The Journal of Behavioral Science (TJBS)

Original Article

The Short-Term Assessment of Risk and Treatability Scale: Reliability and Validity in a Thai Forensic Psychiatric Inpatient Setting

Weerapong Sanmontree[1] and Peera Wongupparaj[2*]

Author Affiliation

[1] Physician, Galya Rajanagarindra Institute, Department of Mental Health, Ministry of Public Health, Bangkok, Thailand.

[2] Lecturer, Cognitive Science and Innovation Research Unit, College of Research Methodology and Cognitive Science, Burapha University, Thailand.

[*] Corresponding author e-mail: peera.wo@go.buu.ac.th

Abstract

The Short-Term Assessment of Risk and Treatability (START) scale is deemed the most appropriate instrument for assessing risks and for managing violent behaviors due to its balanced perspective between protective factors (i.e., strength) and dynamic risk (i.e., vulnerability). Although several facets of reliability and validity of this tool are strong, its reliability, validity, and implementation in the Asian population lacks investigation. This study aimed to examine inter-rater reliability, internal consistency reliability, construct validity and predictive validity of the START in the Thai context. The participants consisted of 118 Thai forensic psychiatric inpatients hospitalized at a public hospital. Trained mental health professionals (i.e., psychiatrists, forensic nurses, clinical psychologists, social workers, and occupational therapists) assessed each participant across the 20 domains with 40 items of the Thai START. Each domain was rated for both degrees of strength and vulnerability on two three-point scales (0-1-2) format. Inter-rater and internal consistency reliabilities and the construct and predictive validities were used to examine psychometric properties of the scale. The main finding demonstrated the good-to-excellent level of inter-rater reliability, an acceptable level of internal consistency reliability. The data supported two relational constructs (i.e., strength vs vulnerability subscales) of the Thai START. The poor-to-excellent predictability of the Thai START items on violent behaviors was also evident. This study also provided the guideline for implementing the tool in non-Western forensic psychiatric populations. Thus, the Thai START is a promising tool for understanding, predicting, and managing violent/challenging behavior in forensic psychiatric inpatients.

Violent and unwanted behaviors among forensic psychiatric inpatients are leading to certain concerns among families and communities. The Short-Term Assessment of Risk and Treatability (START) is a structured assessment, which was developed in 2004, for addressing short-term or acute risk of violent behaviors in adolescents and forensic psychiatric settings (Nicholls et al., 2006; Webster et al., 2004). This tool for assessing violence risk is widely used and regarded as a comprehensive and practical instrument for clinical risk assessment along with other standardized structured risk/forensic instruments (Cartwright et al., 2018; Hogan & Olver, 2018). For example, a global survey (Singh et al., 2014) was conducted across 44 countries on the use of measures for organizing assessment, guiding treatment plans and interventions, and monitoring violence risk. The result from 2,135 mental health professionals (e.g., psychiatrist, psychologists, and nurses) demonstrate that the Historical, Clinical, Risk Management-20 (HCR-20) ranks first, and START places fifth out of 12 instruments. A gradual increase is observed in the publication on risk assessment, that is, from 2 articles in 2006 to 76 articles in 2022. Furthermore, a survey report (De

Beuf et al., 2021) on forensic clinical staff members regarding the implementation of START finds that the items are easy to complete and facilitating communication among staff members.

Although tools for risk assessment are widely used for identifying high- and low-risk behaviors; determining detention, sentencing, and releasing; and planning and managing treatments, their contribution to non-western populations (i.e., Asia, South America, and Africa) and contexts continue to warrant investigation in terms of the psychometric properties of the scale (Fazel et al., 2012; Kikuchi et al., 2021; Zhong et al., 2021). In addition, more than half of all the literature on START use small samples (less than 100 patients), which may not ensure external validity or adequate power to examine the predictive validity of the scale (Bakker et al., 2016). Alternatively, forensic mental health services in Thailand have been gradually developed since 1971. Nevertheless, these services do not incorporate the use of standardized risk assessment tools in the assessment and management of forensic psychiatric patients (Sanmontree, 2022).

The initiation of implementing and training Thai mental health professionals on the Thai version of the HCR-20 proved the tool useful with acceptable sensitivity and specificity. However, a copyright issue, poorly filed validity in clinical forensic psychiatry, and cultural issues also warranted limitation in terms of use (Thamakosit, 2007; Tully, 2017). Furthermore, the other limitations of the HCR-20 include difficulty in predicting rare violence and ignorance of strengths during the assessment of patients in routine practices (Silva, 2020). With the introduction of the START in western forensic settings, growing evidence demonstrated a comparable validity between HCR-20 and the START in the prediction of violent behaviors among forensic psychiatric patients (Cartwright et al., 2018). The START is one of such instruments that address the strengths and vulnerabilities of patients, assist in their assessment/management, and facilitate effective communication between forensic psychiatric clinicians and officers of the criminal justice systems in Thailand. Thus, the main objective of the current study was to investigate the psychometric properties of the START in Thai context.

Literature Review

This section describes violence risk assessments in the international literature and the START and its psychometric properties. The section concludes with the purposes of the current study, which establishes the specific objectives and hypotheses of this investigation.

Violence Risk Assessments

Presently, more than 200 structured tools can be used for assessing the risk of violent and unwanted behaviors in forensic psychiatric settings (Singh et al., 2014). Many studies indicate practitioners most commonly use the Psychopathy Checklist-Revised and the HCR-20 as risk assessment instruments (Hurducas et al., 2014). The first report on the use of risk assessment tools was in 1928 to determine the success or failure of a population of parolees (Burgess, 1928). Since then, the research on risk assessment and management has proliferated with violence risk identification tools and techniques devised and utilized in many populations and settings.

Systematic methods of assessing risks of violent and unwanted behaviors in patients with mental illnesses have been developed over the last several decades (Monahan, 2008). Moreover, various tools for assessing risks of violent and unwanted behaviors have been devised and introduced into clinical practice (Hvidhjelm et al., 2022). The following six key considerations are deemed beneficial in conducting risk assessment (Heilbrun et al., 2010):

1. Applicability: usefulness of a specialized instrument according to its applicability;
2. Approach: approaches that involve actuarial, structured professional judgment and anamnestic assessment;

3. Context: crucial and various contexts, such as clinical, school/workplace, legal (or quasi-legal), and threat to victims;

4. Parameters: consideration of outcome(s) to predict, frequency of outcome(s), probability or type of risk in particular settings, period of time, and the nature of risk and protective factors related to the appraisal;

5. Population: significant differences in risk assessment across populations, which may vary in terms of risk and protective factors, base rates of violent behaviors, and risk-relevant interventions as well as age, gender, mental health status, and location; and

6. Objective: a contemporary risk assessment and risk management model that includes three related domains, namely, risk, needs, and responsivity.

Short-Term Assessment of Risk and Treatability (START)

Based on the START guideline, experienced raters are required to assess 20 dynamic items in terms of risk or *vulnerabilities* and protective factors or *strengths* (Webster et al., 2009). The inclusion of protective factors is regarded as a critical advantage of this tool in that it may improve the positive perception of forensic psychiatric patients and reduce the unethical and biased portrayal of offenders (Braithwaite et al., 2010; Rogers, 2000). The 20 behavioral domains of START are social skills, relationships, occupational, recreational, self-care, mental state, emotional state, substance use, impulse control, external triggers, social support, material resources, attitudes, medical adherence, rule adherence, conduct, insight, plans, coping, and treatability items.

The START was developed to address short-term or acute risk of violent/challenging behaviors in clinical settings. The short-term risk assessment was intended with a time frame of three months for assessment and prediction (Cartwright et al., 2018). The instrument was designed to aid in the organization of clinical assessments, such as evaluating mental illness, monitoring progress, planning therapeutic intervention, and initiating the process of evaluating future risk to self and others on a regular basis. Thus, scholars recommended that experienced and trained clinicians should conduct assessment using the START, because they are expected to be competent in assessing individuals with mental disorders and to become accustomed to similar or related schemes (Webster et al., 2009).

The START shares a similarity with the HCR-20 in terms of the intention to support forensic mental health teams and researchers in identifying sources of violence risk toward others. Nonetheless, the START improves inter-rater reliability, and its items address issues regarding its application to several settings (i.e., civil, forensic psychiatric, and correctional), potential risks to self and others, and inclusive assessment of risk, positive, and protective factors (Webster et al., 2004, 2009). In addition, a recent finding of Cartwright et al. (2018) indicated that the START vulnerability subscale outperformed other standardized tools for risk assessment, such as the HCR-20 and Static-99R in predicting institutional aggression among sexual offenders.

Psychometric Properties of the START

The extant literature on the psychometric properties of the START in terms of reliability and validity suggested moderate-to-strong inter-rater reliability via intraclass correlation coefficient (ICC) values (Abidin et al., 2013; Lowder et al., 2019; Viljoen et al., 2011; Whittington et al., 2014; Wilson et al., 2010) strong-to-exellent internal consistency reliability for each START subscales, moderate concurrent validity with other standardized risk assessment scales, the HCR-20 (Cartwright et al., 2018; Desmarais et al., 2012; Hogan & Olver, 2018), the Structured Assessment of Violence Risk in Youth (Viljoen et al., 2011), the Structured Assessment of Protective factors (Abidin et al., 2013), and the Dangerousness, Understanding, Recovery and Urgency Manual (Abidin et al., 2013). The predictive validity of the START on violent outcome measures suggested moderate-to-good predictability (Braithwaite et al., 2010; Desmarais et al., 2010; Nicholls et al., 2006; Nonstad et al., 2010; Wilson et al., 2010). However, vulnerability subscale of

the START displayed a better predictability than did the strength subscale; nevertheless, predictive validity as indexed by the AUCs of the strength subscale, improved over time in comparison with that of the vulnerability subscale, which relatively decreased overtime (Quinn et al., 2013).

Construct validity is defined as an integration of empirical evidence and theoretical rationale that influence the interpretation or meaning of test scores (Messick, 1989; Strauss & Smith, 2009). However, research that tested the structural relationships between each item and its subscales and between the two subscales of the START (i.e., strength and vulnerability) at the latent factor level (McArdle, 1996) remains lacking. The significant structural relationships between the two START subscales suggest construct validity (Atkinson et al., 2011; Tavakol & Dennick, 2011), which has not been considered and examined thus far. Several researchers and clinicians supported the increment validity of the START for incorporating the strength subscale into the risk assessment framework (Braithwaite et al., 2010; Desmarais et al., 2012; O'Shea et al., 2016; Paetsch et al., 2019). The concept of incorporating the protective factor or the strength subscale of the START is subject to debate at the fundamental level in terms of whether or not it is a unique entity or is simply the reverse of a risk factor or the vulnerability subscale of the START (Haines et al., 2018). Recently, Whittington et al. (2022) conducted a multinational study on 685 forensic practitioners in Scandinavia and the United Kingdom and suggested the unidimensionality of the START. However, three and four items for the strength and vulnerability subscales, respectively, failed to load on the expected component or subscales, which suggests the unclear construct of the START. This conundrum also influenced a START report on the summed scores of the total scale or the separate summed scores of the subscales (Whittington et al., 2022).

Furthermore, whether or not the two subscales may benefit clinical practices remains unclear, because a few items of the two subscales possessed strong correlation coefficients (Braithwaite et al., 2010). With these strong relationships, certain items may not be distinguishable between the two subscales, such that it would appear to undermine the construct validity of the tool. Hence, the given issue may also influence the interpretation of results and the monitoring and management of patients. Moreover, an evident research gap, which emerged from a systematic review and a meta-analysis on the START, reveals the insufficient application of the START across contexts and populations, especially in Asian populations (Douglas et al., 2017; O'Shea & Dickens, 2014). Hence, several studies demonstrate the lower prevalence of violent behaviors in the context of Asian populations (Kikuchi et al., 2021; Zhou et al., 2016). Thus far, concluding that the reports on the reliability coefficients of the START suggested strong reliability and predictive validity but unclear construct validity is possible. Furthermore, currently little evidence exists to support the use of the current instruments for violence risk assessment among forensic psychiatric patients in Thailand and other Asian countries. As such, investigating the psychometric properties of the Thai version of the START (Thai START) is accordingly a research priority.

Objectives of the Study

Evidence on psychometric properties across cultural contexts, settings, and populations could support the generalizability of a tool. In summary, scholars called for further research prior to envisaging established conclusions regarding the psychometric properties and generalizability of the START across forensic psychiatric populations. Moreover, the lack of a standardized tool for the forensic risk assessment of violent behavior in Thailand among Thai forensic professionals for managing risk in forensic psychiatric patients required an immediate need for the purposes. In this regard, the current study aimed to examine the psychometric properties of the Thai START, that is, inter-rater reliability and internal consistency reliability, among the mental health professionals of 118 Thai forensic psychiatric inpatients (Figure 1) using ICC and Cronbach's alpha (α) values, respectively.

Figure 1
Schematics of Inter-Rater Reliability
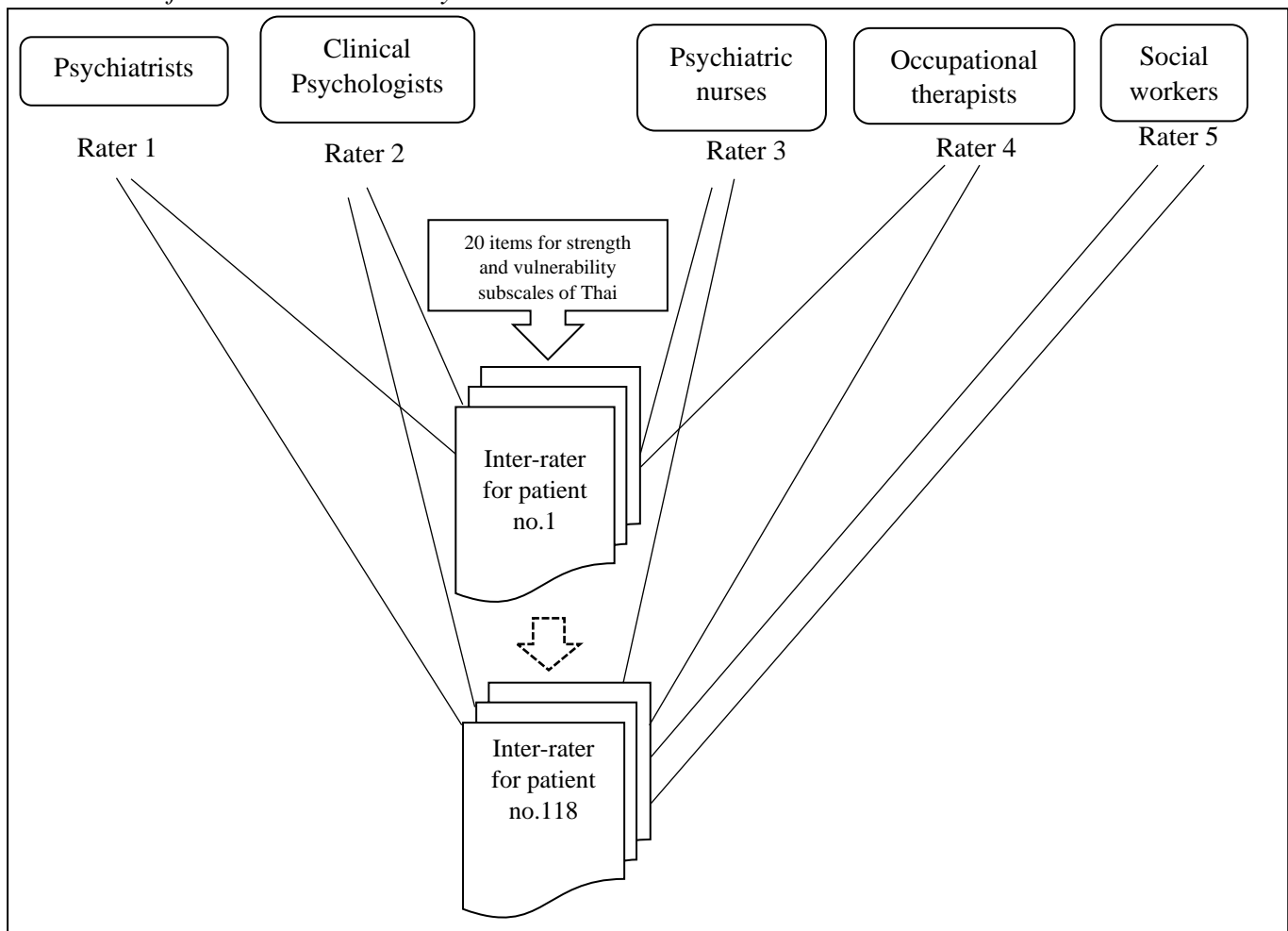
Construct validity was evaluated using ordinal confirmatory factor analysis (CFA) with the weighted least squares mean and variance adjusted (WLSMV) estimator to address a few categories of ordinal data from the scores of the Thai START (DiStefano & Morgan, 2014; Flora & Curran, 2004) (see Figure 2). Moreover, the study examined the predictive validity of the START on violent/challenging behaviors as measured by the START Outcome Scale (SOS; Nicholls et al., 2007) on Thai forensic psychiatric inpatients using the receiver operation curve (ROC). The strong reliability and validity of the Thai START will support its utilization as the first standardized tool for violence risk assessment in forensic psychiatric populations in Thailand. A reliable and valid tool will be of benefit for the management and prevention of violent/challenging behaviors. Thus, the study hypothesizes that the Thai START may yield acceptable reliability and validity in the current context.

Method

Participants

The study recruited a total of 118 participants between January 2014 and January 2016. The participants are forensic psychiatric inpatients hospitalized in the forensic psychiatric unit of the largest forensic psychiatric hospital in Thailand, which serves as the national academic center for forensic psychiatry. Furthermore, this national hospital provided services for forensic inpatients; as such, it could also offer a suitable setting for the observation and examination of a prospective investigation of the START. The required sample size for the present study was 118, which were calculated using four parameters. Specifically, the study considered an expected Area Under the ROC Curve (AUC) of .70 from a previous meta-analysis on the START (O'Shea & Dickens, 2014) together with a power of .80, a base

rate of .17 produced by a pilot study, and an alpha level of .05. The software for calculating the sample size was MedCalc version 12.1.4.0 (Medcalc software, Belgium).

Additionally, the determination of the sample size was considered in line with that of Kyriazos (2018), whose general rule of thumb is 100–200 participants for CFA with multiple indicators (i.e., > 3) per factor or latent variable. The inclusion and exclusion criteria were as follows: eligible age for study (> 18 years) and gender (male and female). The participants were considered eligible if they were hospitalized at any time from June 1, 2015 to June 30, 2016 and excluded if they were admitted to the hospital but discharged after less than two weeks. Notably, however, no participant was excluded from the study on these grounds.

The majority of the patients were male (87.3%), and a large proportion were unemployed (44.1%) followed by employees (30.50) and merchants (8.5%). The major diagnoses were paranoid schizophrenia (37.2%), undifferentiated schizophrenia (17.8%), and adjustment disorder (9.4%). The mean age of the participants, which ranged from 19 to 86 years, was 39.4 years with a standard deviation of 13.1 years.

Instruments

The START consists of 20 dynamic items, namely, social skills, relationships, occupational, recreational, self-care, mental state, emotional state, substance use, impulse control, external triggers, social support, material resources, attitudes, medical adherence, rule adherence, conduct, insight, plans, coping, and treatability. Each item is rated for strength and vulnerability using two sets of three-point scales (0–1–2 format); 0 = no/minimal evidence of vulnerability or strength; 1 = moderate vulnerability/strength; and 2 = high vulnerability/strength (Webster et al., 2009).

The SOS pertains to incident report forms of violent/challenging behaviors among forensic psychiatric inpatients. It consists of 12 items, namely, unauthorized leave, self-neglect, being victimized, self-harm, verbal aggression, aggression against property, stalking, physical aggression against others, sexual aggression, suicide ideation and planning, suicide behaviors, and substance abuse (Nicholls et al., 2007). Each behavior is scored according to frequency during the hospital stay of the participants. Two clinical psychologists, who were blind to the result of the administration of the Thai START, conducts the rating.

Procedure

The procedure is composed of two major parts, namely, translating and adapting the START and administrating and scoring the START and SOS, respectively.

The study adopted the forward and backward translation procedure, which was conducted by Thai senior psychiatrists and a professional native English–Thai bilingual and bicultural translator. Adaptation to the wording was made in terms of the functional equivalence of words (Hui & Triandis, 1985). The content validation of the scale was conducted through the expert judgment of the translators and the research team after reaching consensus. Three additional experts, namely, two consultant forensic psychiatrists from the Department of Forensic and Neurodevelopmental Science of King's College London and one of the authors of the START, Professor Johann Brink, from the University of British Columbia, Canada, rechecked the completed back-translated version.

Descriptive information on socio-demographic characteristics, such as age, gender, marital status, and psychiatric characteristics, such as psychiatric diagnosis, number of previous hospitalizations, and number of visits to the psychiatric emergency unit, were collated from medical records. The potential participants were given an information sheet that included (i) a brief description of the study (objectives); (ii) criteria used to determine eligibility; (iii) a brief list of significant benefits and plausible risks (direct or indirect); (iv) name, address, and workplace of the principal researcher; and (v) location of the research and name of the person to contact for further information.

A psychiatric clinical team composed of seven senior psychiatrists, five clinical psychologists, eighty-five psychiatric nurses, two occupational therapists, and four social workers voluntarily participated and

administrated the Thai START to each patient. The scores of the items for all mental health professionals (maximum of five staff per patient) were derived using the median value. Prior to using the Thai START, all psychiatric clinical staff participated in a START training and workshop in the ward of the hospital for forensic psychiatric inpatients in Bangkok. The training modules included violence risk assessment, START coding and practices, and risk formulation of Thai forensic case studies. For the SOS, trained clinical psychologists, who were not involved in the Thai START administration, coded the number of SOS behaviors that occurred within each item across times 1 (two weeks) and 2 (six weeks) after the completion of the coding for the Thai START. The justification for recording the SOS behaviors at two and six weeks after assessment using the Thai START intended to reflect predictability for two- and six-week ranges of the Thai START (Viljoen et al., 2011).

The study assessed inter-rater reliability or agreement among the psychiatric clinical team (i.e., psychiatrists, forensic nurses, clinical psychologists, social workers, and occupational therapists) using the ICC with a two-way mixed model for ordinal variables with multiple coders (Hallgren, 2012). The ICC was used to compute the inter-rater reliability using test scores with ordinal variables and multiple raters (Fleiss & Cohen, 1973; Marasini et al., 2016). The critical values of the ICC were interpreted using the method of with Cicchetti (1994), where ICC $\geq .75$ = excellent, between .74 and .60 = good, between .59 and .40 = fair, and less than .40 = poor. Cronbach's α coefficients were interpreted according to Tavakol and Dennick (2011), which suggests that acceptable α values should be between .70 and .95. The study used the IBM Statistical Package for the Social Sciences (SPSS) version 27 to analyze descriptive statistics (i.e., median, interquartile range, skewness, and kurtosis), the ICC, and Cronbach's α.

To test construct validity, the underlying structure or construct of the Thai START, that is, the two-factor model with 20 indicators was examined. The strength and vulnerability subscales reflect the two-factor model, where each subscale contains 20 test items. STATA version 15 was used to analyze the polychoric correlation coefficients among test items with ordinal data. The study used Mplus 7.0 to conduct CFA with the robust WLSMV adjusted estimator (DiStefano & Morgan, 2014). CFA model fit was evaluated by multiple fit indices, such as Chi-squared (χ^2) test, χ^2/df, root mean square error of approximation (RMSEA), comparative fit index (CFI), and Tucker–Lewis index (TLI). A value of $\chi^2 > .05$ indicates good fit. The χ^2/df value <2.0 is considered an acceptable fit. The cut-off values for good model fits were adopted as RMSEA $< .08$, and CFI and TLI $> .90$ (Brown, 2015).

The predictive validity or predictability of the Thai START subscales on the SOS across times were computed via AUC values from ROC analysis, which was adopted, because it is reasonably unaffected by the base rates of violent behaviors relative to other methods (Mossman, 1994). AUC values are suitable in the case of the low base rates of violent/challenging behaviors in the current study (Table 2; Chu et al., 2011; O'Shea et al., 2016). SOS was transformed from the frequency of violent/challenging behaviors to present or non-present target behaviors. AUCs represented degrees of separability or predictability between present or non-present violent behaviors for two and six weeks after the Thai START assessment. An AUC value of .50 indicated that the predictive accuracy of the Thai START on the SOS is no better than chance, while an excellent accuracy of the given measures is represented by AUC scores of .90 to 1.00 (Swets, 1986).

Ethical Consideration

Approval from the Institutional Review Board of the Galya Rajanagarindra Institute (IRB No. 17/2557) was obtained prior to conducting data collection. The participants signed consent forms. In addition, the participants were assured regarding the anonymity and confidentiality of their data.

Results

The ICC values across the individual test items of the Thai START displayed good-to-excellent reliabilities (.80 to .99). Item scores were approximately normally distributed as indicated by skewness and kurtosis values between −2 to +2 (Table 1; George & Mallery, 2010). Cronbach's α coefficients exhibited

comparable and acceptable values for the strength (.83) and vulnerability (.81) subscales. The polychoric correlation coefficients among indicators demonstrated strong-to-excellent and negative relationships between the same items of the strength and vulnerability subscales of the Thai START. Nonetheless, the study found null relationships ($p > .05$) for several items within the Thai START subscales. Thus, an overall correlation coefficient was suitable to examine the construct validity using the CFA.

Table 1
ICC, Median (IQR), Skewness, and Kurtosis Across the Thai START Items.

Thai START domains	ICC	Median (IQR)	Skewness	Kurtosis
Strength/Vulnerability				
Social skills (SSS/SSV)	.89/.91	1.00(0.00)/1.00(0.00)	0.31/0.11	1.95/0.28
Relationships (RS/RV)	.80/.90	1.00(0.00)/1.00(0.00)	-0.06/-0.02	1.65/0.25
Occupational (OS/OV)	.89/.94	1.00(1.00)/1.00(1.00)	0.28/-0.36	-0.63/-1.09
Recreational (RCS/RCV)	.88/.94	1.00(0.00)/1.00(0.00)	-0.30/0.03	0.26/-0.51
Self-care (SCS/SCV)	.96/.99	1.00(2.00)/1.00(2.00)	0.04/-0.27	-1.29/-1.62
Mental state (MSS/MSV)	.94/.95	1.00(1.00)/1.00(1.00)	0.44/-0.78	-0.70/-0.59
Emotional state (ESS/ESV)	.92/.94	1.00(0.00)/1.00(1.00)	0.02/-0.17	-0.28/-0.81
Substance use (SUS/SUV)	.93/.95	1.00(1.00)/1.00(1.00)	-0.16/0.04	-0.64/-1.06
Impulse control (ICS/ICV)	.93/.96	1.00(0.00)/1.00(1.00)	-0.02/0.01	-0.21/-0.99
External triggers (ETS/ETV)	.91/.93	1.00(0.00)/1.00(0.00)	0.01/-0.03	-0.07/-0.61
Social support (SOSS/SOSV)	.95/.94	1.00(0.00)/1.00(0.00)	-0.05/-0.02	-0.61/-0.61
Material resources (MRS/MRV)	.94/.94	1.00(1.00)/1.00(1.00)	0.24/-0.23	-0.56/-0.95
Attitudes (AS/AV)	.88/.95	1.00(1.00)/1.00(1.00)	-0.26/0.53	-0.75/0.08
Medical adherence (MAS/MAV)	.85/.93	1.00(1.00)/1.00(1.00)	0.32/-0.35	-1.25/-1.75
Rule adherence (RAS/RAV)	.92/.98	1.00(1.00)/1.00(1.00)	1.09/-0.36	-0.76/-0.85
Conduct (CS/CV)	.93/.97	1.00(1.00)/0.00(1.00)	0.81/0.36	-0.47/-0.65
Insight (IS/IV)	.92/.92	1.00(1.00)/1.00(1.00)	-0.14/0.56	-0.89/-0.64
Plans (PS/PV)	.93/.92	1.00(1.00)/1.00(1.00)	-0.58/-0.12	-0.74/-0.59
Coping (COS/COV)	.92/.89	1.00(1.00)/1.00(1.00)	0.28/0.18	-1.18/-1.96
Treatability (TS/TV)	.89/.88	1.00(0.00)/1.00(0.00)	0.18/-0.05	0.55/0.53

Figure 2 demonstrates the CFA result of the two-factor model that represents the two subscales (i.e., strength vs. vulnerability) of the Thai START with 40 indicators. The overall fit indices suggested an acceptable fit (N = 118; $\chi^2 = 1205.15$, $df = 706$, $p < .01$; $\chi^2/df = 1.71$; RMSEA = .07, CFI = .90, TLI = .90). All factor loadings were statistically significant ($ps < .05$ and .01), except for SUV and ETV for the vulnerability subscale and SUS and ETS for the strength subscale. The relationship between strength and vulnerability was $-.77$ ($p < .01$).

For SOS items at time 1, the most frequently recorded violent/challenging behaviors of the 118 inpatients were self-neglect followed by verbal aggression and physical aggression against others. The study did not observe being victimized, stalking, and suicide behaviors. For time 2, self-neglect was recorded as the most frequent violent/challenging behaviors for 93 inpatients followed by sexual aggression and physical aggression against others. However, stalking was not reported during the hospital stay (Table 2).

The AUCs for the strength subscale of the Thai START were generally between .52 (substance abuse) and .76 (unauthorized leave) at time 1 and .53 (suicide behavior) and .99 (substance abuse) at time 2. These AUC values indicated poor-to-moderate and poor-to-excellent predictabilities of the strength subscale at times 1 and 2, respectively. Furthermore, the AUCs for the vulnerability subscale of the Thai START ranged from .58 (self-harm) to .85 (suicide ideation and planning) at time 1 and .64 (sexual aggression) to .99 (being victimized) at time 2. Thus, it is suggested that the AUCs for the vulnerability subscale of the Thai START demonstrated poor-to-moderate and poor-to-excellent predictabilities at times 1 and 2, respectively.

Figure 2

Path Diagram of the Bi-Factor Model with 40 Indicators, and All Factor Loading Coefficients are Shown in Standardized Scores (ns = non-significant)
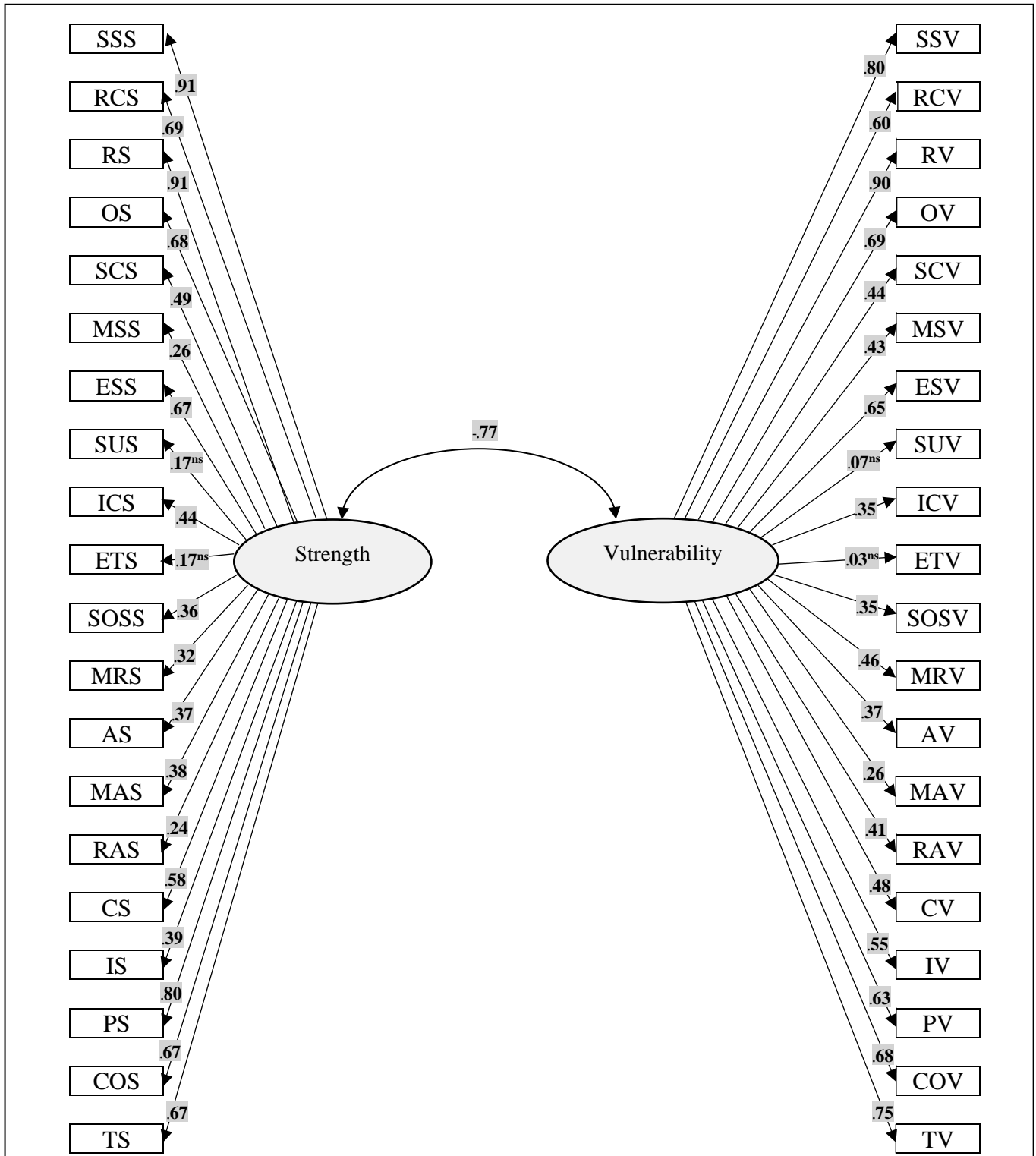
Note. SSS/SSV = Social Skill Strength/Vulnerability, RS/RV = Relationship Strength/Vulnerability, OS/OV = Occupational Strength/Vulnerability, RCS/RCV = Recreational Strength/Vulnerability, SCS/SCV = Self- Care Strength/Vulnerability, MSS/MSV = Mental State Strength/Vulnerability, ESS/ESV = Emotional State Strength/Vulnerability, SUS/SUV = Substance Use Strength/Vulnerability, ICS/ICV = Impulse Control Strength/Vulnerability, ETS/ETV = External Triggers Strength/Vulnerability, SOSS/SOSV = Social Support Strength/Vulnerability, MRS/MRV=Material Resources Strength/Vulnerability, AS/AV=Attitude Strength/Vulnerability, MAS/MAV = Medical Adherence Strength/Vulnerability, RAS/RAV = Rule Adherence Strength/Vulnerability, CS/CV=Conduct Strength/Vulnerability, IS/IV=Insight Strength/ Vulnerability, PS/PV = Plan Strength/Vulnerability, COS/COV = Coping Strength/Vulnerability, TS/TV = Treatability Strength/Vulnerability

Specifically, the highest AUCs of the strength subscale of the Thai START was .76 (unauthorized leave) followed by .75 (sexual aggression) and .69 (self-neglect) at time 1 and .99 (substance abuse) followed by .98 (being victimized) and .93 (self-harm) at time 2. Finally, the strongest predictability using the AUCs of the vulnerability subscale of the Thai S(Timmins et al., 2018)TART was .85 (suicide ideation and planning) followed by .82 (verbal aggression) and .77 (unauthorized leave) at time 1 and .99 (being victimized and suicide ideation/planning) followed by .94 (substance abuse) and .92 (aggression against property) at time 2 (Table 2).

Table 2

Frequency of Total Target Behaviors Recorded for the SOS and AUC for the Subscales of the Thai START for All Professions at Time 1 (n = 118) and 2 (n = 93)

| Item | SOS | | AUCs of Thai START | | | |
| | Frequency | | Strength | | Vulnerability | |
	Time 1	Time 2	Time 1	Time 2	Time 1	Time 2
Unauthorized leave	4	3	.76	.70	.77	.76
Self-neglect	302	154	.69	.56	.65	.68
Being victimized	NA	1	NA	.98	NA	.99
Self-harm	3	1	.53	.93	.58	.80
Verbal aggression	42	4	.62	.76	.82	.86
Aggression against property	2	1	.55	.79	.65	.92
Stalking	NA	NA	NA	NA	NA	NA
Physical aggression against others	40	6	.53	.60	.73	.71
Sexual aggression	6	17	.75	.79	.68	.64
Suicide ideation and planning	9	2	.58	.97	.85	.99
Suicide behavior	NA	1	NA	.53	NA	.80
Substance abuse	13	1	.52	.99	.59	.94

Discussion and Conclusion

This study aimed to examine the underlying structures of the translated version of the START within a forensic psychiatric institution in Thailand. The study recruited a sample of 118 forensic psychiatric inpatients. Previous studies provided evidence that violence risk assessment is largely derived from instruments for risk assessment, which were developed in western clinical or criminal justice populations and evaluated in similar settings (O'Shea & Dickens, 2014). This research investigated the generalizability of these advances in the practice of forensic mental health in an Asian country, where cultural, criminal justice, and clinical factors markedly differ from those of the western environment. The study observed good-to-excellent inter-rater reliability for individual START items and acceptable internal consistency of the subscales of the Thai START. These findings support the previous claim of the generalization of the strong reliability of the tool (De Beuf et al., 2021). Furthermore, the general concordance among disciplines or mental health professionals in the current study may stem from the intensive workshop of the Thai START prior to its implementation. The intensive training conducted by experts in the START enabled clinicians to gain direct experiences with inpatients, to discuss the concepts, and to administer the Thai START to several case studies with confidence. Alternatively, clinical experiences in the forensic setting of individual raters may also influence the scoring decision (Timmins et al., 2018). Accordingly, the excellent inter-rater reliability of the Thai START was plausibly due to the comparable training and clinical experiences of the raters.

The current result relatively supports the two-factor model of the START given the strong relationship between the strength and vulnerability factors. The bi-factor construct of the Thai START confirms its multidimensional nature and supports its distinguishable subscales (i.e., strength and vulnerability; Webster et al., 2009). Although the Thai START demonstrated good-to-excellent reliability and acceptable construct validity, the items substance use and external trigger did not correlate with the subscales. This result may imply great clinical relevance for the selection of key items for predicting violent/challenging behaviors and treatment outcomes (Paetsch et al., 2019). Individual items of the START can be selected to predict specific violent behaviors in men or women in other clinical settings; however, substance use and external trigger differ from other START items in predominantly male inpatients (O'Shea & Dickens, 2014; Paetsch et al., 2019). Notably, this finding uniquely reflects the perceptions of the Thai mental health professionals. An important contribution and implication of the study is that implementing the START in forensic setting should consider relevant cultural factors in a clinical environment, where serious violence is rarer compared with those of a previous research for western samples (Zhou et al., 2016).

Furthermore, previous research indicated duplicated and redundant constructs between the strength and vulnerability subscales of the START (Abidin et al., 2013; O'Shea & Dickens, 2014; Wilson et al., 2010). In contrast, this study is in line with several studies that demonstrate that both subscales are distinguishable and should be retained due to their potential to inform therapeutic relationships and risk management plans (De Beuf et al., 2021; de Ruiter & Nicholls, 2011; Nonstad et al., 2010). Thus, this result suggests a unique contribution of each subscale to the Thai START. Another possibility exists that specific variations in each subscale may lead to differential predictive powers over time, which warrants future investigation. Finally, previous studies indicated that the vulnerability scale of the START possesses better predictive power than the strength subscale, but AUC measures of the strength subscale may increase over time in contrast to the findings on vulnerability, which may decrease over time (Quinn et al., 2013). The current study provides evidence that the strength and vulnerability subscales may produce different predictive powers in terms of the type of violent/challenging behaviors at varying follow-up periods (two and six weeks). The strength subscale outperformed the vulnerability subscale in terms of predicting sexual aggression. In addition, notably, that the subscales of the Thai START generally demonstrated a better predictability at six-week follow-up relative to the two-week follow-up and exhibited downward trends for violent/challenging behaviors. Similar results in terms of optimal predictability were found at the six-week follow-up in Braithwaite et al. (2010) and Nicholls et al. (2006). This result is also consistent with that of Gunenc et al. (2018) in that the improvement in violent/challenging behaviors over six weeks may be attributable to the implementation of treatment (e.g., medication or psychotherapy) and partially informed by START assessment and planning. Thus, concluding that the main benefits of the Thai START should be utilized to manage, predict, and prevent violent/challenging behaviors among Thai forensic psychiatric patients is possible. Moreover, the Thai START should be embedding in the national guidelines and work manuals for forensic psychiatric mental health professionals.

Limitations
This study is the first to examine the factorial structure of the START using CFA. Nonetheless, the study presents certain limitations. Although the sample consisted of 118 participants with neurological and psychiatric disorders (e.g., schizophrenia and adjustment disorder), the current results cannot be generalized to other forensic psychiatric samples. Thus, testing the invariance of the scale across gender would be beneficial. Specifically, evidence to support its use among populations of female forensic psychiatric inpatients is limited. The current study did not assess the intellectual functioning of the patients. To address this concern, further studies should investigate patients with intellectual disabilities. This aspect may contribute to a more complete visualization of forensic psychiatric populations. In the present study, the small subsample size did not enable the testing of measurement invariance with sufficient statistical power. Finally, further research with a larger sample is required to improve the generalizability of the START.

Implications for Behavioral Science

Similar to other Asian countries, the forensic mental health system in Thailand remains in its infancy. This system was systematically established in 2002 with the provision of forensic psychiatric services. However, the main obstacles to be addressed in this system were the shortage of mental health workers, limited supplies of medications and tools, a poor understanding of psychiatric disorders among patients and their families, and a lack of confidence in the assessment and management of psychiatric patients among practitioners. This scenario was in line with the demand for a standardized risk assessment tool among forensic mental health professionals in Thailand to ensure the appropriate assessment and management of risk for violent/challenging behaviors posed by offenders with mental illnesses. At present, the Thai START is presumably the only risk assessment tool validated in the Thai forensic setting. From the theoretical point of view, the Thai START offers multidimensional frameworks for assessing forensic psychiatric inpatients in terms of strength and weakness and for encouraging communication among mental health professionals in terms of planning the treatment and monitoring of patients. From the practical point of view, the methodology and main results can be considered as a guideline for validating cross-cultural assessment of violence risk. This validated tool can be adopted in clinical and criminal justice settings. Moreover, the risk of violent/challenging behaviors by forensic patients or prisoners can be identified, and the outcome may be used to determine prediction, management, detention, sentencing, and release.

Conclusion

The present investigation examined the psychometric properties (i.e., reliability and validity) of the Thai START in the forensic mental health setting. In general, the results demonstrated high levels of reliability and acceptable levels of validity. However, no firm conclusions can be drawn about the generalizability of the results to other populations. The current study indicates that within the forensic setting, the Thai START may be a useful adjunct for effective multidisciplinary assessment, treatment, and risk management. Thus, adapting from the full- to the short-form of the Thai START may be necessary to facilitate daily clinical use. Using the translated version of a risk assessment instrument intended for western populations could neglect the role of risk variables specific to or of greater relevance to violent/challenging behaviors of the Thai population. This aspect could prove a fertile area for future research in Thailand and in other Asian clinical settings. The methodology outlined in this study could then be reapplied to evaluate future risk assessment tools that are locally generated for the benefit of patients who receive psychiatric care and of public safety.

References

Abidin, Z., Davoren, M., Naughton, L., Gibbons, O., Nulty, A., & Kennedy, H. G. (2013). Susceptibility (risk and protective) factors for in-patient violence and self-harm: prospective study of structured professional judgement instruments START and SAPROF, DUNDRUM-3 and DUNDRUM-4 in forensic mental health services. *BMC Psychiatry, 13*(1), 197. https://doi.org/10.1186/1471-244X-13-197

Atkinson, T. M., Rosenfeld, B. D., Sit, L., Mendoza, T. R., Fruscione, M., Lavene, D., Shaw, M., Li, Y., Hay, J., Cleeland, C. S., Scher, H. I., Breitbart, W. S., & Basch, E. (2011). Using confirmatory factor analysis to evaluate construct validity of the Brief Pain Inventory (BPI). *Journal of Pain Symptom Management, 41*(3), 558-565. https://doi.org/10.1016/j.jpainsymman.2010.05.008

Bakker, M., Hartgerink, C. H. J., Wicherts, J. M., & van der Maas, H. L. J. (2016). Researchers' intuitions about power in psychological research. *Psychological Science, 27*(8), 1069-1077. https://doi.org/10.1177/0956797616647519

Braithwaite, E., Charette, Y., Crocker, A. G., & Reyes, A. (2010). The predictive validity of clinical ratings of the Short-Term Assessment of Risk and Treatability (START). *International Journal of Forensic Mental Health, 9*(4), 271-281. https://doi.org/10.1080/14999013.2010.534378

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (2nd ed.). Guilford.

Burgess, E. W. (1928). *Factors Determining Success or Failure on parole*. State Board of Parole.

Cartwright, J. K., Desmarais, S. L., Hazel, J., Griffith, T., & Azizian, A. (2018). Predictive validity of HCR-20, START, and static-99R assessments in predicting institutional aggression among sexual offenders. *Law and Human Behavior, 42*(1), 13-25. https://doi.org/10.1037/lhb0000263

Chu, C. M., Thomas, S. D. M., Ogloff, J. R. P., & Daffern, M. (2011). The predictive validity of the Short-Term Assessment of Risk and Treatability (START) in a secure forensic hospital: Risk factors and strengths. *International Journal of Forensic Mental Health, 10*(4), 337-345. https://doi.org/10.1080/14999013.2011.629715

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284-290. https://doi.org/10.1037/1040-3590.6.4.284

De Beuf, T. L. F., de Ruiter, C., Edens, J. F., & de Vogel, V. (2021). Taking "the boss" into the real world: Field interrater reliability of the Short-Term Assessment of Risk and Treatability: Adolescent Version. *Behavioral Sciences & the Law, 39*(1), 123-144. https://doi.org/https://doi.org/10.1002/bsl.2503

de Ruiter, C., & Nicholls, T. L. (2011, 2011/07/01). Protective factors in forensic mental health: A new frontier. *International Journal of Forensic Mental Health, 10*(3), 160-170. https://doi.org/10.1080/14999013.2011.600602

Desmarais, S. L., Nicholls, T. L., Read, J. D., & Brink, J. (2010). Confidence and accuracy in assessments of short-term risks presented by forensic psychiatric patients. *The Journal of Forensic Psychiatry & Psychology, 21*(1), 1-22. https://doi.org/10.1080/14789940903183932

Desmarais, S. L., Nicholls, T. L., Wilson, C. M., & Brink, J. (2012). Using dynamic risk and protective factors to predict inpatient aggression: Reliability and validity of START assessments. *Psychological Assessment, 24*(3), 685-700. https://doi.org/10.1037/a0026668

DiStefano, C., & Morgan, G. B. (2014). A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal, 21*(3), 425-438. https://doi.org/10.1080/10705511.2014.915373

Douglas, T., Pugh, J., Singh, I., Savulescu, J., & Fazel, S. (2017). Risk assessment tools in criminal justice and forensic psychiatry: The need for better data. *European Psychiatry, 42*, 134-137. https://doi.org/10.1016/j.eurpsy.2016.12.009

Fazel, S., Singh, J. P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: Systematic review and meta-analysis. *BMJ: British Medical Journal, 345*, e4692. https://doi.org/10.1136/bmj.e4692

Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement, 33*(3), 613-619. https://doi.org/10.1177/001316447303300309

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychol Methods, 9*(4), 466-491. https://doi.org/10.1037/1082-989x.9.4.466

George, D., & Mallery, P. (2003). SPSS for windows step-by-step: A simple guide and reference, 14.0 update (7th ed.). *http://lst-iiep.iiep-unesco.org/cgi-bin/wwwi32.exe/[in=epidoc1.in]/?t2000=026564/(100)*

Gunenc, C., O'Shea, L. E., & Dickens, G. L. (2018). Structured risk assessment for reduction of multiple risk outcomes in a secure mental health setting: Use of the START. *Crim Behav Ment Health, 28*(1), 61-71. https://doi.org/10.1002/cbm.2036

Haines, A., Brown, A., Javaid, S. F., Khan, F., Noblett, S., Omodunbi, O., Sadiq, K., Zaman, W., & Whittington, R. (2018). Assessing protective factors for violence risk in U.K. general mental

health services using the structured assessment of protective factors. *International Journal of Offender Therapy and Comparative Criminology, 62*(12), 3965-3983. https://doi.org/10.1177/0306624x17749449

Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8*(1), 23-34. https://doi.org/10.20982/tqmp.08.1.p023

Heilbrun, K., Yasuhara, K., & Shah, S. (2010). Violence risk assessment tools: Overview and critical analysis. In *Handbook of violence risk assessment* (pp. 1-17). Routledge/Taylor & Francis Group.

Hogan, N. R., & Olver, M. E. (2018). A prospective examination of the predictive validity of five structured instruments for inpatient violence in a secure forensic hospital. *International Journal of Forensic Mental Health, 17*(2), 122-132. https://doi.org/10.1080/14999013.2018.1431339

Hui, C. H., & Triandis, H. C. (1985). Measurement in cross-cultural psychology: A review and comparison of strategies. *Journal of Cross-Cultural Psychology, 16*(2), 131-152. https://doi.org/10.1177/0022002185016002001

Hurducas, C. C., Singh, J. P., de Ruiter, C., & Petrila, J. (2014). Violence risk assessment tools: A systematic review of surveys. *International Journal of Forensic Mental Health, 13*(3), 181-192. https://doi.org/10.1080/14999013.2014.942923

Hvidhjelm, J., Brandt-Christensen, M., Delcomyn, C., Møllerhøj, J., Siersma, V., & Bak, J. (2022). Effects of Implementing the Short-Term Assessment of Risk and Treatability (START) for Mechanical restraint in a forensic male population: A stepped-wedge, cluster-randomized design. *Frontiers in Psychiatry, 13*, Article 822295. https://doi.org/10.3389/fpsyt.2022.822295

Kikuchi, A., Soshi, T., Kono, T., Koyama, M., & Fujii, C. (2021). Validity of Short-Term Assessment of Risk and Treatability in the Japanese forensic probation service. *Frontiers in Psychiatry, 12*, Article 645927. https://doi.org/10.3389/fpsyt.2021.645927

Kyriazos, T. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology, 09*, 2207-2230. https://doi.org/10.4236/psych.2018.98126

Lowder, E. M., Desmarais, S. L., Rade, C. B., Johnson, K. L., & Van Dorn, R. A. (2019). Reliability and validity of START and LSI-R assessments in mental health jail diversion clients. *Assessment, 26*(7), 1347-1361. https://doi.org/10.1177/1073191117704505

Marasini, D., Quatto, P., & Ripamonti, E. (2016). Assessing the inter-rater agreement for ordinal data through weighted indexes. *Statistical Methods in Medical Research, 25*(6), 2611-2633. https://doi.org/10.1177/0962280214529560

McArdle, J. J. (1996). Current directions in structural factor analysis. *Current Directions in Psychological Science, 5*(1), 11-18. https://doi.org/10.1111/1467-8721.ep10772681

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (pp. 13-104). American Council on Education & Macmillan.

Monahan, J. (2008). Structured risk assessment of violence. In R. I. Simon & K. Tardiff (Eds.), *Textbook of Violence Assessment and Management.* American Psychiatric Publishing.

Mossman, D. (1994). Assessing predictions of violence: Being accurate about accuracy. *Journal of Consulting and Clinical Psychology, 62*(4), 783-792. https://doi.org/10.1037//0022-006x.62.4.783

Nicholls, T., Gagnon, N., Crocker, A., Brink, J., Desmarais, S., & Webster, C. (2007). *START Outcome Scale (SOS).* BC Mental Health & Addiction Services.

Nicholls, T. L., Brink, J., Desmarais, S. L., Webster, C. D., & Martin, M. L. (2006). The short-term assessment of risk and treatability (START): A prospective validation study in a forensic psychiatric sample. *Assessment, 13*(3), 313-327. https://doi.org/10.1177/1073191106290559

Nonstad, K., Nesset, M. B., Kroppan, E., Pedersen, T. W., Nøttestad, J. A., Almvik, R., & Palmstierna, T. (2010). Predictive validity and other psychometric properties of the Short-Term Assessment of Risk and Treatability (START) in a Norwegian high secure hospital. *International Journal of Forensic Mental Health, 9*(4), 294-299. https://doi.org/10.1080/14999013.2010.534958

O'Shea, L. E., & Dickens, G. L. (2014). Short-term assessment of risk and treatability (START): Systematic review and meta-analysis. *Psychological Assessment, 26*(3), 990-1002. https://doi.org/10.1037/a0036794

O'Shea, L. E., Picchioni, M. M., & Dickens, G. L. (2016). The predictive validity of the Short-Term Assessment of Risk and Treatability (START) for multiple adverse outcomes in a secure psychiatric inpatient setting. *Assessment, 23*(2), 150-162. https://doi.org/10.1177/1073191115573301

Paetsch, A., van Os, T. W. D. P., Troquete, N. A. C., & van den Brink, R. H. S. (2019). Single-item predictive validity of the Short-term assessment of risk and treatability (START) for violent behaviour in outpatient forensic psychiatry. *The Journal of Forensic Psychiatry & Psychology, 30*(4), 630-641. https://doi.org/10.1080/14789949.2019.1581831

Quinn, R., Miles, H., & Kinane, C. (2013). The validity of the short-term assessment of risk and treatability (START) in a UK medium secure forensic mental health service. *The International Journal of Forensic Mental Health, 12*, 215-224. https://doi.org/10.1080/14999013.2013.832714

Rogers, R. (2000). The uncritical acceptance of risk assessment in forensic practice. *Law and Human Behavior, 24*(5), 595-605. https://doi.org/10.1023/A:1005575113507

Sanmontree, W. (2022). *The validity and reliability of the START (Short-Term Assessment of Risk and Treatability) in the Thai forensic psychiatric population* [Unpublished doctoral dissertation]. Institute of Psychaitry, Psychology, and Neuroscience, King's College London.

Silva, E. (2020). The HCR-20 and violence risk assessment - will a peak of inflated expectations turn to a trough of disillusionment?. *BJPsych Bulletin, 44*(6), 269-271. https://doi.org/10.1192/bjb.2020.14

Singh, J. P., Desmarais, S. L., Hurducas, C., Arbach-Lucioni, K., Condemarin, C., Dean, K., Doyle, M., Folino, J. O., Godoy-Cervera, V., Grann, M., Ho, R. M. Y., Large, M. M., Nielsen, L. H., Pham, T. H., Rebocho, M. F., Reeves, K. A., Rettenberger, M., de Ruiter, C., Seewald, K., & Otto, R. K. (2014). International perspectives on the practical application of violence risk assessment: A global survey of 44 countries. *International Journal of Forensic Mental Health, 13*(3), 193-206. https://doi.org/10.1080/14999013.2014.922141

Strauss, M. E., & Smith, G. T. (2009). Construct validity: Advances in theory and methodology. *Annual Review of Clinical Psychology, 5*, 1-25. https://doi.org/10.1146/annurev.clinpsy.032408.153639

Swets, J. A. (1986). Indices of discrimination or diagnostic accuracy: Their ROCs and implied models. *Psychological Bulletin, 99*(1), 100-117.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53-55. https://doi.org/10.5116/ijme.4dfb.8dfd

Thamakosit, W. (2007). khwām trong kh̜ǫng bǣp pramœ̄n khwāmsīang tǭ kān kœ̄t phruttikam runrǣng [Validity of the HCR-20 violence risk assessment scheme]. *Journal of Mental Health of Thailand, 15*(2), 101-109. [in Thai]

Timmins, K. L. E., Evans, L., & Tully, R. J. (2018). Inter-rater reliability of the short-term assessment of risk and treatability (START). *The Journal of Forensic Psychiatry & Psychology, 29*(6), 968-988. https://doi.org/10.1080/14789949.2018.1523945

Tully, J. (2017). HCR-20 shows poor field validity in clinical forensic psychiatry settings. *Evidence-Based Mental Health, 20*(3), 95-96. https://doi.org/10.1136/eb-2017-102745

Viljoen, S., Nicholls, T., Greaves, C., de Ruiter, C., & Brink, J. (2011). Resilience and successful community reintegration among female forensic psychiatric patients: A preliminary investigation. *Behavioral Sciences & the Law, 29*(5), 752-770. https://doi.org/https://doi.org/10.1002/bsl.1001

Webster, C., Martin, M., Brink, J., Nicholls, T., & Desmarais, S. (2009). *Short-Term Assessment of Risk and Treatability (START) version 1.1.* BC Mental Health and Addiction Services Coquitlam.

Whittington, R., Bjørngaard, J. H., Brown, A., Nathan, R., Noblett, S., & Quinn, B. (2014). Dynamic relationship between multiple START assessments and violent incidents over time: A prospective cohort study. *BMC Psychiatry, 14*(1), 323. https://doi.org/10.1186/s12888-014-0323-7

Whittington, R., Pollak, C., Keski-Valkama, A., Brown, A., Haines-Delmont, A., Bak, J., Hvidhjelm, J., Almvik, R., & Palmstierna, T. (2022). Unidimensionality of the strengths and vulnerabilities Scales in the short-term assessment of risk and treatability (START). *International Journal of Forensic Mental Health, 21*(2), 175-184. https://doi.org/10.1080/14999013.2021.1953193

Wilson, C. M., Desmarais, S. L., Nicholls, T. L., & Brink, J. (2010). The role of client strengths in assessments of violence risk using the short-term assessment of risk and treatability (START). *International Journal of Forensic Mental Health, 9*(4), 282-293. https://doi.org/10.1080/14999013.2010.534694

Zhong, S., Yu, R., Cornish, R., Wang, X., Fazel, S., & FoVOx Group (2021). Assessment of violence risk in 440 psychiatric patients in China: Examining the feasibility and acceptability of a novel and scalable approach (FoVOx). *BMC Psychiatry, 21*(1), 120. https://doi.org/10.1186/s12888-021-03115-3