply, 1 = item applies in some respects, 2 = item applies). Items are summed to yield four facet scores (Facet 1 = Interpersonal, Facet 2 = Affective, Facet 3 = Lifestyle, Facet 4 = Antisocial); two factor scores (Factor 1 = Interpersonal/Affective, Factor 2 = Social Deviance); and a total score. The items in the original PCL-R and their loading on the facets and factors are presented in Table 12. The psychometric properties of the K-PCL-R appear to be equivalent to those of the original PCL-R (Cho & Lee, 2008).

Table 12. Items in the PCL-R

Item		Facet	Factor
1.	Glibness/superficial charm	1	1
2.	Grandiose sense of self worth	1	1
3.	Need for stimulation/proneness to boredom	3	2
4.	Pathological lying	1	1
5.	Conning/manipulative	1	1
6.	Lack of remorse or guilt	2	1
7.	Shallow affect	2	1
8.	Callous/lack of empathy	2	1
9.	Parasitic lifestyle	3	2
10.	Poor behavioral controls	4	2
11.	Promiscuous sexual behavior	--	--
12.	Early behavioral problems	4	2
13.	Lack of realistic, long-term goals	3	2
14.	Impulsivity	3	2
15.	Irresponsibility	3	2
16.	Failure to accept responsibility for own actions	2	1
17.	Many short-term marital relationships	--	--
18.	Juvenile delinquency	4	2
19.	Revocation of conditional release	4	2
20.	Criminal versatility	4	2

3.1.4.2 K-PPI-R

The K-PPI-R (Lee & Park, 2009) is the Korean translation of the widely-used PPI-R (Lilienfeld & Widows, 2005). The PPI-R is a 154-item self-report measure designed to assess different facets of psychopathic personality. Items are declarative statement phrased in first-person, singular. Respondents rate the degree to which each statement applies to them on a 4-point Likert-type scale (0 = false, 1 = mostly false, 2 = mostly true, 3 = true). Items are summed to yield a total score, as well as eight subscale scores (Machiavellian Egocentricity, Social Influence, Coldheartedness, Carefree Nonplanfulness, Fearlessness, Blame Externalization, Rebellious Nonconformity, and Stress Immunity) and three factor scores (Self-Centered Impulsivity, Fearless Dominance, and Coldheartedness). The psychometric properties of the K-PPI-R appear to be very similar to those of the original PPI-R (Lee & Park, 2009; Lee, Gong, & Park, 2007; Jung, 2015; Lee, 2012).

3.1.4.3 K-CAPP-IRS

The K-CAPP-IRS is the Korean translation of the CAPP-IRS (Cooke, et al., 2004), which relied heavily on Korean translation of the CAPP model (i.e., the K-CAPP). Experts

rated the K-CAPP-IRS symptoms based on interviews with participants (conducted using a translation of the standard CAPP-IRS interview), observations by staff who knew the participants well, and review of their institutional records. Each of the 33 symptoms was rated in terms of trait extremity and functional impairment on a 4-point scale (0 = none, 4 = severe) and global symptom severity on a 7-point scale (0 = none, 6 = very severe) for a timeframe of lifetime, 6 months, and 12 months. The whole process of the interview was generally followed by guideline suggested by Cooke and Logan (2015). Specifically, the researchers are prompted to discuss each symptom-related part via one or more starter questions and responses from the client are followed up by a series of additional probes and interview techniques.

3.1.5. Training and Supervision of Experts

In terms of the raters' background, the correctional professionals had worked in institutional settings for a mean of five years, ranging from three to twelve years. They all were clinically licensed psychologists approved by the KPA. Along with this, they annually took formal training programs for therapeutic techniques and risk assessments provided by the Legal Research and Training Institute (LRTI) in the Ministry of Justice. The experts were all experienced working with inmates with PPD, mostly in the context of conducting assessments of risk or treatability for the purpose of making placement decisions, but none had prior experience using the CAPP-IRS.

Considerable effort was put into training and supervision of experts for the purpose of this study. Prior to data collection, they attended a 3-day training in the use of the CAPP-IRS, along with completing 3 training cases under the supervision of SDH and SJH; and also attended a 2-day training in the use of the PCL-R, followed by completion of 5 training cases under the supervision of SDH and SJH.

3.1.6. Data Analysis

To evaluate the interrater reliability of the CAPP-IRS, I started by calculating simple intraclass correlation coefficients (ICC; mixed, two-way; raters random, measure fixed, absolute agreement method; Shrout & Fleiss, 1979) to test the agreement between three raters for 89 cases for ratings of the CAPP symptoms, domains, and total scores. I calculated ICCs for single ratings (ICC1, reflecting the average reliability between pairs of

raters) and averaged ratings (ICC2, reflecting the estimated reliability between groups of raters). ICCs were interpreted using the guidelines recommended by Fleiss (1981): < .40 = poor; .40 to .59 = fair; .60 to .74 = good; and > .75 = excellent. Landis and Koch (1977) also suggested the following guideline: below 0 is “poor,” .00 to .20 is “slight,” .21 to .40 is “fair,” .41 to .60 is “moderate,” .61 to .80 is “substantial,” and .81 to 1.00 is “almost perfect.”

I also evaluated the interrater reliability of the K-CAPP-IRS within the framework of Generalizability (G) Theory (Cronbach, Gleser, Nanda, & Tajaratnam, 1972). G theory has several advantages over Classical Test Theory (Bogels et al., 1985; Jackson & Paunonen, 1980; Mitchell, 1979). It allows researchers to simultaneously examine unreliability due to several sources of systematic error, some of which were considered unsystematic error in CTT. It also provides information concerning how measurement can be simplified or improved in more detail. G theory also provides a single index of the adequacy of measurement, whereas CTT theory may require that many indices of reliability and their relative importance be evaluated.

Principal results of a G-study are estimates of the variance components (see Brennan, 1999a, 1992b; Cronbach, Gleser, Nanda, & Rajaratnam, 1972; Shavelson & Webb, 1991). For example, the person variance component is an estimate of the variance across persons of person-level mean scores, where the mean is taken across all tasks and raters in the universe. The task component is the estimated variance of task mean scores, where each mean is taken across all items, persons, occasions, and raters. The rater component is the variance of rater mean scores, where each mean is across persons and tasks. I conducted G Theory analyses using Genova 3.1 for complete and balanced design (Brennan, 2001; Crick & Brennan, 1982, upgraded in 2001), with a fully crossed p x o x r x i design, where p was persons (89 offenders), o was occasions (2 assessment sessions, 3 months apart), r was raters (3 professionals), and i was items (33 K-CAPP-IRS symptom ratings). In my research, I did not use Ur Genova and m Genova for unbalanced random design because they provide only some restricted class of design and don’t inform D-study. See Brennan’ book (Generalizability Theory) for more information and also see the information presented in the Table 13 for details.

Table 13. p x i x o x r in G-study – Full Crossed Radom Model (FCRM)

2 Occasions (a 3-month apart) [a]		*1*			*2*		
3 Same Raters [b]		*1*	*2*	*3*	*1*	*2*	*3*
33 Same Items [c]		*1····33*	*1····33*	*1····33*	*1····33*	*1····33*	*1····33*
89 Persons	1						
	2						
	3						
	.						
	.						
	89						

Note: a time interval between two occasions was three months; b raters are same as first raters by occasions; c all raters rated the same 33 items based on the same materials.

3.2. Results

The K-CAPP-IRS total, domains, and item ratings score descriptive can be seen in Table 14. The mean score in the K-CAPP total is 78.88 (SD = 34.13). As for mean score in domains, Attachment, Behavioural, Cognitive, Dominance, Emotional, and Self are 10.47 (SD = 6.12), 14.58 (SD = 7.84), 13.62 (SD = 6.88), 12.35 (SD = 6.81), 12.96 (SD = 6.35), and 14.90 (SD = 7.96), respectively. All of the ratings in K-CAPP-IRS were obtained from the first assessment occasion.

Table 14. Distribution of CAPP-IRS Total, Domain, and Item Ratings

CAPP-IRS	*Max*	*M*	*SD*
TOTAL	*198*	*78.88*	*34.13*
Attachment	*24*	*10.47*	*6.12*
Detached	6	2.73	1.91
Uncommitted	6	2.58	1.82
Unempathic	6	2.86	1.79
Uncaring	6	2.42	1.73
Behavioural	*36*	*14.58*	*7.84*
Lacks Perseverance	6	1.89	1.76
Unreliable	6	2.95	1.80
Reckless	6	3.40	1.95
Restless	6	1.36	1.59
Disruptive	6	2.32	1.97
Aggressive	6	2.67	1.93
Cognitive	*30*	*13.62*	*6.88*
Suspicious	6	3.11	1.77
Lacks Concentration	6	1.83	1.72
Intolerance	6	2.54	1.90
Inflexible	6	3.00	1.77

Lacks Planfulness	6	3.14	2.05
Dominance	*36*	*12.35*	*6.81*
Antagonistic	6	2.05	1.82
Domineering	6	1.96	1.71
Deceitful	6	2.26	1.63
Manipulative	6	1.78	1.57
Insincere	6	2.41	1.63
Garrulous	6	1.89	1.80
Emotional	*30*	*12.96*	*6.35*
Lacks Anxiety	6	1.80	1.73
Lacks Pleasure	6	2.49	2.07
Lacks Emotional Depth	6	2.54	1.72
Lacks Emotional Stability	6	2.67	1.82
Lack Remorse	6	3.46	1.92
Self	*42*	*14.90*	*7.96*
Self-Centered	6	3.07	1.82
Self-Aggrandizing	6	1.68	1.63
Sense of Uniqueness	6	1.62	1.49
Sense of Entitlement	6	1.65	1.49
Sense of Invulnerability	6	1.27	1.47
Self-Justifying	6	3.19	1.85
Unstable Self-Concept	6	2.42	1.97

3.2.1. K-CAPP-IRS Ratings

To determine the consistency of the domains in measuring in the construct of PPD, internal consistency reliability of the CAPP domains scores was assessed by using Cronbach's alpha. Acceptable levels of internal consistency reliability as measured with Cronbach's alpha typically exceed .70, meaning that they explain at least 50% of the variance of the construct (Vogt, 2007). The alpha for three domain scores (e.g., Behavioural, Dominance, Self) as a scale measuring the construct of psychopathy was not acceptable, ranging from .613 to .684. All of the alpha values changed significantly at this point, indicating that all six domains were not reliably contributing towards the internal consistency of the CAPP scale (see Table 15).

Table 15. Internal Consistency Reliability of Domain Scores from Items

Domains	Alpha (a)	Alpha (a) if item deleted scale
Attachment	.767	.871
Behavioural	.684	.883
Cognitive	.822	.860
Dominance	.613	.892
Emotional	.802	.865
Self	.664	.887

As shown in Table 16 internal consistency was not perfect. Almost all of items problematic with the exception of five items (e.g., four items in Attachment and one item in Emotion) as the analysis suggested that their removal may improve the reliability of the scale. For example, the analysis suggested by removing the item 'Detached', reliability could be increased to a=.84. Similarly, dropping the time 'Lacks Perseverance' from domain of Behavioural would result from .57 to .74. In such a way, all items could be increased by removing the item itself.

Table 16. Internal Consistency Reliability of Item Scores from Items

CAPP-IRS Symptoms	Alpha (*a*)	Alpha (*a*) if item deleted scale
Detached	.767	.851
Uncommitted	.789	.843
Unempathic	.741	.863
Uncaring	.750	.852
Lacks Perseverance	.610	.763
Unreliable	.585	.768
Reckless	.551	.777
Restless	*.302*	*.823*
Disruptive	.647	.753
Aggressive	.672	.747
Suspicious	.535	.776
Lacks Concentration	.574	.765
Intolerance	.692	.726
Inflexible	.659	.739
Lacks Planfulness	*.475*	*.800*
Antagonistic	*.282*	*.778*
Domineering	.534	.708
Deceitful	.617	.686
Manipulative	.605	.702
Insincere	.606	.708
Garrulous	*.441*	*.734*

Lank Anxiety	*.294*	*.732*
Lacks Pleasure	.513	.652
Lacks Emotional Depth	.727	.581
Lacks Emotional Stability	.404	.701
Lack Remorse	.518	.657
Self-Centered	.670	.753
Self-Aggrandizing	.653	.759
Sense of Uniqueness	.643	.763
Sense of Entitlement	.698	.755
Sense of Invulnerability	.488	.788
Self-Justifying	*.341*	*.818*
Unstable Self-Concept	*.386*	*.812*

Table 17. Bivariate Correlation between Domains and CAPP Scores

Domain	CAPP Domains						
	A	B	C	D	E	S	Total
A	-						
B	.60**	-					
C	.81**	.72**	-				
D	.41**	.51**	.50**	-			
E	.82**	.63**	.75**	.44**	-		
S	.54**	.44**	.56**	.71**	.61**	-	
Total[a]	.83**	.80**	.87**	.73**	.87**	.80**	-

Note. * = p<.05, **= p<.01, ***= p<.001 (2-tailed); a C > E > A > B > S > D in order by the level of bivariate correlation; A = Attachment domain, B = Behavioral domain, C = Cognitive domain, D = Dominance domain, E = Emotional domain, and S = Self domain.

Psychopathy as suggested by Cooke et al. (2015) is a superordinate construct underpinned by six domains of functioning. Thus we could expect that all of six domains will be moderately related to each other and more correlated to the construct of psychopathy measured by total scores. As a result, as expected, all domains had significant associations with total scores (rang=.73-.87) (see Table 17). Though two of six domains had a strong relationship with each other (e.g., r=0.81 between Attachment and Cognitive, r =0.82 between Attachment and Emotional), the remains also showed large relationships between each other in accordance with previous studies (McCormick, 2007). This, as Vincent (2002) suggested, indicates that psychopathy could be defined as unidimensional, as the included items all pertain to one overarching construct, psychopathy. Thus these domains all properly shared a common construct or coherent syndrome, showing large correlations with each other.

Compared to other relationships, even significantly related and not very large, two relationships (between Attachment and Dominance, between Behavioural and Self) were relatively lower than the others identified (r=.41, r=.54, respectively). Given the results presented in Table 17, domains showing the strongest relationship with K-CAPP total score were Emotional and Self (r=.87 equally). On the whole, all other domains had significantly large relationships with total score, but Dominance was smaller than other domains and total score correlations.

3.2.2. Interrater Reliability of the K-CAPP-IRS Ratings

3.2.2.1 ICC Analyses

I used intraclass correlation coefficient (ICC; mixed, two-way; raters random, measure fixed, absolute agreement) to test the agreement among three raters for 89 cases for symptom, domain, and total ratings. I calculated both the single measure ICC1 (reliability of a group of raters) and the average measure ICC2 (reliability of a group of raters). ICC1 would be applicable to a context in which a single professional is using the CAPP-IRS, whereas ICC2 would be applicable when groups of raters are using it, and relying on averaged ratings.

The ICCs for all symptoms of six domains with exception of Restless were ‘almost perfect,” ranging from .82 to .90, seen in Table 18.

Table 18. Symptom Level Intraclass Correlation Coefficient for Raters

Domains		ICCs			
Symptoms		*n*	Raters	ICC_1	ICC_2
Attachment (A)					
	Detached	89	3	.82 [.75-.88]	.93 [.90-.95]
	Uncommitted	89	3	.90 [.86-.93]	.96 [.95-.97]
	Unempathic	89	3	.85 [.79-.89]	.94 [.91-.96]
	Uncaring	89	3	.82 [.76-.88]	.93 [.90-.95]
Behavioural (B)					
	Lacks Perseverance	89	3	.89 [.84-.92]	.96 [.94-.97]
	Unreliable	89	3	.91 [.87-.94]	.96 [.95-.98]
	Reckless	89	3	.83 [.77-.88]	.93 [.91-.96]
	Restless	89	3	.72 [.62-.80]	.88 [.83-.92]
	Disruptive	89	3	.86 [.80-.90]	.95 [.92-.96]
	Aggressive	89	3	.93 [.91-.95]	.97 [.96-.98]
Cognitive (C)					
	Suspicious	89	3	.85 [.80-.90]	.94 [.92-.96]
	Lacks Concentration	89	3	.85 [.79-.90]	.94 [.92-.96]
	Intolerance	89	3	.90 [.85-.93]	.96 [.94-.97]
	Inflexible	89	3	.87 [.81-.91]	.95 [.93-.96]
	Lacks Planfulness	89	3	.90 [.86-.93]	.96 [.95-.97]
Dominance (D)					
	Antagonistic	89	3	.88 [.83-.92]	.95 [.93-.97]
	Domineering	89	3	.87 [.82-.91]	.95 [.93-.97]
	Deceitful	89	3	.88 [.83-.92]	.95 [.93-.97]
	Manipulative	89	3	.90 [.86-.93]	.96 [.95-.97]
	Insincere	89	3	.86 [.81-.91]	.95 [.93-.96]
	Garrulous	89	3	.91 [.87-.94]	.96 [.95-.98]
Emotional (E)					
	Lank Anxiety	89	3	.84 [.78-.89]	.94 [.91-.96]
	Lacks Pleasure	89	3	.91 [.88-.94]	.97 [.95-.98]
	Lacks Emotional Depth	89	3	.88 [.83-.92]	.95 [.93-.97]
	Lacks Emotional Stability	89	3	.84 [.78-.89]	.94 [.91-.96]
	Lack Remorse	89	3	.88 [.83-.92]	.95 [.94-.97]
Self (S)					
	Self-Centered	89	3	.90 [.87-.93]	.96 [.95-.97]
	Self-Aggrandizing	89	3	.87 [.82-.91]	.95 [.93-.97]
	Sense of Uniqueness	89	3	.87 [.81-.91]	.95 [.93-.96]
	Sense of Entitlement	89	3	.82 [.75-.87]	.93 [.90-.95]
	Sense of Invulnerability	89	3	.86 [.81-.91]	.95 [.93-.96]
	Self-Justifying	89	3	.88 [.83-.92]	.95 [.93-.97]
	Unstable Self-Concept	89	3	.88 [.83-.92]	.95 [.93-.97]

Note. 95% confidence intervals in parentheses; ICCs calculated using 2-way random effects models, absolute agreement; ICC=intraclass correlation coefficient; the lowest level in ICCs is bolded; Domains of the respective are indicated in brackets

With respect to domain and total scores, the results indicated “almost perfect” interrater reliability, with ICC1 ranging from .91 to .93 and ICC2 ranging from .96 to .97 (see Table 19).

Table 19. Intraclass Correlation Coefficients for K-CAPP-IRS Total and Domain Ratings

K-CAPP-IRS	*n*	Raters	ICC_1	ICC_2
Domain A: Attachment	89	3	.91 [.88, .94]	.97 [.95, .98]
Domain B: Behavioural	89	3	.92 [.89, .94]	.97 [.96, .98]
Domain C: Cognitive	89	3	.93 [.90, .95]	.97 [.96, .98]
Domain D: Dominance	89	3	.93 [.90, .95]	.97 [.96, .98]
Domain E: Emotional	89	3	.92 [.89, .95]	.97 [.96, .98]
Domain S: Self	89	3	.92 [.89, 94]	.97 [.96, .98]
Total Score	89	3	.91 [.87, .94]	.96 [.95, .98]

Note. 95% confidence intervals in parentheses; ICC = intraclass correlation coefficient; ICCs calculated using 2-way random effects models, absolute agreement.

3.2.3. Generalizability (G) Theory Analyses

Generalizability theory provides a framework for evaluating measurement reliability that is more comprehensive than that of classical test theory (Bogels et al., 1995; Cronbach, Gleser, Nanda, & Tajaratnam, 1972). The estimates of the variance components in G-study were obtained using analysis of variance procedures suggested by Brennan (1992a, 1999b, 2001). In the current study, as noted previously, the design is a fully crossed p x o x r x i design, with 89 Persons, 2 Occasions, 3 Raters, and 33 Items. The mean squares, sum of squares, and estimated variance components from the p x o x r x i ANOVA are presented in Table 20.

3.2.3.1 G-study Variance Components

Table 20. Estimated variance component for the G-study [p x o x r x i Design – Full Crossed Random Model]

Effect	*d.f.*	Sum of Squares	Mean Square	Estimated variance component	Percentage of total variance
p [a]	88	14,431.65	180.40	0.82	25.66
o [b]	1	3.87	3.87	<0.01	<0.01
r [c]	2	362.64	181.32	0.02	0.69
i [d]	32	2,114.49	66.078	0.03	1.06
po [e]	88	161.87	2.09	0.02	0.49
pr [f]	176	2,688.87	16.81	0.21	6.54
pi [g]	2,880	9,367.57	3.68	0.10	3.03
or [h]	2	1.48	0.74	<0.01	<0.01
oi [i]	32	15.98	0.50	<0.01	<0.01
ri [j]	64	3,133.19	48.96	0.28	8.89
por [k]	160	64.19	0.40	<0.01	0.06
poi [l]	2,560	1,012.13	0.40	0.02	0.63

pri [m]	5,120	15,458.63	3.02	1.34	42.27
ori [n]	64	32.72	0.51	<0.01	0.06
pori [o]	5,120	1,712.27	0.33	0.33	10.53
Total	16,037	50,561.55		3.18	

Note a p person effect; b o occasion effect; c r rater effect: d i item effect; e po person by occasion effect; f pr person by rater effect; g pi person by item effect; h ri rater by item effect; i oi occasion by item effect; j ri rater by item effect; k por person by occasion by rater effect; l poi person by occasion by item effect; m pri person by rater by item effect; n ori occasion by rater by item effect; o pori person by occasion by rater by item effect, plus error; Largest proportion in variance is bolded.

Looking first at main effects, we should expect a large variance component associated with p, indicating that persons differ. This variance component is analogous to true-score variance in CTT and is not considered to be error. As Table 20 indicates, p accounted for 25.66% of the total variance, a proportion that can be considered large, but not strongly homogenous (Shavelson & Webb, 1991; Shavelson et al., 1989). p usually means individual difference, which is ideally large percentage of total variance. In comparison, the variance components for o and r were negligible, indicating a high level of consistency of scores across occasions and raters; and the variance component for i was small, indicating considerable consistency of scores across items (although this effect is of limited interest).

Looking next at two-, three-, and four-way interactions, 6 of 11 were negligible and 5 of 11 were not. Of the latter, one was the two-way interaction p x i, indicating that item scores differed across persons, which is not unexpected and of limited interest. But the other 4 interactions involved r, indicating some important lack of consistency across raters; indeed, the three-way interaction p x r x i accounted for the largest percentage of total variance (42.27%). This means that raters may evaluate items by person very differently.

Overall, then, although the G Theory findings suggested overall good reliability of K-CAPP-IRS ratings across experts and time, they also revealed that unreliability due to raters contributes substantially to the variability in observed score, a finding that consistent with previous G Theory research on measures of PPD (see Schroeder et al., 1983).

3.2.3.2 D-Study Variance Components

D-study variance components are used to estimate various variance errors and reliability-like coefficients. As can be seen in Table 21, I used the comparisons between universe score and relative/absolute error variance to obtain Generalizability Coefficients (GCs) and Dependability Coefficients (DCs). GC is a universe score divided by relative

error variance and Phi (i.e., Dependability) coefficient is a universe score divided by absolute error variance. Absolute Error Variance is the difference between a Pearson's observed mean score (over items, occasions, and raters) and universe score, including all but the person component. Relative Error Variance is all the errors not including o, i, r, oi, or, ir, and oir.

Table 21. G-coefficient and Phi-coefficient in summary of D study results (pxOxIxR Design)

P INF.	O INF.	R INF.	I INF.	Universe Score	Observed Score	Relative Error Variance	Absolute Error Variance	*GC*	*Phi*
Fixed – Item=33									
89	1	1	33	0.815	1.095	0.280	0.312	0.744	0.723
89	1	2	33	0.815	0.965	0.150	0.166	0.845	0.831
89	1	3	33	0.815	0.921	0.106	0.118	0.885	0.874
89	1	4	33	0.815	0.899	0.084	0.093	0.906	0.897
89	2	1	33	0.815	1.081	0.266	0.298	0.754	0.733
89	2	2	33	0.815	0.954	0.138	0.155	0.854	0.840
89	2	3	33	0.815	0.911	0.959	0.107	0.894	0.883
89	2	4	33	0.815	0.899	0.076	0.083	0.915	0.906
89	3	1	33	0.815	1.076	0.260	0.292	0.757	0.735
89	3	2	33	0.815	0.949	0.134	0.151	0.858	0.843
89	3	3	33	0.815	0.907	0.092	0.103	0.898	0.886
98	3	4	33	0.815	0.886	0.071	0.080	0.913	0.910
Random 1									
89	1	1	40	0.815	1.085	0.270	0.300	0.750	0.730
89	1	2	33	0.815	0.964	0.149	0.166	0.844	0.830
89	1	3	20	0.815	0.934	0.119	0.133	0.872	0.859
89	1	4	10	0.815	0.936	0.121	0.130	0.869	0.855
89	2	1	40	0.815	1.072	0.257	0.287	0.760	0.739
89	2	2	33	0.815	0.953	0.138	0.154	0.854	0.840
89	2	3	20	0.815	0.923	0.108	0.121	0.883	0.869
89	2	4	10	0.815	0.923	0.108	0.124	0.882	0.867
89	3	1	40	0.815	1.067	0.252	0.282	0.763	0.742
89	3	2	33	0.815	0.949	0.151	0.151	0.858	0.843
89	3	3	20	0.815	0.919	0.118	0.118	0.886	0.873
89	3	4	10	0.815	0.919	0.120	0.120	0.886	0.871
Random 2 (Minimum item number at the level of acceptable GC)									
89	1	1	800	0.815	1.043	0.228	0.251	0.780	0.764
89	1	2	13	0.815	1.009	0.194	0.219	0.807	0.788
89	1	3	8	0.815	0.985	0.170	0.193	0.827	0.807
89	1	4	4	0.815	1.017	0.202	0.234	0.822	0.801
89	2	1	600	0.815	1.034	0.219	0.242	0.783	0.766
89	2	2	10	0.815	1.013	0.198	0.227	0.820	0.800
89	2	3	5	0.815	1.014	0.199	0.232	0.803	0.777
89	2	4	4	0.815	0.996	0.181	0.212	0.818	0.792
89	3	1	600	0.815	1.031	0.216	0.239	0.800	0.768
89	3	2	10	0.815	1.007	0.192	0.221	0.808	0.786

89	3	3	5	0.815	1.007	0.192	0.225	0.809	0.783
89	3	4	4	0.815	0.989	0.174	0.205	0.824	0.800

Note. GC (Generalizability Coefficient), comparing the universe score to relative score error; Phi (Dependability) coefficient, comparing universe score to the absolute error variance; Absolute Error Variance = the difference between a person's observed mean score (over items, occasions, and raters) and universe score, including all but the person component; Relative Error Variance= all the errors not including o, i, r, oi, or, ir, and oir.

The G- and Phi-coefficient were calculated for each design, based on the fixed – 33 item presented in Table 21. G-coefficient ranged from .74 to .88, while Phi-coefficient ranged from .72 to .88. A G-coefficient > .80 has often been used as a benchmark for acceptable overall score reliability in GT analysis (e.g., Mushquash & O'Connor, 2006), although this criterion has never been validated against independent benchmarks (Brennan & Johnson, 1995).

Through D-study results, as seen in Table 21, it is clear that the number of occasions has little effect since the components of the occasion facet are small and thus there appears to be little advantage to using more than one occasion. On the other hand, due to the large person-item-rater component as well as the non-negligible person-rater component, the number of raters has a great effect on the magnitude of the SEM estimates and the Generalizability and Dependability Coefficients. For example, if we use 33 items, 89 people, 2 occasions, and one rater, the GC is .74 and Phi is .72. Usually raters do not assess the same case twice or three times, by the reason, if we use 89 people, three raters, 33 items and one occasion, the GC is .88 and Phi is .87, which is acceptable range. By increasing from one to two occasions at the same number of rater, the G-coefficient very slightly increased by 0.04 points, whereas increasing from one rater to two at the same of occasion, G-coefficient remarkably increased by 0.11 points, which is an increase almost ten times as large (Phi-coefficient also showed similar pattern with GC) (Brown et al., 2004). Thus, the number of raters would have a great impact on score reliability (Figure 8 and Table 21). If we want to get an acceptable level of G coefficient (>.80), we should have at least more than two raters regardless of the number of occasion (see Figure 8). We also calculated the Generalizability coefficient by the number of item because in the current study we also focused on the number of item based on the facets on raters and occasions as well as their interactions (see Shavelson & Webb, 2005 for the conditions of fixed model). As a result, based on facets in random model, we could find out that regardless of the number of occasions, two-rater needs more than 13 items for over acceptable GC, three-rater needs more than 8 items for over acceptable GC, four-rater needs more than 4 items for over acceptable GC, and so on. In other words, more than

two-rater remarkably decreased the number items at the level of acceptable GC (e.g., one third or one fourth of the CAPP symptoms).

To summarize, principally considered improving overall generalizability, the logical choice would be to either have more raters or have the raters assess inmates more often. Since the variance component of pri is larger than poi, for the better decision, it would increase the rater's number and if you increase the number to four, the GC would remarkably better increase. As well, if we intend to decrease the number of items due to the over-inclusiveness, we might also increase the number of raters, instead of the number of occasions.

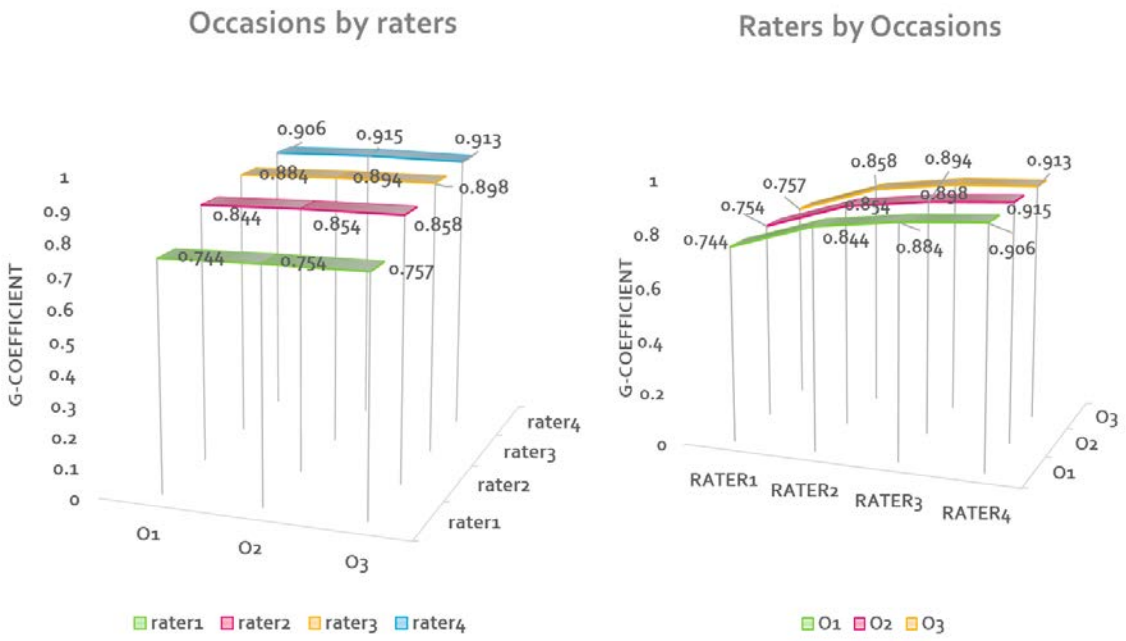

Figure 8. G-Coefficient's shift due to the increase of the number of Occasion and Rater on full cross version.

Note. GC (Generalizability Coefficient) and Phi-C (Dependent coefficient) are acceptable in case of '> 0.8'.

3.2.4. Concurrent Validity of the K-CAPP-IRS

3.2.4.1 K-CAPP-IRS with K-PCL-R scale

The K-PPI-R and the K-PCL-R scale were used to assess the concurrent validity of the K-CAPP-IRS. The correlation between the K-CAPP-IRS total score and K-PCL-R total score was all significantly correlated, ranging from r = .392 to r = .685 (see Table 21). Among six domains, B was most strongly correlated with PCL-R total score (r = .685),

whereas S was most weakly correlated with PCL-R total score (r = .392). The associations between each of the K-CAPP-IRS domains and the four K-PCL-R facets were also examined. Results revealed significant correlations between those measures and domains A through E. Specifically, A was correlated positively with facet 2 (r = .575) and facet 3 (r = .424). B was positively correlated with facet 2 (r = .492), facet 3 (r = .646), and facet 4 (r = .521). C was positively correlated with facet 2 (r = .573), facet 3 (r = .490), and facet 4 (r = .407). D was positively correlated with facet 1 (r = .469) and facet 2 (r = .337). E was positively correlated with facet 2 (r = .543), and facet 3 (r = .420). Domain S was positively correlated with facet 1 (r = .387) and facet 2 (r = .383). However, contrary to expectations, Facet 1 was not correlated with four CAPP domains such as Domain A, B, C, and E.

Table 22. Bivariate Correlations between the K-PCL-R and K-CAPP-IRS

K-PCL-R	K-CAPP-IRS						
Facet	A	B	C	D	E	S	Total
Facet 1	.187	.069	.085	.469**	.065	.387**	.263**
Facet 2	.575**	.492**	.573**	.337**	.543**	.383**	.597**
Facet 3	.424**	.646**	.490**	.248*	.420**	.225	.498**
Facet 4	.279*	.521**	.407*	.199	.242	.092	.354**
Factor 1	.454**	.325**	.413**	.514**	.336**	.417**	.501**
Factor 2	.345**	.647**	.416**	.132	.331**	.032	.406**
Total	.550**	.685**	.610**	.438**	.504**	.392*	.647**

Note. * p < .05, ** p < .01, *** p < .001 (2-tailed); Large correlations between K-CAPP-IRS and K-PCL-R are bolded; facet 1 = interpersonal, facet 2 = affective, facet 3 = lifestyle, and facet 4= antisocial; factor 1 = facet 1 + facet 2, factor 2 = facet 3 + facet 4; all ratings are obtained at the first assessment occasion.

3.2.4.2 K-CAPP-IRS and K-PPI-R

The correlation between the K-CAPP-IRS and K-PPI-R total scores was significant but only moderate in magnitude, r = .350 (see Table 23). Also, between K-CAPP-IRS total score and K-PPI-R factor scores, there were significant correlations (r = .481 between K-CAPP-IRS total scores and Self-Centered Impulsivity, r = .340 between K-CAPP-IRS total score and Coldheartedness). Finally, there were significant correlations between the K-CAPP-IRS total scores and the K-PPI-R content scale scores. Namely the K-CAPP-IRS total score was correlated with the Rebellious Nonconformity (RN), r = .331; Blame Externalization (BE), r = .472; Carefree Nonplanfulness (CN), r = .287; and Coldheartedness (C), r = .340.

In addition, the associations between each of the K-CAPP-IRS domains and the K-PPI-R total score, factor scores, and content scale scores were also examined. Attachment was significantly correlated with Rebellious Nonconformity, r = .399; Blame Externalization, r = .460; Carefree Nonplanfulness, r = .363; Coldheartedness, r = .287; the Self-Centered Impulsivity factor, r = .487); the Coldheartedness factor, r = .287; and the K-PPI-R total score, r = .311. Behavioural was significantly correlated positively with Rebellious Nonconformity (r=.228) with Blame Externalization (r=.341), with Carefree Nonplanfulness (r=.339), with Coldheartedness (r=.262), with Self-Centered Impulsivity factor (r=.401), Coldheartedness factor (r=.262), as well as with the K-PPI-R total (r=.287). Cognitive was significantly correlated with Rebellious Nonconformity, r = .356; Blame Externalization, r = .485; Carefree Nonplanfulness, r = .375; Coldheartedness, r = .387; Self-Centered Impulsivity factor, r = .541; Coldheartedness factor, r = .387; and K-PPI-R total, r = .333. Dominance was not significantly correlated with K-PPI-R scales except Blame Externalization, r = .286. Emotional was significantly correlated with Rebellious Nonconformity, r = .245; Blame Externalization, r = .426; Carefree Nonplanfulness, r = .297; Social Influence (negatively), r = -.299; Coldheartednes, r = .334; Self-Centered Impulsivity factor, r = .412); and Coldheartedness factor, r = .334. Self was significantly correlated positively with Rebellious Nonconformity, r = .266; Blame Externalization, r = .353; Coldheartedness, r = .274; Self-Centered Impulsivity factor, r = .342; Coldheartedness factor, r = .274; and PPI-R total, r = .333.

Table 23. Bivariate Correlations Between the K-PPI-R and K-CAPP-IRS

K-PPI-R	K-CAPP-IRS Domains						
Scales	A	B	C	D	E	S	Total
ME	.261*	.196	.263*	.082	.156	.183	.228*
RN	.399**	.228*	.356**	.154	.245*	.266*	.331**
BE	.460**	.341**	.485**	.286*	.426**	.353**	.472**
CN	.222*	.339**	.375**	.042	.297**	.132	.287**
SOI	-.285**	-.192	-.311**	-.017	-.299*	-.018	-.218
F	-.010	-.022	-.045	-.014	-.002	-.033	.000
STI	-.168	-.156	-.257*	-.066	-.099	-.049	-.159
C	.345**	.262*	.387**	.141	.334**	.274**	.340**
SCI_Factor[a]	.487**	.401**	.541**	.207	.412**	.342**	.481**
FD_Factor[b]	-.198	-.133	-.261*	-.038	-.171	-.007	-.157
C_Factor[c]	.287**	.262**	.387**	.141	.334**	.274**	.340**
Total[d]	.311**	.287*	.333**	.167	.284*	.333**	.350**

Note: * p < .05, ** p < .01, *** p < .001 (2-tailed); significant correlations between CAPP total and PPI-R total are in bold; a SCI=ME+RN+BE+CN; b FD=SOI+F+STI; c C=C; d Total: ME+RN+BE+CN+SOI+F+STI+C; ME=Machiavellian Egocentricity: RN=Rebellious Nonconformity: BE=Blame Externalization: CN=Carefree Nonplanfulness: SOI=Social

Influence: F=Fearlessness: STI=Stress Immunity: C=Coldheartedness: SCI=Self-Centered Impulsivity: FD=Fearlessness: C=Coldheartedness; all ratings are obtained at the first assessment occasion.

As evident in Table 24, there were positive correlations between K-PPI-R facets and the K-PCL-R facets, but on the whole, they were not large. Specifically, K-PCL-R total scores were positively correlated with: RN, r = .255; BE, r = .254; CN, r = .201; SCI, r = .335; C, r = .221; and K-PPI-R total scores, r = .264. Antisocial (Facet 4) scores in the K-PCL-R were only correlated with K-PPI-R total scores, r = .303.

Table 24. Bivariate Correlations Between the K-PCL-R and K-PPI-R

K-PCL-R	K-PPI-R											
Facet	ME	RN	BE	CN	SOI	F	STI	C	SCI	FD	C	Total
Facet 1	.050	.018	.165	-.167	.072	-.113	-.042	-.112	.025	.031	-.112	-.028
Facet 2	.105	.167	.366**	.211**	-.226	-.089	-.138	.172	.295**	.187	.172	.142
Facet 3	.153	.188	.197	.240*	-.280**	.150	-.143	.147	.282*	.113	.147	.184
Facet 4	.218	.211*	.065	.204	-.031	.215*	-.036	.236*	.248*	.061	.236*	.303**
Total	.221	.255*	.254*	.201*	-.164	.095	-.125	.221*	.335**	.079	.221*	.264*

Note. * p < .05, ** p < .01, *** p < .001 (2-tailed); see Table 22 for complete names of scales; all ratings are obtained at the first assessment occasion.

Overall, as expected based on previous literature (Cook et al., 2013), Facet 1 (Interpersonal) was correlated most highly with Dominance and Self; Facet 2 (Affective) with Attachment and Emotional; Facet 3 (Lifestyle) with Behavioural and Cognitive; and Facet 4 (Antisocial) with Behavioral. Of the K

-PPI-R scales, CAPP domains had positive abd large correlations with SCI factor and conversely negative associations with FD factor. Generally, of the K-CAPP domains, Cognitive and Self had somewhat larger correlations with others scales, but the differences were small.

Chapter 4. General Discussion

The issue of cross-cultural validity of forensic assessment instruments, including measures of PPD and actuarial risk assessment tools, is currently a hot topic in the field (e.g., Cooke et al., 2012; Olver et al., 2016; Grisso, 2016). For example, in Canada there is concern about the cross-cultural validity of forensic assessment instruments for use with Aboriginal offenders, as evident in the case of Ewert v. Canada (2015) that recently was appealed to the Supreme Court of Canada (e.g., Hart 2016; Olver et al., 2016). Even in Korea, there is concern about the use of Korean actuarial risk assessment tools (e.g., the KORAS-G; Lee & Ko, 2011) for use with foreign-born offenders, such as South African migrant workers, although these concerns were dismissed by the Korean Supreme Court.

The current study is important as it adds to the growing body of literature examine the nature and extent of cross-cultural variance with respect to forensic assessment instruments, and in particular it is the first study to assess cross-cultural invariance in an Eastern country, namely, Korea. Overall, the results provided some support for the cross-cultural invariance and potential clinical utility of Korean translations of the CAPP model of PPD and the CAPP-IRS measure of PPD. The main findings are discussed below.

4.1. Prototypicality Analysis

The aim of the first study was to evaluate the content validity of the K-CAPP model through the use of prototypicality analysis. The primary findings were these. First, the large majority of K-CAPP symptoms were evaluated as prototypical of PPD, and on average more prototypical than the foils. Second, with respect to K-CAPP domains, Attachment, Dominance, and Self were judged as more prototypical of psychopathy than Emotional, Behavioral, and Cognitive, which is reasonably consistent with Western cultures. Third, the prototypicality ratings made by experts were generally quite similar to those made by lay people. Finally, the prototypicality ratings made by experts and lay people were comparable to those made by experts and lay people in Western countries. Overall, the findings are consistent with the cross-cultural generalizability of the CAPP concept map of PPD in Korea, and also support the validity of the lexical hypothesis underlying the development of the CAPP model insofar as it was possible to translate the CAPP model into Korean.

The current study suggested that Korean sample perceived four K-CAPP symptoms (Unempathic, Lacks Emotional Depth, Self-Centered, and Lacks Remorse) as highly prototypical symptoms of PPD, with mean scores around 6-point. The same results in English version (Robinson, 2017; Kreis et al., 2012; Lim et al., 2014; Murray, 2014), German (Hanna et al., 2018), Sweden (Sörman et al., 2014; Pauli et al., 2018), Spanish (Florez et al., 2015) and Norwegian (Hoff et al., 2012) were also found. The researchers discriminated the symptoms' ratings according to the two groups (experts versus lay people) and then compared the ratings to each other. It revealed that the groups (experts versus lay people) perceived a prototypical psychopath as self-absorbed and egocentric, dominant and manipulative, callous and unattached, and especially being swallow and superficial in emotions. In terms of the domains, the both groups regarded the symptoms from the Attachment, Dominance, and Self domains as being the most prototypical of PPD.

On the whole, these findings were compatible with Western studies (Kreis et al., 2012; Florez et al., 2015; Hoff et al., 2014; Robinson, 2017; Edens et al., 2013a; Robison, 2017; Murray, 2014; Lim et al., 2014 and so forth) and thus could contribute to the fact that the CAPP symptoms allow for better generalizability across different cultures, languages, and nationalities (Sörman et al., 2014; Stoll et al., 2011; Lim, 2016). Still, in this study, some of the K-CAPP symptoms were considered to be less prototypical characteristic of PPD. The groups (expert and lay people) in participants viewed Lacks Perseverance, Lacks Planfulness, Restless, Lacks Concentration, and Lacks Pleasure as least central to the concept of PPD. The Western studies also showed that Lacks Planfulness, Lacks Perseverance, and Lacks Pleasures were viewed as least central concept (Robinsion, 2017; Smith et al., 2014; Kreis et al. 2012) and similar studies by Kreis et al. (2012) (International sample), Hoff et al. (2012) (Norwegian sample), and Sorman et al. (2014) (Swedish sample) suggested that Lacks Pleasure and Lack Concentration turned out be least prototypical of PPD. The domains such as Behavioral, Cognitive, and Emotional were considered to be least prototypical of PPD, which is also very in accordance with previous studies (Hoff et al., 2012; Kreis et al., 2012; Sörman et al., 2014; Robinson, 2017; Lim et al., 2014). However, some of CAPP symptoms also have cultural difference with other Western societies, for example, the Unstable self-concept was not the least prototypical symptom for the Self domain, rating moderate prototypical in our sample which is regardless of expertise. Also Lacks anxiety in the

Emotional domain was to be moderate prototypical which is also different from the Western countries.

The significant differences in 16 symptoms and 4 domains by expertise were found. Compared to experts, the lay people rated the foil symptoms Perfectionistic and Strange as more (at least moderate) prototypical of psychopathy, which is similar to ratings from clinical staffs in the study Hoff et al. (2012) and Sörman et al. (2014) as well as from lay people in the study by Robinson (2017). As expected from previous studies, experts made protoypicality ratings that tended to discriminate or differentiate K-CAPP symptoms from foils to a greater extent than did the prototypicality ratings made by lay people.

What can our findings add to the debate on what is and what is not prototypical of psychopathy? First, Cleckley (1988) and many scholars agreed with findings that the central conception of psychopathy consists of lacks of remorse, pathological egocentrism, and emotional deficiency than antisocial and criminal behavior (Blackburn, 2007; Lilienfeld, 1994; Skeem & Cooke, 2010). In this perspective, our findings could support this idea as the Attachment, Dominance, and Self domains were echoed by two samples as a better way to conceptualize core psychopathy than the Behavioral domain which is more related to antisocial and behavioral problems.

Through the least typical symptoms (Lacks perseverance and Reckless in Behavioral domain, Lacks concentration and Lacks planfulness in Cognitive domain, Garrulous in Dominance domain) which are strongly correlated with the construct of impulsivity (see Florez et al., 2015; Baratt, Stanford, Kent, & Alan, 1997), it also suggests that impulsivity might not be the essential feature of psychopathy that it has been seen as (Hart & Dempster, 1997; Poythress, & Hall, 2011). Interestingly, Lacks pleasure from the Emotional domain was very marginally perceived as prototypical. This is in line with first studies to most current studies through various cultures (Kreis et al., 2012; Hoff et al., 2012). Thus, the lack of prototypicality of the symptoms might suggest that the participants in the current study do not see people with PPD as being completely without emotion (Reidy, Zeichner, & Seibert, 2011; Blair & Mitchell, 2009). Instead, considering Lacks Anxiety which is relatively less prototypical of PPD, they’re more likely to feel improperly or wrongly.

I examined concurrent validity, reliability of the K-CAPP-IRS, and test-retest validity within the frame of G theory for the purpose of its practical utility. Base on the results, it could be suggested that the suitability of K-CAPP-IRS as a whole for assessment of PPD seems to be clear. Main observations of the findings drawn from this study were followed below.

4.2. Concurrent Validity of the K-CAPP-IRS

The concurrent validity of the K-CAPP-IRS was evaluated by taking into consideration the extent to which various aspects of psychopathy as measured by the K-PPI-R and K-PCL-R relate to psychopathy as measured by the K-CAPP-IRS. Overall, the results support the concurrent validity of K-CAPP-IRS and those measures (Robinson, 2017; Cook et al., 2013; Cook et al., 2013; Nikolova, 2009; Lilienfeld & Andrews, 1996; Cook et al., 2013). The K-CAPP-IRS total score was associated with the Self-Centered Impulsivity (SCI), Coldheartedness (C) factor and total score on the K-PPI-R. Generally, positive associations were revealed between K-CAPP-IRS domains and Rebellious Nonconformity (RN), Blame Externalization (BE) and Coldheartedness (C) content scales. There are also correlations between five CAPP domains and Self-Centered Impulsivity (SCI) and Coldheartedness (C) factor as measured by the K-PPI-R. Also, there were no significantly positive correlations between the domains of the K-CAPP-IRS and the Fearlessness Dominance (FD) factor in comparison to the Self-Centered Impulsivity (SCI) factor.

On the converse, there were some negative correlations between K-CAPP-IRS domains and content scales of the K-PPI-R scale. Namely, three domains (A, C, and E) were negatively correlated with Social Influence (SOI) content scale as well as domain C was negatively associated with Fearlessness Dominance (FD) factor of the PPI-R. Compared to previous studies (e.g., Cook et al., 2013; Nikolova, 2009; Lilienfeld & Andrews, 1996), the correlation pattern with the PPI-R in the current study showed some differences by showing several negative associations with FD factor and SOI content scale of the K-PPI-R. However, according to a few literatures (Lilienfeld et al., 2012; Miller & Lynam, 2012; Lynam & Miller, 2012; Sellbom & Phillips, 2013), it argued against the factor of FD inclusion and rather the factor could be regarded as a “protective factor against psychopathology and a measure of stable extraversion.” Indeed, researchers found out

the individuals with high FD score are more likely to be higher in intelligence and executive functioning, and engage in less criminal behavior (Wall, Sellbom, & Goodwin, 2013).

In terms of K-PCL-R facets and total score, as expected in line with previous studies using PCL: SV and the PCL: YV (Stoll et al., 2011; Watkins et al., 2007; Sandvick et al., 2014; Cook et al., 2013), the correlation between K-PCL-R total score and all the K-CAPP-IRS domains as well as total score were associated. Overall, all of the K-CAPP-IRS domains were correlated with facet 2 (Affective) of the K-PCL-R. Next, among the domains, the Attachment domain, the Cognitive domain, and the Emotional domain were strongly related to this K-PCL-R facet. This may be because these domains included similar items related to affective function and emotional relationship with others. Furthermore, it may indicate that the CAPP model could be emphasized by an affective and emotional feature. There was a weak association between the K-PCL-R facet 4 (Antisocial) and the K-CAPP-IRS total score. This result is generally in accordance with previous research (Stoll et al., 2011), which found lower correlations between CAPP-IRS and PCL factor 2, compared to factor 1. Moreover, this finding is in line with the developers' purpose for the CAPP-IRS, focusing of a measure of psychopathic personality disorder rather than criminal behaviors itself which is suggested by Cooke et al. (2004, 2012) and Stoll et al. (2011). Overall, the high association found in the study indicates that they both assess the same cardinal concept, regardless of how they emphasize antisocial and criminal behavior, which is in accordance with previous studies (Cooke et al., 1999; Hare & Neumann, 2010).

Except for some differences, these findings are generally in accordance with prior research on the CAPP-IRS (Nikolova, 2009). Many studies support for the association between the CAPP-IRS and the scores of other measures of psychopathy, which are the PPI-R (Cook et al., 2013; Cook et al., 2013), TriPM (Cook, Viljoen, Hart, Layden, Murray, & McGinnis, 2013) and the PCL-R (Corrado et al., 2006; Corrado et al., 2007; McCormick et al., 2008; Watkison et al., 2007; Sadvick et al., 2014). The current study showed a strong support for such association.

Additionally, Ray et al. (2009) suggested that PPI-R also has correlation with characteristics of PPD as measured by the PCL-R (Skeem & Lilienfeld, 2004). While some of studies has suggested stronger associations between the PPI-R and the PCL-R (i.e., Factor 1), others have showed stronger associations with Factor 2 (Skeem & Lilienfeld,

2004). For example, the PPI-R factor I (Fearlessness Dominance) was correlated strongly with the PCL-R factor 1, while PPI-R factor II (Self-Centered Impulsivity) was associated strongly with factor 2 of the PCL-R (Derefinko & Lynam, 2006)

4.3. Reliability of the K-CAPP-IRS

Given that there are only rarely studies investigating the interrater reliability of the CAPP-IRS, the finding that the interrater reliability in the current study appeared to be satisfactory is meaningful, in line with a few studies (Corrado et al., 2006). In fact, our findings were achieved more than expected, compared to previous studies (Nikolova, 2009; Perdersen, Kunz, Rasmussen, & Elsass, 2010; Dawson et al., 2012; McCormick, Corado, Hart, & Cohen, 2008; Kreis & Cooke, 2011). According to the guideline by Landis and Koch (1977), it ranged from 'substantial' to 'almost perfect' (range = .81 to 1.00). Moreover, by the criteria suggested in Fleiss (1981), interrater reliability in almost all symptoms (e.g., items) except Restless (ICC1 = .72, ICC2 = .88) was 'excellent' (> .75). This result is somewhat different from McCormick's (2007) findings that the symptom-level interrater reliability was substantially lower than the domain-level interrater reliability. Compared to the PCL-R item reliability (Lynam et al., 2007), our results showed more consistent ICCs range such as over .75.

As previous studies suggested (Vincent, 2002; Lynam et al., 2007; Salekin et al., 2006; Jones et al., 2006; Skeem & Cauffman, 2003), interrater reliability analyses using intraclass correlation coefficients with measures obtained through administration of PCL scales typically ranges between .79 and .98 for total scores. Similarly, in the current study, the K-CAPP-IRS total score interrater reliability was 'excellent' (> .75; Fleiss, 1981) with an overall average ICC1 = .91 and ICC2 = .92. Similarly, the analyses of factor scores of the PCL instruments in the previous studies were between .75 and .95 for the PCL Factor 1 and Factor 2, respectively (Skeem & Cauffman, 2003). The interrater reliability for all six domains of the K-CAPP-IRS were within 'almost perfect' (> .81; Landis & Koch, 1977) along with 'excellent range' (> .75; Fleiss, 1981). This exceeded the PCL-R factor scores' interrater reliability identified by Skeem and Cauffman (2003).

In the current study, the reliability could potentially be strengthened by the use of more comprehensive files, semi-structured interview, systematic information given from the Korean correctional information system (so-called Borami Management System),

CAPP Informant-forms completed by correctional staffs, and daily observation records written up during the interview process. Along with the interview, experts also double-checked (as a fact-check process, using a list of data) whether all have case information at the same or similar level before blind and independent ratings. This could exclude other external variables to contribute to confounding and then could guarantee pure interrater reliability based on the same information. Along with this, all cases were involved in face-to-face contact (interviewing) with the examinee, which is recommended by APA (2013). This is an essential part of the study. Overall, it seems likely that these factors such as raters with similar backgrounds, FTF interviews and the DC (double-check) step may be attributable to more stability in reliability, which is compatible with previous studies (e.g., Jackson, 2016).

One of strengths in the study is the recruitment of professionals representing a breadth of skills, training, and experience accumulated in forensic mental health and correctional psychological treatment services. As already discussed earlier, Hanson et al. (2014) suggested that the level of raters' expertise tends to be positively correlated with the reliability of risk ratings, in line with Smid, Kamphuis, Wever, & Van Beek (2014). Most importantly, participants were new to this type of rating, meaning that the research design captures the all-important "first contact" between clinicians and inmates, which might possibly be missed in other studies (Hart, 2003; Wilson, 2013; Darjee et al., 2016; Watt, 2006). From a methodological perspective, it means that the raters were not yet familiar with participants, and hence could not have been influenced by prior contacts or assessments (Douglas & Belfrage, 2014). Last, we acknowledge how the professional experience has an impact on the pattern of ratings. In fact, Sutherland et al. (2012) pointed out that caution should be exercised when used by less qualified assessors with particular items and in cases of middling complexity and risk. In the current study, all experts were professionally trained in clinical settings for a certain period time (ranging from one to twelve years) and obtained sufficient case information from diverse sources for assessing PPD. In this context, the CAPP-IRS (and K-CAPP-IRS) should be cautiously used in conjunction with various information sources such as court, police, correctional staffs, management system, inmates, counselors and medical doctors.

4.4. Generalizability: Consequences of Different Numbers of Raters and Occasions

It is clear from the G Theory analyses that, notwithstanding the fact that the interrater reliability of the K-CAPP-IRS ratings was quite high when indexed using simple ICC analyses, the number of raters has a substantial impact since the components for the rater facet were very large. Thus, there appears to be a great advantage to using more than one rater. On the other hand, due to the small person-occasion and person-occasion-item, the numbers of occasions has a little effect on the magnitude of the SEM estimated and further the generalizability and dependability coefficients. Thus, it is clear that tolerable values for the estimates require a significant number of rater per person. For example, with one occasion, two raters are required to achieve a generalizability coefficient in excess of .80. For comparability purposes, it is clearly important that scores be stable for occasion to occasion. In particular, it is important that the ratings of items should be stable over time. Presumably the reason why components for the rater facet or interactions (ri, pri) with rater facet across items are not small could be due to either not well-defined scoring manual or not well-trained raters or any unidentified complex interactions. Principally, experts must understand not only what is being tested but also the standards and procedures used to assign scores. In our case, it seems that it is not related to raters' training and expertise because the training included case supervision and sequentially several follow-up educations. Further investigation to find out unidentified complex interactions among facets should be followed.

In summary, the application of G Theory extensively helped to determine the sources of variance. The variance among various interactions in the current study could explain how these variables were interacted and further which types of error variance are likely to affect CAPP-IRS ratings (Brennen, 2001; Ryan, 2016).

4.5. K-CAPP-IRS: Clinical Utility

How can professionals utilize the K-CAPP-IRS in corrections? One approach was to introduce the K-CAPP-IRS to professionals most likely to get benefits from its application, especially who are in charge of treatment programs for individuals with PPD. Avoiding just a site-wide implementation is very important because this could cause a within-group resistance by staffs much less motivated for the K-CAPP-IRS. Correctional

psychologists in the Classification Review Board (CRB) could implement K-CAPP-IRS to classify the level of inmates with PPD for institutional decision-making such as extension, cell allocation, classifying examination for the intervention and additional management strategies (Hubson & Shines, 1998; Rasmussen, Storater, & Levander, 1999; García, Moral, Frias, Valdivia, & Diaz, 2012). Moreover, professionals in prison regularly implement involuntary programs for seriously violent offenders with PPD and the K-CAPP-IRS could provide opportunities to understand not merely how the PPD symptoms were fluctuated in an individual, but also how these symptoms mechanically motivate, disinhibit and destabilize to engage in serious offenses.

Another important issue recently raised in the Korean correction is to testify the efficacy of group-based programs and components. These programs were principally designed based on the both Risk-Need-Responsivity model (RNR; Polaschek, 2012; Taxman & Marlowe, 2006; Andrew, Bonta, Wormith, 2011) and Good Life Model (GLM; Gannon, King, Lockerbie, & Wills, 2011) with diverse mixtures of the well-proven programs (Yoon, 2014) without any cross-cultural validation. Assuring whether these programs (taking a substantial time and effort) function effectively in decreasing the extremity of PPD over certain period is still of great concern. In this perspective, K-CAPP-IRS not only provides 33 symptoms with comprehensive scope of PPD but ensures what is being individually assessed in current functioning as opposed to just average functioning across a lifespan. Thus, K-CAPP-IRS can potentially provide an in-depth picture of symptoms' changes by individuals over a certain period of time, which could be quite pragmatic to examine the efficacy and additional modification needed to components of the programs.

Also, typical measures (e.g., PCL-R) to asses PPD generally take a considerable time and effort to complete the ratings. This lengthy protocol or convoluted guidelines of manuals could lead to assessor's resistance especially under the burden of workload. Thus, the CAPP model has another measure version, the CAPP-Staff Rating Form. In this version, staff (e.g., correction officers, counselors, physicians, and Correctional Rapid Patrol Team officers) can complete ratings quickly based what they know about a specific individual with PPD. It could be also preferable when an inmate refuses to engage in the K-CAPP-IRS interview. The CAPP-Staff form is guided by the same illustrative indicators used in the CAPP-IRS. The ratings between CAPP-Staff form and CAPP-IRS interview were highly correlated (Nikolova, 2009; Cooke, Hart, & Logan, 2012; Cooke, Hart, Michie,

& Logan, 2004). To date, research on the CAPP-Staff Rating Form as an alternative measure of PPD is very limited.

More recently, one of the critical issues in women's penitentiaries is how to assess female offenders with PPD. Professionals in prison have difficulty in understanding prototypical characteristics for women but merely have a clinical limitation in intervention. Even though the PCL-R scale has some symptoms related to female prototype, this doesn't well capture the distinctively psychopathic women (e.g., Unstable self-concept and Lacks emotional stability) (Kries et al., 2011). As women with PPD use of more subtle and relational ways of exploiting and dominating other people in domestic settings against close intimates and dependents, their violent acts are easily missed by the officers (Forouzan & Cooke, 2005; Verona & Vitale, 2006). The PCL-R and derivatives relying strongly on officially recorded criminality and of more male typical presentations are likely to miss a great proportion of psychopathic women. But, the CAPP-IRS, as suggested by Cooke et al. (2004, 2012), does not focus merely on criminal and antisocial behavior and provides a more gender-sensitive conceptualization of psychopathy. This could help professionals to catch the core construct of psychopathy well among psychopathic women.

Finally, as Korean professionals treat and evaluate inmates as a team-base working, K-CAPP-IRS is useful in terms of facilitating them to communicate about their own ratings. The K-CAPP-IRS rating system deals with more comprehensive than that of the PCL tests, which simply allows for the absence, likely presence, and definite presence of a trait. Specifically, seven-point scale (global ratings), two different types of sub-scales (trait strength and functional impairment), and summary table to form the basis of clinical formulation could encourage the mutual interchange of diverse opinions and perspectives. Professionals might understand sequential changes in symptoms over time after treatment programs and it promotes to communicate about why the symptoms' severity decreased or increased, not just echoing presence or absence.

4.6. Limitations

As suggested by several studies (see Robison, 2017; Hoff et al., 2012; Florez et al., 2015), translating the original to the Korean version was a tough process because linguistic bias could potentially not be excluded. Nonetheless, presumably given the high

degree of similarity between other different nations (e.g., Norwegian study versus Spanish: Hoff et al., 2012 and Florez et al., 2012; English versus Spanish: Robinson, 2017) and Korean version, it seems to support the lexical approach to personality used by the CAPP model in Korea.

Perfectionist, Strange, and Shy foils were rated as more typical of PDD by the lay people, which is similar to views of clinical ward staff in the study by Hoff et al. (2012) and Sorman et al. (2014). This could be attributable to the fact that lay people might not have experienced any training on psychopathy or forensic psychological evaluation. Also, those who the mass media portrayed (e.g., Dexter Morgan) and fictional television broadcasted like intelligent serial killers (e.g., Mr. Brooks; On Country for the Old Men) who play 'mind games' with the investigators (e.g., Hannibal Lector) were broadly acknowledged as being psychopathic in terms of Perfectionist and thus these characters might be unconsciously imprinted as a "prototypical psychopath" (Smith et al., 2014; Edens et al., 2013a; Edens et al., 2013 b; Park, 2014; Park & Joe, 2013).

In line with Western samples (Robinson, 2017; Florez, 2015), both expert and lay people samples rated Garrulous as moderately common among psychopaths, but the expert group significantly rated higher than lay people. As Smith et al. (2014) pointed out, this may because the Garrulous may be more difficult for laypersons (e.g., jurors) to understand or capture the conception as a psychopathy in a way to give the item a neutral or leave question blank.

Nonetheless, as Kreis and Cooke (2012) suggested, these items viewed as being less typical among psychopaths should not be necessarily excluded from the CAPP model. Rather should be it more explored or investigated as some of items low rated may be possibly more hard for others to observe or only evident at very high level psychopaths, though clearly relevant to the concept of psychopathy. Even the symptoms viewed as less rated may not be related to the core symptoms of psychopathy, it could be still correlated with the concept characteristics of psychopathy. Future research should be needed to test hypothesis with various samples such as different participants, nations, ages, sex, and cultures.

The rating for the both K-PCL-R and the K-CAPP-IRS were obtained by the different raters and K-PPI-R rating was filled out along with other interview, on average,

spending almost four to five hours to each participant. This may increase serious level of tense, fatigue and low diligence effect, but timely care was given to the participants whenever participants want to take rests between sessions. We also tried to include the female sample, but very small number of female sample may limit the generalizability of the findings. Previous studies pointed out that the psychopathic women could be different for psychopathic men and it means they might have partially different psychopathy construct. Given small number of female sample in the study, it was not possible to determine whether would have an effect on the rating of the K-CAPP symptoms and domain and total score.

In terms of G-theory, we may also think about occasion bias. For example, occasion interval could not be long enough, so the first assessment could may impact on the second assessment in such a way like that interviewers might have confirmation bias or prior talks may lead to implicit or unspoken consensus because the interviewers worked in a same team in a same center in a same prison. Thus, we need to extend the time-interval to 6, 9, 12 months for complete separated session as well as restrain possible connection among raters.

4.7. Future Directions

Studies with more diverse samples and additional languages can provide further support for the content validation of the CAPP-IRS. Future studies should further examine gender differences in the expression of psychopathy, which can add to the research support gender-sensitive items included in the CAPP. To operationalize a gender-neutral psychopathy risk assessment, significant symptoms diagnosing psychopathic men and women should be identified. Korean population seems to share analogous conception of psychopathy with Western samples with few differences. Though our findings support for the lexical hypothesis and in numerous settings, follow-up research should be continued. According to research by Ramstedt (1949; for related studies see Miller, 1984 and Lee, 1978 for details), the Korean language (categorized as Koreanic language; language isolate) is dissimilar linguistically to Iranian (Indo-Iranic), English (Eastern Germanic language), Spanish (categorized as Normanic language) and Norwegian (categorized as North-Germanic language) belongs to Altaic family which consists of Turkish language, Mongolian language, and Manju-Tungus language. Later, Korean as a fourth language was added to the Altaic family. However, English originated from East-Germanic language

has many differences with Koreanic isolate. For example, word order, final and ending, and subject ellipse are representatively different. For these reasons, a degree of linguistic bias and errors in the future researches should be re-investigated.

Through the concurrent validity, the use of a small and limited sample may limit generalizability. Thus replication research with increased sample sizes and number of raters to better analyze interrater reliability in different settings should be necessary before robust conclusions can be confirmed.

The very initial results provided by this study lend strong support to the interrater reliability of the CAPP for total and domain scores. However, in a Canadian setting, the comparatively low interrater reliability for the Self domain was found (McCormick, 2007). Future research should keep on exploring the range of interrater reliability in additional populations, whether adult custody, forensic psychiatric settings, or community samples. Furthermore, the variables to have impact on the level of reliability should be explored. Three raters are clinically trained in correctional settings and jointly participated in the interview process to every case. These variables might be recommended to increase the reliability across raters in clinical and incarcerated settings. Additionally, a few studies approved the predictive utility in the CAPP (Pedersen et al., 2010), nonetheless, the CAPP should be endeavored to address prediction ability of future behaviors such as non-violent and violent crime.

In terms of G-theory, we may also add some other interesting facets. For example, we could classify interview method into several ways like jointly interviews (like in my research), independent interviews, single interview with video-observers, and single interview with blind observers. Along with this, information availability or information type for symptom rating could be considered for adding facets.

On conclusion remarks, as aforementioned, the case of Ewert v. Canada (2015), raised cultural issues for legal challenges to forensic risk assessment and other structured assessments to cultural bias. Many scholars also generated considerable debate concerning the extent to which these tools may be considered valid for use across diverse culture (Hart, 2016). Indeed, especially a big question whether tools originated from Western could be equal to Eastern societies could be the most challenging future problem facing the clinical setting of correctional psychology (Olver et al., 2016; Hart, 2016; Haag,

Boyes, Cheng, & MacNeil, 2016). Perhaps, this study could address the potentiality and limitation for the cross cultural validation and application of forensic assessment instruments, in this case, the CAPP-IRS.

References

Ackerman, T. A. (1992). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *Journal of educational measurement*, 29(1), 67-91.

American Psychological Association. (2013). Specialty guidelines for forensic psychology. *American Psychologist*, 68, 7-19.

Andrews, D. A., Bonta, J., & Wormith, J. S. (2011). The risk-need-responsivity (RNR) model: Does adding the good lives model contribute to effective crime prevention?. *Criminal Justice and Behavior*, 38(7), 735-755.

Assadi, S. M., Noroozian, M., Pakravannejad, M., Yahyazadeh, O., Aghayan, S., Shariat, S. V., & Fazel, S. (2006). Psychiatric morbidity among sentenced prisoners: prevalence study in Iran. *The British Journal of Psychiatry*, 188(2), 159-164.

Barratt, E. S., Stanford, M. S., Kent, T. A., & Alan, F. (1997). Neuropsychological and cognitive psychophysiological substrates of impulsive aggression. *Biological psychiatry,* 41(10), 1045-1061.

Billioud, S. & Joel T. (2015). *The Sage and the People: The Confucian Revival in China*. UK, London: Oxford University Press.

Bishopp, D., & Hare, R. D. (2008). A multidimensional scaling analysis of the Hare PCL-R: Unfolding the structure of psychopathy*. Psychology, Crime & Law*, 14(2), 117-132.

Blackburn, R. (2007). Personality disorder and antisocial deviance: Comments on the debate on the structure of the psychopathy checklist-revised. *Journal of personality disorders*, 21(2), 142.

Blair, R. J. R., & Mitchell, D. G. V. (2009). Psychopathy, attention and emotion*. Psychological Medicine*, 39(4), 543-555.

Boduszek, D., & Debowska, A. (2016). Critical evaluation of psychopathy measurement (PCL-R and SRP-III/SF) and recommendations for future research. *Journal of Criminal Justice*, 44, 1-12.

Cai, L., Thissen, D., & du Toit, S. H. C. (2011). *IRTPRO for Windows* [Computer software]. Lincolnwood, IL: Scientific Software International.

Cheung, F. M., Kwong, J. Y., & Zhang, J. (2003). Clinical validation of the Chinese personality assessment inventory*. Psychological Assessment*, 15(1), 89.

Cho, E., & Lee, S. J. (2008). *Hare Psychopathy Checklist–Revised: Professional guidebook* (2nd ed.) (Translation for the Korean Edition). Hakjisa: Korea Psychological Services.

Choi, Y., Gahng, T., Cho, E. (2014). The role and relevance of the Psychopathy Checklist-Revised in U.S. Courts (2005-2012). *Korean Criminological Review*, 100, 375-414.

Cleckley, H. (1941*). The Mask of Sanity,* 1988. Previous editions copyrighted, 1950(1955), 1964.

Cleckley, H. (1976*). The Mask of Sanity*. 1941. St Louis, Mosby.

Cleckley, H. (1988*). The mask of sanity*. Augusta, GA: Emily S. Cleckley.

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and *Psychological Measurement*, 20, 37-46.

Cooke, A. N., Viljoen, S., Hart, S. D., Layden, B. K., Murray, A. A., McGinnis, C, R. (2013). *Validity of the PPI:SF and the TriPM using the CAPP as a Concept Model.* In 2013 AP-LS annual meeting, Portland, USA.

Cooke, D. J. & Logan, C. (2015). Capturing clinical complexity: Towards a personality-oriented measure of psychopathy*. Journal of Criminal Justice*, 43(4), 262-273.

Cooke, D. J. & Michie, C. (1997). An Item Response Theory evaluation of Hare's Psychopathy Checklist*. Psychological Assessment*, 9, 2-13.

Cooke, D. J. & Michie, C. (1999). Psychopathy across cultures: North America and Scotland compared. *Journal of Abnormal Psychology*, 108, 55-68.

Cooke, D. J., & Michie, C. (2001). Refining the construct of psychopathy: Towards a hierarchical model. *Psychological Assessment*, 13, 171-188.

Cooke, D. J., & Michie, C., Hart, S. D., & Clark, D. A. (2005a). Assessing psychopathy in UK: concerns about cross –cultural Generalizability. *The British Journal of Psychiatry*, 186, 335-341.

Cooke, D. J., Kosson, D. S., & Michie, C. (2001). Psychopathy and ethnicity: Structural, item, and test generalizability of the Psychopathy Checklist—Revised (PCL-R) in Caucasian and African American participants*. Psychological assessment*, 13(4), 531.

Cooke, D. J., Michie, C., & Skeem, J. (2007). Understanding the structure of the Psychopathy Checklist–Revised An exploration of methodological confusion. *The British Journal of Psychiatry*, 190(49), s39-s50.

Cooke, D. J., Michie, C., Hart, S. D., & Clark, D. A. (2004). Reconstructing psychopathy: Clarifying the significance of antisocial and socially deviant behavior in the

diagnosis of psychopathic personality disorder. *Journal of Personality Disorders*, 18, 337-357.

Cooke, D. J., Michie, C., Hart, S. D., & Clark, D. A. (2005b). Searching for the pan-cultural core of psychopathic personality disorder. *Personality and Individual Differences*, 39, 283-295.

Corrado, R. R., Watkison, A. V., Hart, S. D., & Cohen, I. M. (June, 2006). *Comprehensive Assessment of Psychopathic Personality Disorder: Initial results with a serious and violent youth sample*. Paper presented at the 6th annual conference of the International Association for Forenisc Mental Health Society, Amsterdam, the Netherlands.

Corrado, R. R., Watkison, A. V., Hart, S. D., & Cohen, I. M. (June, 2007). *The interrater reliability of the Comprehensive Assessment of Psychopathic Personality Disorder: Initial results with a serious and violent youth sample*. Paper presented at the International Congress of Psychology and Law, Adelaide, Australia.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. *The adapted mind*, 163-228.

Cowen, N. (2010). Comparisons of Crime in OECD countries based on the UN Survey of Crime Trends (UN-CTS). Retrieved from http://www.civitas.org.uk/content/files/crime_stats_oecdjan2012.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.

Darjee, R., Russell, K., Forrest, L., Milton, E., Savoie, V., Baron, E., & Stobie, S. (2016). Risk for Sexual Violence Protocol (RSVP): *A real world study of the reliability, validity and utility of a structured professional judgement instrument in the South East Scotland*. Edinburgh: NHS Lothian Sex Offender Liaison Service

Dawson, S., McCuish, E., Hart, S. D., & Corrado, R. R. (2012). Critical issues in the assessment of adolescent psychopathy: An illustration using two case studies. *International journal of forensic mental health*, 11(2), 63-79.

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20V3: Assessing risk for violence – User guide.* Burnaby, Canada: Mental Health, Law, and Policy Institute, Simon Fraser University.

Douglas, K. S., & Belfrage, H. (2014). Interrater reliability and concurrent validity of the HCR-20 Version 3. *International Journal of Forensic Mental Health*, 13(2), 130-139.

Douglas, K. S., Strand, S., Belfrage, H., Fransson, G., & Levander, S. (2005). Reliability and validity evaluation of the Psychopathy Checklist: Screening Version (PCL: SV) in Swedish correctional and forensic psychiatric samples. *Assessment*, 12(2), 145-161.

Edens, J. F., Clark, J., Smith, S. T., Cox, J., & Kelley, S. E. (2013a). Bold, smart, dangerous and evil: Perceived correlates of core psychopathic traits among jury panel members. *Personality and Mental Health*, 7(2), 143-153. doi: 10.1002/pmh.1221

Edens, J. F., Davis, K. M., Smith, K. F., & Guy, L. S. (2013b). No sympathy for the devil: Attributing psychopathic traits to capital murders also predicts support for executing them. *Personality Disorders: Theory, Research, and Treatment*, 4(2), 175-181.doi:10.1037/a0026442

Embretson, S. E. & Reise, S. P. (2000). *Item Response Theory for Psychologists*. Mahwah, NJ: Lawrence Erlbaum.

Ewert v. Canada, FC 1093 (2015).

Fathi, N. (2004). *Taarof* [Electronic version]. Open democracy. Retrieved September. 24, 2007, from http://www.opendemocracy.net/arts/article_2117.j.

Fiske, A. P., Kitayama, S., Markus, H. R., & Nisbett, R. E. (1998). The cultural matrix of social psychology. In D. T. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 915–981). New York: McGraw-Hill.

Forouzan, E., & Cooke, D. J. (2005). Figuring out la femme fatale: Conceptual and assessment issues concerning psychopathy in females. *Behavioral Sciences & the Law*, 23(6), 765-778.

Gannon, T. A., King, T., Miles, H., Lockerbie, L., & Willis, G. M. (2011). Good lives sexual offender treatment for mentally disordered offenders. *The British Journal of Forensic Practice,* 13(3), 153-168.

García, C. H., Moral, J., Frías, M., Valdivia, J. A., & Díaz, H. L. (2012). Family and socio-demographic risk factors for psychopathy among prison inmates. *The European journal of psychology applied to legal context*, 4(2), 119.

Grisso, T. (2016, March). Finding your way without GPS. In S. L. Broksky *The 2-way psych-law time machine: Looking at 1968 from 2016, and 2016 from 2064*. Symposium conducted at the annual meeting of the American Psychology-Law Society, Atlanta, GA.

Haag, A. M., Boyes, A., Cheng, J., MacNeil, A., & Wirove, R. (2016). An introduction to the issues of cross-cultural assessment inspired by Ewert v. Canada. *Journal of Threat Assessment and Management,* 3(2), 65.

Hall, J. R., Benning, S. D., & Patrick, C. J. (2004). Criterion-related validity of the three-factor model of psychopathy personality, behavior, and adaptive functioning. *Assessment*, 11(1), 4-16.

Hanson, R. K., Harris, A. J., Helmus, L., & Thornton, D. (2014). High-risk sex offenders may not be high risk forever. *Journal of Interpersonal Violence*, 29(15), 2792-2813.

Hare, R. D. (1991). *The Hare Psychopathy Checklist -Revised*. Toronto: Multi-Health Systems.

Hare, R. D. (2003). *The psychopathy checklist revised*. Toronto, ON: Multi-Health Systems.

Hare, R. D., & Neumann, C. S. (2005). Structural models of psychopathy. *Current psychiatry reports,* 7(1), 57-64.

Hare, R. D., & Neumann, C. S. (2010). The role of antisociality in the psychopathy construct: comment on Skeem and Cooke (2010). *Psychological Assessment,* 22(2), 446-454.

Hare, R. D., & Neumann, C. S. (2010). The role of antisociality in the psychopathy construct: comment on Skeem and Cooke (2010*). Psychological Assessment*, 22(2), 446-454.

Hare, R. D., & Neumann, C. S. (2010). The role of antisociality in the psychopathy construct: comment on Skeem and Cooke (2010). *Psychological Assessment,* 22, 446–54. doi:10.1037/a0013635

Harpur, T. J., Hare, R. D., & Hakstian, A. R. (1989). Two-factor conceptualization of psychopathy: Construct validity and assessment implications. *Psychological Assessment: A Journal of consulting and clinical Psychology*, 1(1), 6.

Hart, S. D. (2016). Culture and violence risk assessment: The case of Ewert v. Canada*. Journal of Threat Assessment and Management*, 3(2), 76.

Hart, S. D., & Dempster, R. J. (1997). Impulsivity and psychopathy*. Impulsivity: Theory, assessment, and treatment,* 212-232.

Hart, S. D., Cox, D. N., & R [obert] D. Hare. (1995). *The Hare psychopathy checklist: Screening version (PCL: SV)*. MSH-Multi-Health Systems, Incorporated.

Hashemian, M., Roohani, A., & Karami, T. (2016). Study Of Cross-Cultural Differences And Pragmatic Transfer In English And Persian Compliments. *Modern Journal of Language Teaching Methods*, 6(8), 82A.

Heilbrun, K. (2001). *Principles of mental health assessment*, New York: Kluwer Academic/Plenum

Hildebrand, M., Ruiter, C. D., Vogel, V. D., & Wolf, P. V. D. (2002). Reliability and factor structure of the Dutch language version of Hare's Psychopathy Checklist-Revised*. International Journal of Forensic Mental Health*, 1(2), 139-154.

Hill, C. D., Neumann, C. S., & Rogers, R. (2004). Confirmatory factor analysis of the psychopathy checklist: screening version in offenders with axis I disorders. *Psychological Assessment*, 16(1), 90.

Hobson, J., & Shines, J. O. H. N. (1998). Measurement of psychopathy in a UK prison population referred for long-term psychotherapy. *The British Journal of Criminology*, 38(3), 504-515.

Hofstede, G. (2001). Culture's recent consequences: Using dimension scores in theory and research. *International Journal of cross cultural management*, 1(1), 11-17.

Hofstede, G. (2010). *Geert hofstede*. National cultural dimensions. London: McGraw-Hill.

Hofstede, G., Hofstede, G. J., & Minkov, M. (1991). *Cultures and organizations: Software of the mind (*Vol. 2). London: McGraw-Hill.

Holland, P. W., & Wainer, H. (1993*). Differential item functioning*. Hillsdale, NJ, 137-166.

Huang, Y. H. C. (2011). Favor (renqing): characteristics and practice from a resourced-based perspective. *China Media Research*, 7(4), 34-44.

Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). *Item response theory: Application to psychological measurement*. CA: Dorsey Press.

Jackson, R. L., Rogers, R., Neumann, C. S., & Lambert, P. L. (2002). Psychopathy in female offenders an investigation of its underlying dimensions. *Criminal Justice and Behavior*, 29(6), 692-704.

Jackson, W. & Verberg, N. (2007). *Method: Doing Social Research* (4th ed.). Toronto: Pearson Prentice Hall.

Joe U. K. & Lee, S. J. (2008). *Translated version of The Hare Psychopathy Checklist - Revised. Seoul*, KR: Hakjisa Press.

Johansson, P., Andershed, H., Kerr, M., & Levander, S. (2002). On the operationalization of psychopathy: further support for a three-faceted personality oriented model. *Acta Psychiatrica Scandinavica,* 106(s412), 81-85.

Jones, S., Cauffman, E., Miller, J. D., & Mulvey, E. (2006). Investigating different factor structures of the Psychopathy Checklist: Youth version: Confirmatory factor analytic findings. *Psychological assessment*, 18(1), 33.

Kim, S., & Hagtvet, K. A. (2003). The impact of misspecified item parceling on representing latent variables in covariance structure modeling: A simulation study. *Structural Equation Modeling*, 10(1), 101-127.

Kitayama, S., & Uchida, Y. (2005). *Interdependent agency: An alternative system for action. In Cultural and social behavior: The Ontario symposium* (Vol. 10, pp. 137-164). Lawrence KaErlbaum Mahwah, NJ.

Kreis, M. K., & Cooke, D. J. (2011). Capturing the psychopathic female: A prototypicality analysis of the Comprehensive Assessment of Psychopathic Personality (CAPP) across gender*. Behavioral Sciences & the Law*, 29(5), 634-648.

Kreis, M. K., Cooke, D. J., Michie, C., Hoff, H. A., & Logan, C. (2012). The Comprehensive Assessment of Psychopathic Personality (CAPP): content validation using prototypical analysis. *Journal of Personality Disorders,* 26(3), 402.

Lee, J. S., Lee, D.G., & Wee, H. J. (2015). Investigating Two-, Three-, and Four-factor Structures of Korean PCL-R: Applied across Multiform Criminal Offenders. *Journal of Korean Criminological association,* 9(2), 3-32.

Lee, K., Ashton, M. C., Wiltshire, J., Bourdage, J. S., Visser, B. A., & Gallucci, A. (2013). Sex, power, and money: Prediction from the Dark Triad and Honesty–Humility. *European Journal of Personality*, 27(2), 169-184.

Lee, S. & Park, H. (2009). *Psychopathic Personality Inventory-Revised* (Korean Edition). Seoul: Hakjisa Publiser.

Lee, S. I. (1978). *An Examination of" Studies in Korean Etymology*" by GJ Ramstedt. University Microfilms.

Lee, S. J., & Kim, M. J. (2010). The Study of Construct Validity on the Psychopathy Checklist-Revised (PCL-R).The Korean Version*. Korean Journal of Forensic Psychology*, 1(1), 43-55.

Lee, S. J., Ko, R., & Kim, J. K. (2009). The study of construct validity on the Psychopathy Checklist-Revised (PCL-R): Korean version*. Korean Journal of Social and Personality Psychology,* 23(3), 57-71.

Lee, SJ & Ko, RJ. (2011). A study of KORAS-G on the Validity Evidences of Korean Offender Risk Assessment System*. Korean Criminological Review*, 5(2), 221-258.

Lee, SJ & Park, HY. (2007*). The Guidebook of the PPI-R (Korean version).* Seoul, Korea: Hakjisa Publisher.

Lee, SJ., Gong, JS., & Park, HY. (2007). Recidivism Prediction among Inmates Based on KCPI and PPI-R. *The Korean Association of Police Science Review*, 14. 161-186.

León-Mayer, E., Folino, J. O., Neumann, C., & Hare, R. D. (2015). The construct of psychopathy in a Chilean prison population*. Revista Brasileira de Psiquiatria*, 37(3), 191-196.

Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing psychopathic attributes in a noninstitutionalized population. *Journal of personality and social psychology*, 68(1), 151.

Lilienfeld, S. O. (1994). Conceptual problems in the assessment of psychopathy. *Clinical Psychology Review*, 14(1), 17-38.

Lilienfeld, S. O., Waldman, I. D., Landfield, K., Watts, A. L., Rubenzer, S., & Faschingbauer, T. R. (2012). Fearless dominance and the US presidency: implications of psychopathic personality traits for successful and unsuccessful political leadership. *Journal of personality and social psychology*, 103(3), 489.

Lilienfeld, S. O., Widows, M. R., & Staff, P. A. R. (2005). Psychopathic Personality Inventory TM-Revised. *Social Influence (SOI)*, 61(65), 97.

Lim, Y. L. (2016). *Cross-cultural Generalizability of Psychopathic Personality Disorder: Differences Between Individualistic Versus Collectivistic Cultures* (Master dissertation, Arts & Social Sciences: Department of Psychology).

Littlejohn, R. (2010*). Confucianism: An Introduction*. IB Tauris.

Liu, B., Huang, X., & Lv, X. (2010). Fan zui jing shen bing tai de chu bu tan suo [A preliminary research on the psychopathy of criminals]. *Xin Li Ke Xue*, 33, 223-225.

Lynam, D. R., & Miller, J. D. (2012). Fearless dominance and psychopathy: A response to Lilienfeld et al.

Lynam, D. R., & Miller, J. D. (2012). Fearless dominance and psychopathy: A response to Lilienfeld et al.

Lynam, D. R., & Widiger, T. A. (2007). Using a general model of personality to identify the basic elements of psychopathy. *Journal of personality disorders*, 21(2), 160.

Lynam, D. R., Caspi, A., Moffitt, T. E., Loeber, R., & Stouthamer-Loeber, M. (2007). Longitudinal evidence that psychopathy scores in early adolescence predict adult psychopathy. *Journal of abnormal psychology*, 116(1), 155.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological review*, 98(2), 224.

Markus, H. R., & Kitayama, S. (1994). A collective fear of the collective: Implications for selves and theories of selves*. Personality and Social Psychology Bulletin*, 20(5), 568-579.

Markus, H. R., Kitayama, S., & Heiman, R. J. (1996). Culture and" basic" psychological principles. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles*. (pp. 857-913). New York: Guilford.

Marshall, L. A., & Cooke, D. J. (1999). The childhood experiences of psychopaths: A retrospective study of familial and societal factors. *Journal of Personality Disorders*, 13(3), 211-225.

Masui, K., Fujiwara, H., & Ura, M. (2013). Social exclusion mediates the relationship between psychopathy and aggressive humor style in noninstitutionalized young adults*. Personality and Individual Differences*, 55(2), 180-184.

Matsumoto, D., Kudoh, T., Scherer, K., & Wallbott, H. (1988). Antecedents of and reactions to emotions in the United States and Japan. *Journal of Cross-Cultural Psychology,* 19(3), 267-286.

McCord, W., & McCord, J. (1964). *The psychopath: An essay on the criminal mind*.

McCormick, A. (2007). *Interrater reliability of the comprehensive assessment of psychopathic personality disorder among a sample of incarcerated serious and violent young offenders* (Doctoral dissertation, School of Criminology-Simon Fraser University).

McCormick, A. V., Corrado, R. R., Hart, S. D., & Cohen, I. M. (2008*). Interrater reliability and internal consistency reliability of the CAPP-IRS among incarcerated young offenders*. In 8th annual conference of the International Association of Forensic Mental Health Services, Vienna, Austria.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum

Michie, C., & Cooke, D. J. (2005). The Structure of Violent Behavior A Hierarchical Model*. Criminal Justice and Behavior*, 33(6), 706-737.

Miller, R. A. (1984). Korean and Altaic. *Journal of Korean Studies*, 5(1), 143-171.

Millsap, R. E., & Everson, H. T. (1993). Methodology review: Statistical approaches for assessing measurement bias*. Applied psychological measurement*, 17(4),297-334.

Mokros, A., Neumann, C. S., Stadtland, C., Osterheider, M., Nedopil, N., & Hare, R. D. (2011). Assessing measurement invariance of PCL-R assessments from file reviews of North American and German offenders. *International journal of law and psychiatry,* 34(1), 56-63.

Morana, H. C. P., Arboleda-Flórez, J., & Câmara, F. P. (2005). Identifying the cutoff score for the PCL-R scale (psychopathy checklist-revised) in a Brazilian forensic population. Forensic science international, 147(1), 1-8.

Müller, J. L. (2010). Psychopathy—an approach to neuroscientific research in forensic psychiatry. *Behavioral sciences & the law*, 28(2), 129-147.

Muthén, L. K., & Muthén, B. O. (2012*). Mplus user's guide*: Statistical analysis with latent

Neal, T. M., & Sellbom, M. (2012). Examining the factor structure of the hare self-report psychopathy scale*. Journal of personality assessment*, 94(3), 244-253.

Neumann, C. S., Hare, R. D., & Newman, J. P. (2007). The super-ordinate nature of the psychopathy checklist-revised. *Journal of personality disorders*, 21(2), 102.

Neumann, C. S., Hare, R. D., & Pardini, D. A. (2015). Antisociality and the construct of psychopathy: Data from across the globe. *Journal of personality*, 83(6), 678-692.

Neumann, C. S., Johansson, P. T., & Hare, R. D. (2013). The Psychopathy Checklist-Revised (PCL-R), low anxiety, and fearlessness: A structural equation modeling analysis*. Personality Disorders: Theory, Research, and Treatment*, 4(2), 129.

Neumann, C. S., Schmitt, D. S., Carter, R., Embley, I., & Hare, R. D. (2012). Psychopathic traits in females and males across the globe. *Behavioral sciences & the law,* 30(5), 557-574.

Ng, F. F. Y., Pomerantz, E. M., & Deng, C. (2014). Why are Chinese mothers more controlling than American mothers? "My child is my report card". *Child development,* 85(1), 355-369.

Nikolova, N. (2009). *Comprehensive assessment of psychopathic personality disorder-institutional rating scale (CAPP-IRS)–validation* (Doctoral dissertation, Dept. of Psychology-Simon Fraser University).

Olver, M. E., Sowden, J. N., Kingston, D. A., Nicholaichuk, T. P., Gordon, A., Christofferson, S. M. B., & Wong, S. C. P. (2016). Predictive accuracy of Violence Risk Scale–Sexual Offender Version Risk and change scores in treated Canadian Aboriginal and Non-Aboriginal sexual offenders. *Sexual Abuse: A Journal of Research and Treatment. Advance online publication.* doi: 10.1177/1079063216649594

Osumi, T., Kanayama, N., Sugiura, Y., & Ohira, H. (2007). Validation of the Japanese primary and secondary psychopathy scales. *The Japanese Journal of Personality*, 16, 117-120.

Park, J. (2014). People's perception and fear of Psychopaths. *Korea Criminological Review*, 8(2), 145-175.

Park, J., & Joe, J. (2013). A Content Analysis of Media Reports on Psychoapths. *Korean Police Research*, 12(4), 175-200

Patrick, C. J. (2014). *Triarchic Psychopathy Measure* (TriPM).Y. Shou, M. Sellbom, & Q. Liu (Trans.). Retrieved from https://www.phenxtoolkit.org/index.php?pageLink=browse. protocoldetails&id=121601. (Original work published 2010).

Patrick, C. J. (Ed.). (2005). *Handbook of psychopathy*. New York: Guilford Press.

Patrick, C. J., & Drislane, L. E. (2015). Triarchic model of psychopathy: Origins, operationalizations, and observed linkages with personality and general psychopathology. *Journal of personality*, 83(6), 627-643.

Patrick, C. J., Fowles, D. C., & Krueger, R. F. (2009). Triarchic conceptualization of psychopathy: developmental origins of disinhibition, boldness, and meanness. *Development and Psychopathology*, 21, 913–938. doi:10.1017/S0954579409000492

Paulhus, D. L., Hemphill, J. F., & Hare, R. D. (2012*). Self-report psychopathy scale* (SRP-III). Toronto: Multi-Health Systems.

Pedersen, L., Kunz, C., Rasmussen, K., & Elsass, P. (2010). Psychopathy as a risk factor for violent recidivism: Investigating the Psychopathy Checklist Screening Version (PCL: SV) and the Comprehensive Assessment of Psychopathic Personality (CAPP) in a forensic psychiatric setting*. International Journal of Forensic Mental Health,* 9(4), 308-315.

Polaschek, D. L. (2012). An appraisal of the risk–need–responsivity (RNR) model of offender rehabilitation and its application in correctional treatment. *Legal and criminological Psychology,* 17(1), 1-17.

Poythress, N. G., & Hall, J. R. (2011). Psychopathy and impulsivity reconsidered. *Aggression and Violent Behavior*, 16(2), 120-134.

Ramstedt, G. J. (1949). 1953. Studies in Korean etymology.

Rasmussen, K., Storsæter, O., & Levander, S. (1999). Personality disorders, psychopathy, and crime in a Norwegian prison population. *International Journal of Law and Psychiatry*, 22(1), 91-97.

Reidy, D. E., Zeichner, A., & Seibert, L. A. (2011). Unprovoked aggression: Effects of psychopathic traits and sadism. *Journal of personality*, 79(1), 75-100.

Robinson, R. A. (2017). *What is a Psychopath? Cross-National Perceptions of Psychopathy using the Comprehensive Assessment of Psychopathic Personality (CAPP)* (Doctoral dissertation, Alliant International University).

Roju, N. S., Van der Linden, W. J., & Fleer, P. F. (1995). IRT-based internal measures of differential functioning of items and tests. *Applied Psychological Measurement*, 19(4), 353-368.

Ryan, T. J. (2016). *An Examination of the Interrater Reliability and Concurrent Validity of the Spousal Assault Risk Assessment Guide–Version 3 (SARA-V3)* (Doctoral dissertation, Arts & Social Sciences: Department of Psychology).

Ryder, A. G., Sun, J., Dere, J., & Fung, K. (2014). Personality disorders in Asians: summary, and a call for cultural research. *Asian Journal of Psychiatry*, 7, 86-88.

Salekin, R. T., Brannen, D. N., Zalot, A. A., Leistico, A. M., & Neumann, C. S. (2006). Factor structure of psychopathy in youth: Testing the applicability of the new four-factor model. *Criminal Justice and Behavior*, 33(2), 135-157.

Sandvik, A. M., Hansen, A. L., Johnsen, B. H., & Laberg, J. C. (2014). Psychopathy and the ability to read the "language of the eyes": Divergence in the psychopathy construct. *Scandinavian journal of psychology*, 55(6), 585-592.

Santor, D. A., & Ramsay, J. O. (1998). Progress in the technology of measurement: Applications of item response models. *Psychological Assessment*, 10(4), 345.

Schroeder, M. L., Schroeder, K. G., & Hare, R. D. (1983). Generalizability of a checklist for assessment of psychopathy. *Journal of Consulting and Clinical Psychology*, 51(4), 511.

Sea, J., Hart, D. S., & Lee, S. (2018). The generalizability of PCL-R to Korea, comparing with Canadian data. In preparation.

Sea, J., Lee, S., & Cooke, D. (2017). The generalizability of PCL-R based on IRT: How the psychopath is shaped in Korean Culture?. *Korean Criminological Review*, 28(4), 47-87.

Sellbom, M., & Phillips, T. R. (2013). An examination of the triarchic conceptualization of psychopathy in incarcerated and nonincarcerated samples. *Journal of Abnormal Psychology*, 122(1), 208.

Sellbom, M., Cooke, D. J., & Hart, S. D. (2015). Construct validity of the Comprehensive Assessment of Psychopathic Personality (CAPP) concept map: Getting closer to the core of psychopathy. *International Journal of Forensic Mental Health*, 14(3), 172-180.

Shariat, S. V., Assadi, S. M., Noroozian, M., Pakravannejad, M., Yahyazadeh, O., Aghayan, S., & Cooke, D. (2010). Psychopathy in Iran: A cross-cultural study*. Journal of personality disorders*, 24(5), 676.

Shi, J., Lv, Y., Zhang, C., & Han, C. (2013). Bao li xing wei yu jing shen bing tai de xiang guan xing yan jiu [An examination of the relationship between violent behaviors and psychopathy]. *Guangdong Gongan Keji*, 20(4), 31-33.

Shipley, S., & Arrigo, B. A. (2001). The confusion over psychopathy (II): Implications for forensic (correctional) practice. International *Journal of Offender Therapy and Comparative Criminology*, 45(4), 407-420.

Shou, Y., Sellbom, M., & Han, J. (2016 a). Evaluating the Construct Validity of the Levenson Self-Report Psychopathy Scale in China. *Assessment*, DOI: 10.1177/1073191116637421.

Shou, Y., Sellbom, M., & Han, J. (2016 b). Development and Validation of the Chinese Triarchic Psychopathy Measure. *Journal of Personality Disorders*, 30(4), 436-450.

Shou, Y., Sellbom, M., & Xu, J. (2017). Psychometric Properties of the Triarchic Psychopathy Measure: An Item Response Theory Approach. *Personality disorders: Theory, Research, and Treatment.* DOI: 10.1037/per0000241.

Shou, Y., Sellbom, M., Xu, J., Chen, T., & Sui, A. (2016). Elaborating on the Construct Validity of Triarchic Psychopathy Measure in Chinese Clinical and Nonclinical Samples*. Psychological Assessment*. DOI: 10.1037/pas0000398.

Skeem, J. L., & Cauffman, E. (2003). Views of the downward extension: Comparing the youth version of the Psychopathy Checklist with the Youth Psychopathic Traits Inventory. *Behavioral Sciences & the Law,* 21(6), 737-770.

Skeem, J. L., & Cooke, D. J. (2010a). Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. *Psychological Assessment*, 22, 433-445.

Skeem, J. L., & Cooke, D. J. (2010b). One measure does not a construct make: Directions toward reinvigorating psychopathy research—reply to Hare and Neumann (2010). *Psychological Assessment, 22(2), 455-459.*

Skeem, J. L., & Mulvey, E. P. (2001). Psychopathy and community violence among civil psychiatric patients: Results from the MacArthur Violence Risk Assessment Study. *Journal of Consulting and Clinical Psychology*, 69, 358-374.

Skeem, J. L., Edens, J. F., Camp, J., & Colwell, L. H. (2004). Are there ethnic differences in levels of psychopathy? A meta-analysis. *Law and Human Behavior*, 28, 505-527

Skeem, J. L., Mulvey, E. P., & Grisso, T. (2003). Applicability of traditional and revised models of psychopathy to the Psychopathy Checklist: Screening Version*. Psychological Assessment*, 15(1), 41.

Skeem, J., & Lilienfeld, S.O. (2004). *Psychometric properties of self-report psychopathy measures.* Paper presented at the annual meeting of the American Psychology - Law Society, Scottsdale, AZ.

Smid, W. J., Kamphuis, J. H., Wever, E. C., & Van Beek, D. J. (2014). A comparison of the predictive properties of nine sex offender risk assessment instruments*. Psychological assessment,* 26(3), 691.

Smith, S. F., Watts, A., & Lilienfeld, S. (2014). On the trail of the elusive successful psychopath. *Psychological Assessment*, 15, 340-350.

Smith, S. T., Edens, J. F., Clark, J., & Rulseh, A. (2014). “So, what is a psychopath?” Venireperson perceptions, beliefs, and attitudes about psychopathic personality. *Law and human behavior,* 38(5), 490.

Sohn, J. S. & Lee, S. J. (2014). Investigating Two-, Three-, and Four-Factor Structures of the Korean PCL-R in Serious offenders. *International Journal of Offender Therapy and Comparative Criminology*, 8(2), 1-14.

Sohn, J. S., Webb, D., & Lee, S. J. (2010). Can Korean PCL-R predict implicit aggression among Korean inmates? *Asia Pacific Journal of Police & Criminal Justice*, 8(2), 49-69.

Stark, S., Chernyshenko, O., & Dragsow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Towards a unified strategy. *Journal of Applied Psychology*, 91(1292), 1306.

Stoll, E., Heinzen, H., Köhler, D., & Huchzermeier, C. (2011). *Comprehensive Assessment of Psychopathic Personality (CAPP): Validity and practicability of the German version.* Frankfurt am Main, Germany: verlag für Polizeiwissenschaft.

Storey, J. E., Hart, S. D., Cooke, D. J., & Michie, C. (2016). Psychometric Properties of the Hare Psychopathy Checklist-Revised (PCL-R) in a Representative Sample of Canadian Federal Offenders*. Law and Human Behavior*, 40(2), 136-146.

Suh, E., Diener, E., Oishi, S., & Triandis, H. C. (1998). The shifting basis of life satisfaction judgments across cultures: Emotions versus norms. *Journal of personality and social psychology*, 74(2), 482.

Susser, M. (1973). Causal thinking in the health sciences concepts and strategies of epidemiology. In *Causal thinking in the health sciences concepts and strategies of epidemiology.* Oxford, England: Oxford University Press.

Sutherland, A. A., Johnstron, L., Davidson, K. M., Hart, S. D., Cooke, D. J., Kropp, P. R., Logan, C., Michie, C., & Stocks, R. (2012). Sexual violence risk assessment: An investigation of the interrater reliability of professional judgments made using the Risk for Sexual Violence Protocol. *International Journal of Forensic Mental Health*, 11, 119-133.

Taxman, F. S., & Marlowe, D. (2006). Risk, needs, responsivity: In action or inaction?. *Crime & Delinquency, 52, 3-6.*

The Art of Taarof. (2004, 2007.9.24). *Persian mirro*r. Retrieved August 11, 2010 from http://www.persianmirror.com/colture/distinct/distinct.cfm#art.

Thissen, D. & Steinberg, L. (1997). *A response model for multiple-choice items* (pp. 51-65). New York: Springer New York.

Thissen, D. & Steinberg, L. (2010). Using item response theory to disentangle constructs at different levels of generality*.* In Embretson, Susan E. (Ed), *Measuring psychological constructs: Advances in model-based approaches* (pp. 123-144). Washington, DC, US: American Psychological Association.

Thissen, D. (1991). *Multilog User's Guide* (Version 6). Mooresville, IN: Scientific Software.

Thissen, D., Cai, L., & Bock, R. D. (2010). The nominal item response model. In M. Nering & R. Ostini (Eds.), *Handbook of polychromous item response theory models* (pp. 43-75). New York, NY: Routledge.

Ullrich, S., Paelecke, M., Kahle, I., & Marneros, A. (2003). Categorical and dimensional assessment of" psychopathy" in German offenders. Prevalence, gender differences and age factors. *Der Nervenarzt*, 74(11), 1002-1008.

Uno , T .(1991) *Chinese Ideology*, Translated by H. Park. Seoul, South Korea: Daewon . variables. Los Angeles, CA: Author.

Verona, E., & Vitale, J. (2006). Psychopathy in Women: Assessment, Manifestations, and Etiology. In C. J. Patrick (Ed.), *Handbook of psychopathy* (pp. 415-436). New York, NY, US: Guilford Press.

Vincent, G. M. (2004). *Investigating the legitimacy of adolescent psychopathy assessments: Contributions of item response theory*. (Doctoral dissertation, Dept. of Psychology-Simon Fraser University).

Vitacco, M. J., Neumann, C. S., & Jackson, R. L. (2005). Testing a four-factor model of psychopathy and its association with ethnicity, gender, intelligence, and violence. *Journal of consulting and clinical psychology*, 73(3), 466.

Waller, N.G.,Thompson, J. S. & Wenk, E. (2000). Using IRT to separate measurement bias from true group difference on homogenous and heterogeneous scales: an illustration with the MMPI. *Psychological Methods*, 5,125-146.

Watkins, A. V., Corredo, R. R., Lussier, P., Hart, S. D. and Cohen, I. M. (2007 June) *Comprehensive Assesment of Psychopathic Personality Disorder and the Psychopathy Checklist: Youth Version*. Paper presented at the Seventh International Association of Forensic Menatl Health Services (IAFMHS) Annual Conference, Montreal, Canada.

Weaver, C. M., Meyer, R. G., Van Nort, J. J., & Tristan, L. (2006). Two-, three-, and four-factor PCL-R models in applied sex offender risk assessments. *Assessment*, 13(2), 208-216.

Weisz, J. R. & McCarty, C. A. (1999). Can we trust parents' reports on cultural and ethnic differences in child psychopathology? *Journal of Abnormal Psychology*, 108, 598-605.

Weisz, J. R., Suwanlert, S., Chaiyasit, W., & Walter, B. R. (1987). Over-and undercontrolled referral problems among children and adolescents from Thailand and the United States: the wat and wai of cultural differences. *Journal of Consulting and Clinical Psychology*, 55(5), 719.

Weizmann-Henelius, G., Putkonen, H., Grönroos, M., Lindberg, N., Eronen, M., & Häkkänen-Nyholm, H. (2010). Examination of psychopathy in female homicide offenders—Confirmatory factor analysis of the PCL-R. *International journal of law and psychiatry*, 33(3), 177-183.

Widiger, T. A. E., & Lynam, D. R. (l998). *Psychopathy and the five-factor model of personality*. In T. Millon, E. Simonsen, M. Birket- Smith, & R D. Davies (Eds.), Psychopathy: Antisocial, Criminal and violent behavior (pp. 171-187). New York: Guilford Press.

Widiger, T. A., & Lynam, D. R. (1998). Psychopathy and the five-factor model of personality. *Psychopathy: Antisocial, criminal, and violent behavior*, 171-187.

Wierzbicka, A. (1997). *Understanding cultures through their key words: English, Russian, Polish, German, and Japanese* (Vol. 8). London: Oxford University Press.

Williams, K. M., Paulhus, D. L., & Hare, R. D. (2007). Capturing the four-factor structure of psychopathy in college students via self-report. *Journal of personality assessment*, 88(2), 205-219.

Wilson, C. M. (2013). *Reliability and consistency of risk formulations in assessments of sexual violence risk* (Doctoral Thesis). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 38379)

Wilson, M. J., Abramowitz, C., Vasilev, G., Bozgunov, K., & Vassileva, J. (2014). Psychopathy in Bulgaria: The cross-cultural generalizability of the Hare Psychopathy Checklist. *Journal of psychopathology and behavioral assessment,* 36(3), 389-400

Yamagishi, T., Hashimoto, H., & Schug, J. (2008). Preferences versus strategies as explanations for culture-specific behavior. *Psychological Science*, 19(6), 579-584.

Yokota, K. (2012). The validity of a three-factor model in PPI-R and social dominance orientation in Japanese sample*. Personality and Individual Differences*, 53(7), 907-911.

Yoon, J. (2011). *Treating sexual offenders: an examination of Korean Sex Offender Treatment Programs*. Seoul: Korean Institute of Criminology publisher.

Yuki, N. (2008*). Hare Psychopathy Checklist–Revised PCL-R* (2nd ed., Japanese version). Tokyo, Japan: Kaneko Shobo.

Zhang, W., Zou, H., Wang, M., & Finy, M. S. (2015). The role of the Dark Triad traits and two constructs of emotional intelligence on loneliness in adolescents. *Personality and Individual Differences*, 75, 74-79.

Zhou, J., Witt, K., Xiang, Y., Zhu, X., Wang, X., & Fazel, S. (2016). Violence risk assessment in psychiatric patients in China: A systematic review*. Australian & New Zealand Journal of Psychiatry*, 50(1), 33-45.

Žukauskienė, R., Laurinavičius, A., & Čėsnienė, I. (2010). Testing factorial structure and validity of the PCL: SV in Lithuanian prison population. *Journal of Psychopathology and Behavioral Assessment*, 32(3), 363-372.

Zwets, A. J., Hornsveld, R. H., Neumann, C., Muris, P., & van Marle, H. J. (2015). The four-factor model of the Psychopathy Checklist—Revised: Validation in a Dutch forensic inpatient sample. *International journal of law and psychiatry,* 39, 13-22.