An examination of convergent constructs among Level of Service measures and other measures

Sandy Jung, Melissa Daniels, Michael Friesen, Denise Ledi

An Examination of Convergent Constructs Among

Level of Service Measures and Other Measures

Sandy Jung, Grant MacEwan University

Melissa Daniels, University of Alberta

Michael Friesen, Grant MacEwan University

Denise Ledi, Forensic Assessment & Community Services

Correspondence concerning this article should be addressed to Sandy Jung, Department of Psychology, Grant MacEwan University, P.O. Box 1796, Edmonton, Alberta, T5J 2P2, Canada. E-mail address: JungS0@macewan.ca

Abstract

The Level of Service Inventory-Revised and its successor, the Level of Service/Case

Management Inventory are theoretically based risk assessment measures derived from what are

known as the "Central Eight" risk factors. These Level of Service instruments have been

empirically demonstrated to predict recidivism; however, given the importance of using the these

instruments to assess the central eight risk factors, it is also important to ensure that the

measures' subscales are actually assessing the intended constructs. In the present study, files of

219 offenders were coded to investigate the concurrent and discriminant validity by correlating

seven of the Level of Service subscales with construct-relevant scales of the Personality

Assessment Inventory, Cormier-Lang Criminal History Score, Minnesota Multiphasic

Personality Assessment-2, and intelligence measures. The results provide preliminary support for

the construct validity of the domains measured by the Level of Service instruments, with the

exception of the Emotional/Personal subscale.

An Examination of Convergent Constructs Among

Level of Service Measures and Other Measures

Risk assessment has become a routine and essential component in the effective

management and rehabilitation of offenders. Risk assessments not only assist in the judicial

determination of sentences, but are particularly imperative in ensuring that offenders who are

released into the community receive the interventions and supervision that correspond to their

level of risk and need. As a result, a plethora of risk assessment measures are available, many of

which have been empirically demonstrated to predict reoffending behavior to some degree of

accuracy (Andrews, Bonta, & Wormith, 2006; Langton et al., 2007). Concurrent with the use of

validated measures of risk is the need for theoretical models to understand offending behavior.

Widely accepted in guiding offender management and rehabilitation is a theory of criminal

conduct known as the Risk-Need-Responsivity model (Andrews & Bonta, 2010a; 2010b),

comprising three principles; namely, (1) provide intensive services to higher risk offenders, (2)

target criminogenic needs in interventions, and (3) provide the intervention in a way that

corresponds to the offender's abilities. Widespread acceptance of this model has resulted in a

recognized need for a more dynamic assessment tool that includes key domains in risk

assessment. Along these lines, Bonta (2002) prescribed a number of guidelines that assessment

of offenders should follow, including the use of evidence based actuarial measures, the use of

tools to demonstrate predictive validity, addressing criminogenic needs, and basing assessments

on relevant theories of criminal behaviour. The Level of Service Inventory–Revised (LSI-R;

Andrews & Bonta, 1995) and its successor, the Level of Service/Case Management Inventory

(LS/CMI; Andrews, Bonta, & Wormith, 2004) are two measures that have been developed in

following these guidelines. These Level of Service instruments are theoretically based measures

that are derived from the major predictors of criminal behavior known as the "Central Eight" risk

factors, and have been empirically demonstrated to predict recidivism (Andrews & Bonta,

2010a). Although psychometric evaluations of these measures have shown strong empirical

support with respect to reliability, validity evidence is far more limited, as only predictive

validity has been examined. The present study contributes to this gap in the literature by

investigating the concurrent and discriminant validity of the risk domains measured by the Level

of Service instruments.

The original instrument, the LSI-R, assessed 10 domains (i.e., criminal history,

education/employment problems, financial difficulties, family/marital problems,

accommodation, leisure/recreation, antisocial companions, alcohol/drug problems,

emotional/personal factors, attitudes/orientation) (Andrews & Bonta, 1995), but was later revised

and the number of factors reduced to the "central eight" factors. In the revised version of the

LS/CMI, the central eight factors that were included comprised four major risk/need factors (i.e.,

history of antisocial behaviour, antisocial personality pattern, antisocial cognition, antisocial

associates or companions), as well as four risk/need factors that are moderately correlated to

recidivism (i.e., family/marital problems, education/employment problems, problematic or lack

of prosocial leisure/recreation, alcohol/drug problems) (Andrews et al., 2004). Relevant research

showed that some risk factors and their related subscales are not predictive of criminal

behaviour; hence, these items or subscales were removed from the original Level of Service

measure to develop the LS/CMI. Moreover, an antisocial patterns subscale that improves the

measure's ability to predict recidivism and assess offenders' crimogenic needs was added

(Andrews et al., 2006; Bonta, Law, & Hanson, 1998; Vose, 2009).

The predictive validity of the LSI-R and LS/CMI instruments have been well-established in the literature (Andrews, Bonta, & Wormith, 2006), revealing that they have, at minimum, moderate predictive ability to assess risk of general and violent recidivism, and in some instances, they have been shown to be superior to other risk measures (e.g., compared to PCL-R; Campbell, French, & Gendreau, 2009). Accordingly, the predictive validity of the Level of Service measures to assess the risk of both violent and non-violent recidivism has been fairly promising across a variety of studies, across gender, and with varying offender samples (Campbell et al., 2009; Kelly & Welsh, 2008; Manchak, Skeem, & Douglas, 2008; Vose, Lowenkamp, Smith, & Cullen, 2009). In addition to the predictive value, research has recently shown that changes in offenders' LSI-R scores over time, reflecting changes in their scores on the measure's dynamic items, are predictive of their recidivism as well (Vose, 2009). This suggests that, as the criminogenic needs of offenders are addressed, their likelihood of continuing to engage in criminal behavior can be influenced.

In light of the importance of using the Level of Service measures to assess the central eight domains, it is crucial to ensure that the measures' subscales are actually assessing the intended constructs. There is a paucity of empirical literature that has examined the construct validity of risk assessment measures. Examining the extent to which measures capture the constructs of interest is important and necessary, as it should not be taken for granted that any given measure, as designed or as applied by clinicians, adequately measures the intended constructs. In the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999), it is noted that the "analyses of the relationship of test scores to variables external to the test provide another important source of validity evidence" (p. 13) and "when a test provides more than one score, the distinctiveness of the separate scores should be demonstrated"

(p. 20). There are only two published studies to date that have examined the construct validity of risk assessment measures. They include recent publications from Nunes and Babchishin (2011) who examined the construct validity of the Stable-2000 and Stable-2007, which are dynamic risk assessment instruments for sexual offenders, and Jung, Ledi, and Daniels (in press) who examined the construct validity of the scales and items of the HCR-20, a structured professional judgment measure of risk for violent reoffending. There is currently no published research that examines the construct validity of the domains measured by the LSI-R or the LS/CMI.

The current study examines the concurrent validity of the constructs measured in the LSI-R and LS/CMI instruments by statistically assessing their convergence with construct-relevant scales from the Personality Assessment Inventory (PAI), the Cormier-Lang Criminal History Score, the Minnesota Multiphasic Personality Inventory-2 (MMPI-2), and intelligence measures. Specifically, the convergent validity of the Criminal History, Family and Marital, Companions, Alcohol and Drug, Emotional and Personal, Attitude and Orientation, and Antisocial Pattern subscales of the Level of Service instruments was examined. It was expected that these seven risk factors would be correlated to constructs deemed similar or convergent to other related constructs.

Methods

Participants

Clinical files were reviewed for offenders who were referred to a forensic psychiatric outpatient clinic in a medium-sized Canadian city and who were evaluated at the pre-sentence stage following the receipt of a criminal conviction. The clinic offers court-ordered assessment and treatment services that address offending behaviours and mental health and provides correctional services to probation. Files containing a completed Level of Service measure (i.e.,

LSI-R or LS/CMI) were included in the sample. Two hundred and nineteen files were identified

and used in this study. The LSI-R or the LS/CMI in the offenders' files were administered as

part of a court-ordered evaluation. Eighty-one (37%) files contained a completed LSI-R

protocol, and 138 (63%) files contained a completed LS-CMI protocol. The average age of the

offenders at the time of their offence was 29.17 years ($SD = 9.49$, ranging from 17 to 55). Most

of the offenders were male ($n = 177$, 80.8%), while a fifth of the sample were female ($n = 42$,

19.2%).

Of the 99 files that indicated the ethnicity of the offender, almost half were Caucasian

(47.5%), a third were Aboriginal or Métis (32.3%), and the remaining were identified as Asian,

Black, or other (e.g., Middle Eastern). Employment at the time of the index offence was

categorized for a majority of the sample ($n = 214$), revealing that almost half (46.7%) were

unemployed or enrolled as a student, while 32.7% were employed as unskilled labourers. The

average number of years of education was 10.4 years ($SD = 2.20$). Most of the offenders were

single (60.6%), a quarter were married (24.4%), and less than a fifth were separated or divorced

(15.0%); 62% had at least one biological child. With regards to the types of index offences that

the court-ordered assessment was addressing, there was a great deal of variance, with 21%

committing assaults, 14.6% committing robbery, 14.2% committing property offences, and less

than 10% committing possession of a weapon, uttering threats, sexual offences, break and enter,

driving offences, substance-related offences, fraud, and arson.

Measures

Level of Service Inventory-Revised (LSI-R). The LSI-R is a 54 item theoretically based

risk assessment measure that consists of static and dynamic risk factors, and is designed to

categorize offenders according to risk as well as identify potential treatment needs (Andrews &

Bonta, 1995). Items are sorted into 10 subscales: Criminal History, Education/Employment, Financial, Family/Marital, Accommodation, Leisure/Recreation, Companions, Alcohol/Drug problem, Emotional/Personal, and Attitudes/Orientation. Each item is scored as either absent or present with total scores ranging from 0 to 54. A higher score on the LSI-R denotes a higher risk to recidivate. The LSI-R has strong overall interrater reliability (Interclass Correlation (ICC), rs = .80 to .94) as well as good internal consistency (Cronbach's α =.70). The LSI-R's predictive validity for general recidivism has a mean area under the curve (AUC) of .71 (r = .36) and for violent recidivism it has a mean AUC of .64 (r = .25) (Andrews et al., 2006; AUCs calculated using the conversion table of Rice and Harris, 2005). Similar findings demonstrating good psychometric properties of the LSI-R have been reported in independent reviews of the LSI-R (e.g., ICC = .94; $AUC_{General\ Recidivism}$ = .693; $AUC_{Violent\ recidivism}$ = .667; Kroner & Mills, 2001).

In this study, we specifically examine the convergent validity of the constructs measured by the Criminal History, Family/Marital, Companions, Alcohol/Drug Problems, Emotional/Personal, and Attitudes/Orientation subscales. Two subscales were modified to exclude items in order to correspond with the subscales of the LS/CMI. Of the LSI-R, two items were removed from the Criminal History subscale (i.e., escape history, official record of assault), one item removed from the Alcohol/Drug Problem subscale (i.e., other indicators), and one item removed from the Companions subscale (i.e., a social isolate). Moreover, only items from the LSI-R formed the Emotional/Personal subscale examined as it is not included in the LS-CMI.

Level of Service/Case Management Inventory (LS/CMI). The LS/CMI (Andrews et al., 2004) is a 43-item revised version of the LSI-R. Items and subscales that were not found to strongly relate to recidivism were omitted from this version of the LSI (Andrews & Bonta, 2010b). The Accommodation, Financial, and the Emotional/Personal subscales as well as a

number of items from the Criminal History, Education/Employment, and Alcohol/Drug

Problems subscales were omitted. An Antisocial Patterns subscale was also added in this version

of the LSI. The LS/CMI items are mostly identical to the LSI-R items, but given the reduction of

both the items (i.e., from 54 to 43) and subscales (i.e., Accommodations, Financial, and

Emotional/Personal subscales were omitted, while Antisocial Pattern was added, hence reducing

the subscales from 10 to 8), the total score of the LS/CMI ranges from 0 to 43. Currently, there

are no published studies that have provided internal consistency or interrater reliability for this

measure. Andrews et al. (2006) reported the predictive validity for general (AUC = .739) and

violent recidivism (AUC = .666) as fairly strong.

In this study, no individual items were omitted from this measure. As mentioned in the

previous section, the Emotional/Personal subscale is not present in the LS/CMI measure and only

measured using the LSI-R. The items in the Antisocial Pattern subscale were only available in

the LS/CMI, and therefore the subscale was only taken from this measure. Hence, there were six

subscales included in this examination of concurrent validity (Criminal History, Family/Marital,

Companions, Alcohol/Drug Problems, Emotional/Personal, Attitudes/Orientation, and Antisocial

Pattern subscales).

Personality Assessment Inventory (PAI). The PAI (Morey, 2007) is a self-report

personality measure consisting of 344 items organized into 22 non-overlapping categories. It

includes four validity scales, eleven clinical scales, five treatment attitude scales, and two

interpersonal scales. Each PAI item is rated on a 4-point scale (very true, mainly true, slightly

true, and false). Individuals' responses on the items in each category are transformed into t

scores to allow for interpretation relative to a normative sample including both clinical and

nonclinical participants. All scales and subscales have mean t scores of 50 with standard

deviations of 10. An individual's *t* score that is greater than 50 would indicate a higher level of that particular construct (e.g., Anxiety, ANX) than is found in the normative sample. Ninety-eight percent of nonclinical patients will score less than two standard deviations above the mean; thus, scores two or more standard deviations above the mean (70 or greater) are considered areas of clinical significance. The literature has reported that the PAI has strong internal consistency, reporting Cronbach alphas that range from .81 to .86 across normative, clinical, and college samples. The validity of the measure and its scales has been extensively reviewed in the second edition of the manual (Morey, 2007).

Construct-relevant scales of the PAI were included in this study: Antisocial features (ANT), Antisocial behaviour subscale (ANT-A), Egocentricity (ANT-E), Stimulus seeking (ANT-S), Alcohol problem (ALC), Drug problem (DRG), Nonsupport (NON), Anxiety (ANX), Depression (DEP), Mania (MAN), Paranoia (PAR), and Schizophrenia (SCZ).

Cormier-Lang Criminal History Score (CLS). The CLS is used to quantify an offender's history of criminal offences, a current or index offence, or a particular subgroup of offences and can be used when official police information is available (Quinsey, Harris, Rice, & Cormier, 2006). Different points are assigned to offence convictions (e.g., robbery = 7 points) and the score is tallied based on the offender's prior convictions. The CLS is included as part of two risk assessment instruments, the Violence Risk Appraisal Guide (VRAG) and the Sex Offender Risk Appraisal Guide (SORAG). Two types of CLS were calculated for the present study; these were a non-violent CLS scale, which is used in both the SORAG and VRAG, and a violent CLS scale, used only in the SORAG.

Quinsey et al. (2006) reported that the interrater reliability of the 12 items within the VRAG is strong ($r \geq .80$), and this includes the item for the non-violent CLS scale. They also

reported that the predictive validity of the non-violent CLS scale is within an acceptable range

(AUCs = .70 to .74). A multisite examination of the validity of the CLS scores indicated strong

interrater reliability for both the non-violent (r = .98) and violent (r = 1.00) scales (Harris, Rice,

Quinsey, Lalumiere, & Boer, 2003).

  Minnesota Multiphasic Personality Inventory-2 (MMPI-2). The MMPI-2 (Butcher,

Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) is a measure used to determine various

clinical and personality traits. It is one of the most commonly used measures for personality

assessment due to its extensive empirical validation (Wygant, Sellbom, Gervais, Ben-Porath, &

Stafford, 2010). The MMPI-2 is a 567-item self-report questionnaire that contains 38 scales

covering a wide breadth of personality structures. For the purposes of our study, five clinical

basic scales (i.e., Depression, Paranoia, Schizophrenia, Hypomania, and Social Introversion) and

the five subscales of Psychopathic Deviate (i.e., Familial Discord, Authority Problems, Social

Imperturbability, Social Alienation, Self-Alienation) were used.

  Wechsler Adult Intelligence Scale (WAIS-III) or Wechsler Abbreviated Scale of

Intelligence (WASI). Wechsler measures of intelligence are available in two forms for adult

populations. For this study, if a measure of intelligence was available in the forensic files, it

would most likely be one of these two measures. The WAIS-III (Wechsler, 1997a) is a battery of

measures designed to assess the mental abilities of adults. The WASI (Wechsler, 1999) is a

reliable, brief measure of intellectual ability that has been demonstrated to correlate with scores

on the WAIS-III. Both measures have demonstrated reliability and validity (for a comprehensive

discussion of the psychometric properties of the measures, see Wechsler, 1997b). For the

present study, the full scale intelligence quotient (FSIQ), verbal intelligence quotient (VIQ), and

performance intelligence quotient (PIQ) were included in the analyses to examine the

discriminant validity of the Level of Service subscales.

Procedure

Data was coded retrospectively from clinical files of offenders who were referred for a

court-ordered assessment between 2001 and 2009. Risk measures (LSI-R or LS/CMI) were

completed by the clinician (i.e., doctoral psychologist or psychiatrist) at the time of assessment.

Psychometric measures (PAI, MMPI-2, WAIS-III, WASI) were administered as a questionnaire

to the offender and scored by a psychometrist. A search of files that contained a Level of

Service measure (i.e., LSI-R or LS/CMI) was conducted. Offender files contained a court-

ordered assessment report, criminal records, case notes, offence description reports, demographic

information, and some description of victim information. The items from the Level of Service

measures were coded directly from the measure's protocol. Also, the *t* scores on the

psychometric measures, PAI and MMPI-2, were directly coded from the computerized profile,

the IQ scores (FSIQ, PIQ, VIQ) were recorded from the scored booklets, and the CLS scores

were calculated for each offender from information available in offenders' files. Not all

variables were available in the files for all offenders; therefore, sample sizes varied depending on

the analysis.

Results

This study's primary objective was to examine the concurrent validity of seven subscales

measured by the Level of Service inventories—six of which are considered to be part of the

Central Eight risk factors predictive of future offending (Andrew & Bonta, 2010a). First, the

internal consistency and the intercorrelations among the subscales were examined. Second,

using other measures of constructs that are expected to be related to the subscales, we

investigated whether the five risk factors assessed by the Level of Service measures had

concurrent validity. Third, we examined whether the five risk factors demonstrated discriminant

validity with constructs that should be unrelated. For the latter two objectives, Pearson product-

moment correlation coefficients were calculated, given the continuous nature of the variables

included in the analyses. To correct for multiple correlational analyses, a conservative alpha of

.01 was used. Given that the offender files did not consistently include the same information and

measures, there are discrepancies in sample sizes for some analyses from the total sample of 219

offenders. Descriptive statistics (means, standard deviations, range of scores) are listed in Table

1 for the Level of Service subscales (Criminal History, Family/Marital Problems, Alcohol/Drug

Problems, Emotional/Personal, Attitudes/Orientation, and Antisocial Patterns).

Internal Validity of Each Level of Service Subscale

The intercorrelations between and internal consistencies of the Level of Service subscales

are presented in Table 1. We found significant positive correlations between all of the measures

with a few notable exceptions. First, the LSI-R's Emotional/Personal subscale did not

significantly correlate with any other LSI-R subscales included in this study. It can be noted that

the Emotional/Personal subscale was dropped in the revised LS/CMI version; hence, no

correlations could be calculated between the Emotional/Personal of the LSI-R and the Antisocial

Pattern subscales of the LS/CMI. Second, the Family/Marital subscale was not correlated to the

Alcohol/Drug Problems subscale or the Antisocial Patterns subscale. The internal consistencies

of each subscale ranged from poor (e.g., Attitudes/Orientation, Cronbach's α = .07) to strong

(e.g., Criminal History, Cronbach's α = .80)

Concurrent Validity of the Level of Service Subscales

In the following sections, each Level of Service subscale was evaluated in terms of concurrent validity and discriminant validity by examining its association with other related and unrelated measures. Pearson correlation coefficients are reported in Tables 2 and 3.

Criminal History subscale. The relationship between the Criminal History (CH) subscale and the construct-relevant scales from the PAI (namely, the Antisocial Features scale and its three subscale derivatives), the CLS scores, and the Authority Problems (Pd2) scale of the MMPI-2 was examined. As seen in Table 3, the CH subscale was positively correlated with the Antisocial Features (PAI-ANT) scale and the Antisocial Behaviours subscale (PAI-ANT-A) as expected, but not with the other two subscales, Egocentricity (PAI-ANT-E) and Stimulus Seeking (PAI-ANT-S). The CH subscale was also positively correlated with the violent and the non-violent CLS scores. However, no significant relationship emerged between the CH subscale and the Authority Problems subscale (MMPI2-Pd2). Of the remaining scales and measures, the CH subscale demonstrated discrimination from other unrelated constructs; hence, it was not significantly related to most of the PAI and MMPI-2 scales, or the full scale IQ. However, a few significant positive correlations emerged; the CH subscale was positively correlated with the Alcohol and Drug scales of the PAI and with the Verbal and Performance IQ scores on the intelligence measures.

Family/Marital subscale. We predicted that the Family/Marital (FM) subscale of the Level of Service measures would be significantly correlated with select PAI scales (i.e., Nonsupport, PAI-NON; Warmth, PAI-WRM) and MMPI-2 scales (i.e., Familial Discord, MMPI2-Pd1; Social Alienation, MMPI2-Pd4). Significant correlations in the expected direction emerged with both the PAI and MMPI-2 scales. The FM subscale was positively correlated with perceived lack of social support, as measured by the PAI-NON scale, and negatively correlated

with the extent to which a person is empathic and engaging in relationships, as measured by the

PAI-WRM. Also, with a smaller sample size, the FM subscale was positively correlated with

both the MMPI2-Pd1 and MMPI-Pd4, suggesting that more family and marital problems were

associated with more self-reported familial discord and social alienation. In addition to being

associated with related constructs, the FM subscale was also correlated with unrelated constructs.

Significant positive correlations emerged with the Anxiety, Depression, Schizophrenia, and

Aggression scales of the PAI, both the violent and non-violent criminal history scores of the

CLS, and the Depression scale of the MMPI-2.

     Companions subscale. It was predicted that scores on the Companions (CO) scale of the

Level of Service measures would correlate with scores on the PAI-ANT, along with the three

PAI Antisocial subscales, the PAI-NON, the PAI-WRM (negatively), and two of the subscales

on the MMPI-2, including Social Imperturbability (MMPI2-Pd3) and Social Alienation

(MMPI2-Pd4).

     We found significant positive correlations between the CO scale and the PAI Antisocial

scale and subscales (i.e., PAI-ANT, PAI-ANT-A, PAI-ANT-E, PAI-ANT-S) and the PAI-NON

scale. Moreover, we found a negative correlation between the CO scale and the PAI-WRM as

expected. A significant correlation did not emerge between the CO scale and the two MMPI-2

subscales of the psychopathic deviate construct. We also found relationships between the CO

scale and other constructs that are deemed unrelated, and these include the substance abuse

scales of the PAI (PAI-ALC, PAI-DRG), two clinical psychopathology scales of the PAI (PAI-

DEP, PAI-SCZ), the Aggression scale of the PAI (PAI-AGG), the criminal history scores of the

CLS (both violent and non-violent), and the verbal intelligence scale.

Alcohol/Drug subscale. A relationship was expected between the Alcohol/Drug (AD) subscale of the Level of Service measures and the Alcohol and Drug scales of the PAI (PAI-ALC, PAI-DRG). Although there was a large positive correlation between the AD subscale and the two PAI scales (accounting for variance in the range of 39 to 43%), the subscale was also positively associated with the PAI Antisocial scale and subscales (PAI-ANT, PAI-ANT-A, PAI-ANT-E, PAI-ANT-S), resistance to change as measured by the Treatment Rejection scale (PAI-RXR), the Schizophrenia and Aggression scales (PAI-SCZ, PAI-AGG), and the violent criminal history score of the CLS.

Emotional and Personal subscale. Only available in the LSI-R, the Emotional and Personal (EP) subscale was predicted to correlate with the four psychopathology scales of the PAI (Anxiety, PAI-ANX; Depression, PAI-DEP; Mania, PAI-MAN; Schizophrenia, PAI-SCZ) and the four clinical scales of the MMPI-2 (MMPI2-Dp; Paranoia, MMPI2-Pa; Schizophrenia, MMPI2-Sz; Hypomania, MMPI2-Ma). The EP subscale was not significantly correlated with any of the predicted clinical scales from the self-report measures. To the subscale's credit, it was not related to any of the unrelated constructs used to examine its discriminant ability, except for a negative correlation with performance IQ.

Attitude and Orientation subscale. It was predicted that scores on the Attitudes/Orientation (AO) subscale would positively correlate with other procriminal constructs including the PAI-ANT and its three subscales, the PAI-AGG (i.e., assesses attitudinal and behavioral features relevant to aggression, anger, and hostility), and the authority problems scale of the MMPI-2 (MMPI2-Pd2). We found significant positive correlations between the AO subscale and the PAI-ANT, PAI-ANT-S, and PAI-AGG scales. No significant correlations materialized between the AO subscale and the two other PAI subscales or any of the

Psychopathic Deviate subscales of the MMPI-2. The AO subscale did demonstrate discriminant validity in that it was not associated with any of the other scales of the PAI or MMPI-2, or the criminal history or intelligence measures.

  Antisocial Pattern subscale. The relationship between the Antisocial Pattern (AP) subscale of the Level of Service measures and the Antisocial scale and subscales of the PAI, PAI-AGG, the criminal history scores of the CLS, and the Psychopathic Deviate subscales of the MMPI-2 were examined. As expected, the AP subscale was positively correlated with the PAI-ANT and PAI-AGG scales, the PAI-ANT-A, PAI-ANT-E, and PAI-ANT-S subscales, and the violent criminal history score of the CLS (but not the non-violent CLS score). However, with the small sample size available, no significant correlations emerged between the AP subscale and the MMPI-2 subscales. For the most part, the AP subscale was not correlated with other unrelated constructs, except for a negative association with the verbal intelligence quotient (indicating that the more antisocial behavioural pattern exhibited, the lower the verbal intelligence).

Discussion

  Examining the extent to which a measure and its scales capture intended constructs is of the utmost importance, as it should not be taken for granted that measures adequately assess the constructs of interest (AERA et al., 1999). Hence, the current study is an important and necessary step toward establishing the concurrent and discriminant validity of the constructs measured by the Level of Service measures. The subscales of the LSI-R and LS/CMI were compared to constructs measured in the PAI, CLS, MMPI-2, and intelligence measures. Overall, moderate to strong support was found for the concurrent validity of six of the seven subscales

examined. Notably, these six subscales are part of the latest version of the Level of Service

measures (i.e., LS/CMI; Andrews et al., 2004).

      The Criminal History subscale was moderately related to the other subscales on the Level

of Service measures (with the exception of the Emotional/Personal subscale, discussed later), had

excellent internal consistency, and was associated with several related constructs, demonstrating

good concurrent validity. Of particular note, the subscale was not distinct from self-identified

alcohol or drug problems or intelligence measures, although it was associated with many of the

other constructs, such as psychopathology (e.g., depression) and personality traits (e.g., warmth).

      In examining the Family/Marital subscale, support for its construct validity was quite

strong as all related constructs were associated with this subscale. However, discriminant

validity was not well-supported, as the subscale showed no contrast from intelligence scores,

treatment rejection, criminal history, and psychopathology scales. Interestingly, there was an

association between the Family/Marital subscale and all of the other subscales on the Level of

Service measures, except for the Emotional/Personal and Alcohol/Drug subscales. The lack of

association with the Alcohol/Drug subscale was surprising considering the association between

substance abuse and maladaptive family/marital relationships often seen in the literature

(Hutchison, 1999; Kantor & Strause, 1989). Also, in the present study, the subscale was

associated with the Alcohol and Drug scales on the PAI.

      The association between the Family/Marital subscale and the selected convergent

measures of perceived social support (i.e., nonsupport of the PAI, social alienation of the MMPI-

2) suggests that the subscale is an appropriate conceptualization of this risk factor and

encourages the use of this subscale to identify a relevant criminogenic need. Preliminary

research suggests that interventions aimed at increasing social support to post-release offenders

have been effective (Wilson, Cortoni, & McWhinnie, 2009), while the absence of positive

familial and marital relationships represents a significant risk factor for potential recidivism

(Johnson, Smailes, Cohen, Kasen, & Brook, 2004).

   The association between the Companions subscale and the related constructs of

antisociality, nonsupport, and interpersonal comfort provides support for its concurrent validity.

However, its discriminant validity was more questionable given its additional association with

other unrelated constructs, such as alcohol and drug problems, psycholopathology, criminal

history, and verbal intelligence. Moreover, it had poor internal consistency in this study as

compared to other studies (e.g., Loza & Simourd, 1994). In its favour, the subscale was

correlated with the other subscales (except the Emotional/Personal subscale). Similar to the

Criminal History subscale, the Alcohol/Drug subscale was well-supported with respect to both

concurrent validity and discriminant validity. However, an unusual association between this

subscale and the Schizophrenia scale of the PAI was found. Further, the subscale had a

relatively weak internal consistency, especially when compared to other empirical studies (Kelly

& Welsh, 2008; Loza & Simourd, 1994).

   An examination of the Emotional/Personal subscale of the LSI-R provided support for the

removal of this subscale from the revised version of the measure. Contrary to expectations, the

subscale was not associated with any of the psychopathological scales of the PAI and MMPI-2,

despite having good internal consistency and good discrimination from other construct-irrelevant

variables. Also, the subscale was not associated with any of the other subscales on the Level of

Service measures. These findings support that the Emotional/Personal subscale is not effective

for measuring an offender's emotional and mental well-being. The authors of the Level of

Service measures noted that the Emotional/Personal subscale was omitted from the revised

LS/CMI version of the LSI-R because it was not found to be predictive of recidivism (Bonta et

al., 1998; Vose, 2009). This, in turn, could suggest that the subscale's lack of predictive power

may be a result of it being a poor measure of an individual's mental health. However, it is more

likely that our findings reflect the differences between the PAI clinical scales and the

Emotional/Personal subscale. In particular, the Emotional/ Personal subscale is a measure of an

interviewer's perception of an offender's general mental well-being; in contrast, the PAI scales

are domain specific self-reports. Accordingly, the PAI may be too specific to allow for

comparisons with a general measure of perceived mental health. Although this may account for

our findings, previous research still suggests that mental illness is not strongly related to an

offender's likelihood to reoffend; as such, its exclusion from the revised version, the LS/CMI, is

more likely due to this reason (Bonta et al., 1998).

Despite poor internal consistency among the items, the Attitude/Orientation subscale of

the Level of Service measures demonstrated good concurrent validity and very good discriminant

validity. It also consistently showed associations with the other subscales. Similar findings

emerged with the Antisocial Pattern subscale, which demonstrated very good concurrent and

discriminant validities, although an unexpected association arose between this subscale and

verbal intelligence. Moreover, the Antisocial Pattern subscale was associated with the 5 other

subscales examined from the LS/CMI and showed adequate internal consistency.

Analyses conducted in this study provide overall support for the assumption that the

subscales in the most recent Level of Service measure, the LS/CMI, are associated with the

intended constructs denoted by the "Central Eight" risk factors (Andrews & Bonta, 2010a).

Therefore, for the most part, the Level of Service measures have good construct validity, with the

exception of the Emotional/Personal subscale. These findings add to the evidence supporting the

Level of Service measures and accordingly bolster the confidence that can be placed in this risk assessment tool.

From a clinical perspective, this research has important implications with respect to treatment decisions for offenders. Previous research demonstrating the predictive validity of the LSI-R and the LS/CMI has only addressed the assessment of an offender's risk to recidivate. The inclusion of dynamic factors in the Level of Service measures was intended to allow for the identification of an offender's criminogenic needs and to accordingly highlight areas to target in interventions. This study provides important preliminary support by suggesting that the subscales are actually measuring the static and criminogenic factors they were designed to measure. This can increase the confidence of clinicians that the measures reflect risk factors that are foundationally important to treatment change and for predicting recidivism—in essence, ensuring that the measure is assessing the criminogenic needs that the measure purports to assess.   These findings not only add to the growing body of research on the Level of Service measures in the forensic field, but also highlight areas that future development of the measure could focus upon. For example, by inquiring as to whether the subscales assess only the intended constructs or assess a broader area of risk, enhancement and revision of the subscale items and operational definitions provided in the manual may serve to improve the measure. Of note, the Family/Marital and Companions subscales appear to have strong concurrent validity but are weaker in discriminant validity, and therefore, may be assessing a broader understanding of the risk factor intended. Revision of the existing items to increase specificity and sensitivity to the domain may increase its discriminant validity whilst maintaining its concurrent validity.

There are a number of limitations that should be considered in interpreting the findings of the current study. Because some of the comparisons made in this study used relatively low

sample sizes (e.g., Level of Service subscales and the MMPI-2 clinical scales, ns = 24, 35, 36),

some comparisons should be interpreted cautiously. Low sample sizes may have contributed to

nonsignificant results for analyses involving the MMPI-2 scales. Another methodological

limitation is associated with the archival nature of the study. In particular, not all of the offenders

in the sample were administered all of the scales and subscales examined in this study, and it was

impossible to stratify the data based on the nature of the offence (violent or non-violent

offences). While assaults accounted for the largest single category of index offences (21%), non-

violent crimes accounted for the majority of the index offences (72%). For this reason, the results

of this study should be interpreted cautiously when generalizing beyond general criminality.

Although these considerations are quite common among studies on risk assessment measures,

future research should continue to improve upon and address these limitations.

In conclusion, the present study provides preliminary support for the construct validity of

the LSI-R and the LS/CMI subscales. It is often tempting to take an overly pragmatic stance by

evaluating forensic risk assessment measures solely based on their predictive validity. However,

an examination of these measures must also encompass construct validity, as predictive validity

alone is insufficient. Indeed, that a tool has been shown to be effective does not equate to that

tool being based solely on prediction. As Bonta (2002) has asserted, it is important to select risk

assessment measures the demonstrate predictive validity, but also measures should address and

assess the criminogenic constructs that they purport to measure.

References

American Educational Research Association (AERA), American Psychological Association

(APA), & National Council on Measurement in Education (NCME). (1999). *Standards

for educational and psychological testing.* Washington, DC: American Educational

Research Association.

Andrews, D. A. & Bonta, J. (1995). *The Level of Service Inventory–Revised*. Toronto, Canada:

Multi-Health Systems.

Andrews, D.A. & Bonta, J. (2010a). Rehabilitating criminal justice policy and practice.

Psychology, Public Policy, and Law, 16, 39-55.

Andrews, D.A. & Bonta, J. (2010b). *The psychology of criminal conduct (5th ed.)*. Cincinnati,

OH: Anderson.

Andrews, D.A., Bonta, J., & Wormith, J.S. (2004). *The Level of Service/Case Management

Inventory (LS/CMI): User's Manual*. Toronto, Canada: Multi-health Systems.

Andrews, D.A., Bonta, J., & Wormith, J.S. (2006). The recent past and near future of risk and/or

need assessment. *Crime & Delinquency, 52,* 7-27.

Bonta, J. (2002). Offender risk assessment guidelines for selection and use. *Criminal Justice and

Behavior, 29,* 355-379.

Bonta, J., Law, M., & Hanson, K. (1998). The prediction of criminal and violent recidivism

among mentally disordered offenders: A meta-analysis. *Psychological Bulletin, 123,* 123-

142.

Butcher, J.N., Dahlstrom, W.G., Graham, J.R., Tellegen, A., & Kaemmer, B. (1989). *Minnesota

Multiphasic Personality Inventory (MMPI-2) manual for administration and scoring*.

Minneapolis, MN: University of Minnesota Press.

Campbell, M.A., French, S., & Gendreau, P. (2009). The prediction of violence in adult

offenders: a meta-analytic comparison of instruments and methods of assessment.

Criminal Justice and Behavior, 36, 567-590.

Harris, G.T., Rice, M.E., Quinsey, V., Lalumiere, M., & Boer, D. (2003). A multisite comparison

of actuarial risk instruments for sex offenders. *Psychological Assessment, 15,* 413-425.

Hutchison, I. (1999). Alcohol, fear, and woman abuse. *Sex Roles, 40,* 893-920.

Johnson, J.G., Smailes, E., Cohen, P., Kasen, S., & Brook, J.S. (2004). Antisocial parental

behviour, problematic parenting, and aggressive offspring behaviour during adulthood.

British Journal of Criminology, 44, 915-930.

Jung, S., Ledi, D., & Daniels, M. (in press). Evaluating the construct validity of the HCR-20.

Journal of Risk Research.

Kantor, G. K., & Straus, M.A. (1989). Substance abuse as a precipitant of wife abuse

victimizations. *American Journal of Drug and Alcohol Abuse, 15*, 173-189.

Kelly, C. E., & Welsh, W. N. (2008). The predictive validity of the Level of Service Inventory-

Revised for drug-involved offenders. *Criminal Justice and Behavior, 35,* 819-831.

Kroner, D.G., & Mills, J.F. (2001). The accuracy of five risk appraisal instruments in predicting

institutional misconduct and new convictions. *Criminal Justice and Behavior, 28,* 427-

489.

Langton, C. M., Barbaree, H. E., Seto, M. C., Peacock, E. J., Harkins, L., & Hansen, K. T.

(2008). Actuarial assessment of risk for reoffense among adult sex offenders: Evaluating

the predictive accuracy of the Static-2002 and five other instruments. *Criminal Justice

and Behavior, 34,* 37-59.

Loza, W., & Simourd, D. J. (1994). Psychometric evaluation of the level of supervision

inventory (LSI) among male Canadian federal offenders. *Criminal Justice and Behavior,*

21, 468-480.

Manchak, S.M., Skeem, J.L., & Douglas, K.S. (2008). Utility of the Revised Level of Service

Inventory (LSI-R) in predicting recidivism after long-term incarceration. *Law & Human*

Behavior, 32, 477-488.

Morey, L. C. (2007). *Personality Assessment Inventory: Professional manual. (2nd ed.).* Tampa,

FL: Psychological Assessment Resources.

Nunes, K.L. & Babchishin, K.M. (2011). Construct validity of Stable-2000 and Stable-2007.

Sexual Abuse: A Journal of Research and Treatment, 24, 29-45.

Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders:*

Appraising and managing risk. Washington, DC: American Psychological Association

Rice, M.E., & Harris, G.T. (2005). Comparing effect sizes in follow-up studies ROC area,

Cohen's d and r. *Law and Human Behavior, 29,* 615-620.

Vose, B. (2009). *Assessing the predictive validity of the Level Of Service Inventory-Revised:*

recidivism among iowa parolees and probationers. (Unpublished Doctoral Dissertation).

University of Cincinnati, Ohio.

Vose, B., Cullen, F., & Smith, P. (2008). The empirical status of the Level of Service Inventory.

Federal Probation, 72, 22-29.

Vose, B., Lowenkamp, C. T., Smith, P., & Cullen, F. T. (2009). Gender and the predictive

validity of the LSI-R: A study of parolees and probationers. *Journal of Contemporary*

Criminal Justice, 25, 459-471.

Wechsler, D. (1997a). *WAIS-III administration and scoring manual*. San Antonio, TX:

Psychological Corporation.

Wechsler, D. (1997b). *WAIS-III/WMS-III technical manual*. San Antonio, TX: Psychological

Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The

Psychological Corp.

Wilson, R.J. Cortoni, F., & McWhinnie, A.J. (2009). Circles of Support and Accountability: a

Canadian national replication of outcome findings. *Sexual Abuse: A Journal of Research

and Treatment, 21,* 412-430.

Wygant, D., Sellbom, M., Gervais, R., Ben-Porath, Y., Stafford, K., Freeman, D., et al. (2010).

Further validation of the MMPI-II and MMPI-II-RF Response Bias scale: Findings from

disability and criminal forensic settings. *Psychological Assessment: A Journal of

Consulting and Clinical Psychology, 22*(4), 745-756. doi:10.1037/a0020042

Table 1

Descriptive statistics, intercorrelations, and internal consistencies of the Level of Service subscales

Subscales	N	M	SD	Intercorrelations						
				1	2	3	4	5	6	7
1 Criminal History (8 items)	219	4.07	2.32	(.80)						
2 Family/Marital (4 items)	219	1.63	1.15	.259*	(.26)					
3 Companions (4 items)	219	1.84	1.40	.365*	.223*	(.36)				
4 Alcohol/Drug Problems (8 items)	219	3.89	2.42	.349*	.094	.463*	(-.46)			
5 Emotional/Personal (5 items)	81	2.17	1.46	.066	.152	.166	.149	(.68)		
6 Attitudes/Orientation (4 items)	219	1.13	1.19	.347*	.279*	.442*	.241*	.137	(.07)	
7 Antisocial Patterns (4 items)	137	1.67	1.35	.405*	.172*	.373*	.202*	-	.295*	(.60)

Note. $*p < 0.01$. Internal consistencies (Cronbach's alphas) are reported on the diagonal.
Note. Items for each subscale of the Level of Service measures: Criminal History (prior adult convictions, 2 or more convictions, 3 or more prior convictions, 3 or more present offences, arrested under 16, incarcerated upon conviction), Family/Marital (dissatisfaction with marriage, non-rewarding-parental, non-rewarding-other relatives, criminal-family/spouse), Companions (some criminal acquaintances, some criminal friends, few anti-criminal acquaintances, few anti-criminal friends.), Alcohol/Drug Problem (alcohol-ever, drug-ever, alcohol-currently, drug-currently, law violations, marital-family, school-work, medical), Emotional/Personal (moderate interference, severe interference-active psychosis, mental health treatment-present, psychological assessment indicated), Attitudes/Orientation (supportive of crime, unfavorable toward convention, poor-toward sentence, poor-toward supervision), and Antisocial Pattern (specialized assessment for antisocial pattern, early and diverse antisocial behavior, criminal attitude, pattern of generalized trouble). For the Cormier Lang scale, total violent and nonviolent offence scores are presented. For the Personality Assessment Inventory, *t*-scores are reported in means and standard deviations. For the Cormier Lang scale, total violent and nonviolent offence scores are presented.

Table 2

Correlations between the Level of Service subscales and the Personality Assessment Inventory scales and the Cormier Lang Criminal History Scores

Level of Service subscales	Personality Assessment Inventory (PAI)														Cormier Lang	
	ANT	ANT-A	ANT-E	ANT-S	NON	RXR	WRM	ALC	DRG	ANX	DEP	MAN	SCZ	AGG	Violent	Non-Violent
Criminal History	.318* (133)	.427* (128)	.210 (128)	.140 (128)	.110 (133)	-.206 (133)	-.206 (133)	.309* (133)	.359* (133)	.120 (133)	.192 (133)	.047 (133)	.163 (133)	.280* (133)	.313* (170)	.604* (170)
Family/Marital	.157 (133)	.143 (128)	.055 (128)	.161 (128)	.426* (133)	-.192 (133)	-.313* (133)	.055 (133)	.415* (133)	.449* (133)	.081 (133)	.283* (133)	.293* (133)	.276* (133)	.253* (170)	.110 (170)
Companions	.369* (133)	.304* (128)	.302* (128)	.279* (128)	.224* (133)	.151 (133)	-.297* (133)	.280* (133)	.401* (133)	.202 (133)	.264* (133)	.067 (133)	.272* (133)	.388* (133)	.264* (170)	.265* (170)
Alcohol/Drug Problems	.487* (133)	.429* (128)	.381* (128)	.419* (128)	.119 (133)	-.263* (133)	-.170 (133)	.657* (133)	.622* (133)	.235* (133)	.181 (133)	.169 (133)	.237* (133)	.435* (133)	.217* (170)	.120 (170)
Emotional/ Personal	-.006 (46)	.062 (45)	-.015 (45)	-.095 (45)	-.307 (46)	-.253 (46)	-.267 (46)	.151 (46)	.147 (46)	.226 (46)	.337 (46)	-.326 (46)	.371 (46)	-.075 (46)	.002 (55)	-.212 (55)
Attitude/ Orientation	.258* (133)	.182 (133)	.204 (133)	.245* (133)	.172 (133)	.011 (133)	-.179 (133)	.188 (133)	.150 (133)	.104 (133)	.130 (133)	.082 (133)	.236 (133)	.339* (133)	.181 (170)	.110 (170)
Antisocial Pattern	.383* (86)	.317* (82)	.229* (82)	.396* (82)	.106 (128)	-.113* (128)	-.103 (128)	.167 (128)	.200 (128)	.083 (128)	.132 (128)	.170 (128)	.075 (128)	.349* (128)	.280* (165)	.159 (165)

*p > 0.01

Note. Personality Assessment Inventory scales: Antisocial Features (ANT), Antisocial Behaviours subscale (ANT-A), Egocentricity subscale (ANT-E), Stimulus Seeking subscale (ANT-S), Alcohol Problem (ALC), Drug problem (DRG), Nonsupport (NON), Treatment Rejection (RXR), Warmth (WRM), Anxiety (ANX), Depression (DEP), Manic (MAN), Schizophrenic (SCZ), and Aggression (AGG). Sample size of correlational comparisons indicated within the parentheses.

Table 3

Correlations between the Level of Service subscales and the Minnesota Multiphasic Personality Inventory-Revised (MMPI-2) scales and the intelligence quotient scales

Level of Service subscales	Minnesota Multiphasic Personality Inventory – Revised (MMPI-2)										Intelligence measure		
	Pd 1	Pd 2	Pd 3	Pd 4	Pd 5	Dp	Pa	Sc	Ma	Si	Full Scale	Verbal IQ	Performance IQ
Criminal History	.068 (36)	.363 (36)	-.257 (36)	.368 (36)	.258 (36)	.260 (36)	.026 (36)	.043 (36)	.007 (36)	.421 (36)	-.227 (121)	-.295* (120)	-.250* (120)
Family/Marital	.617* (36)	.078 (36)	-.299 (36)	.458* (36)	.419 (36)	.473* (36)	.392 (36)	.401 (36)	.317 (36)	.410 (36)	-.123 (121)	-.126 (120)	-.188 (120)
Companions	.099 (36)	.258 (36)	-.301 (36)	.380 (36)	.279 (36)	.350 (36)	.155 (36)	.212 (36)	.114 (36)	.418 (36)	-.206 (121)	-.289* (120)	-.121 (120)
Alcohol/Drug Problems	.185 (36)	.243 (36)	-.236 (36)	.334 (36)	.348 (36)	.214 (36)	.110 (36)	.137 (36)	.098 (36)	.170 (36)	.031 (121)	-.035 (120)	-.002 (120)
Emotional/ Personal	-.038 (24)	-.110 (24)	-.080 (24)	.441 (24)	.211 (24)	.371 (24)	.228 (24)	.340 (24)	-.044 (24)	.237 (24)	-.139 (51)	-.210 (51)	-.377* (51)
Attitude/ Orientation	.275 (36)	-.068 (36)	-.264 (36)	.277 (36)	.074 (36)	.212 (36)	.230 (36)	.106 (36)	-.139 (36)	.234 (36)	-.114 (121)	-.189 (120)	-.166 (120)
Antisocial Pattern	.037 (36)	.213 (35)	-.096 (35)	.210 (35)	.214 (35)	.206 (35)	.111 (35)	.141 (35)	.150 (35)	.181 (35)	-.171 (118)	-.273* (117)	-.190 (117)

$*p < 0.01$

Note. Minnesota Multiphasic Personality Inventory – Revised (MMPI-2) scales: Familial Discord (Pd1), Authority Problems (Pd2), Social Imperturbability (Pd3), Social Alienation (Pd4), Self-Alienation (Pd5), Depression (Dp), Paranoia (Pa), Schizophrenia (Sc), Hypomania (Ma), Social Introversion (Si). Intelligence quotients taken from psychometric measures that include the Wechsler Adult Intelligence Scale, 3rd revision (WAIS-III) or Wechsler Abbreviated Scale of Intelligence (WASI). Sample size of correlational comparisons indicated within the parentheses.