Colorado Department of Corrections

Validation of the Level of Supervision Inventory (LSI) for Community Based Offenders in Colorado: Phase II

Maureen L. O'Keefe
Department of Corrections

Kelli Klebe
University of Colorado at Colorado Springs

Scott Hromas
Department of Corrections

April 1998

Acknowledgments

Many individuals have been involved with this pilot project since it commenced in May, 1995. A sincere thanks is extended to the Southeast Parole Region and ComCor. Without the consistent efforts of their parole officers and case managers, a study of this magnitude would not have been possible. Another thank you is expressed to those individuals who provided a peer review of this report, including Sean Ahlmeyer, Jean Peters, Birgit Fisher, Skye Babé, Charity Stoddard, and Karen Uhlmeyer. Finally, the continual support of Joe Stommel on this project is greatly appreciated.

Executive Summary

Until recently, agencies within Colorado's criminal justice system identified and assessed offender substance abuse needs differently. As a result of House Bill 1173 passing into law under Colorado Revised Statute 16-11.5 in 1991, a standardized assessment and treatment process was constructed. All jurisdictions composing the criminal justice system collaborated in its development. To match offenders to treatment services, the assessment battery measures two dimensions: criminal risk and substance abuse needs. The scores from these assessments are used to assign offenders to one of the seven classifications constituting the treatment system.

Purpose of Study

The Level of Supervision Inventory (LSI) was selected to evaluate the risk dimension of the standardized offender assessment. The present research evaluated the validity of the LSI as a risk assessment instrument with community based offenders in Colorado. The sample consisted of transitional community corrections and parole offenders. As currently applied within the Colorado criminal justice system, the LSI is a component of the standardized substance abuse evaluation. Nonetheless, the LSI was designed as a risk assessment instrument *not* restricted to substance abusing populations. Thus, this research considered the utility of the LSI in predicting recidivism rather than identifying and treating substance abusing offenders. The findings of this study will assist the Colorado Department of Corrections in making a decision regarding full implementation of the LSI as a classification system with community based offenders.

Design

The participants in this study included 172 parolees admitted to the southeast parole office and 85 transitional community corrections residents admitted to ComCor, a community corrections center, between November, 1994, and December, 1995. Offenders were assessed upon release to the community and after 6 months of community supervision. At Parole, TASC case managers conducted the initial LSI assessment). Parolees were also assessed on the Wisconsin classification system upon release (Time 1) and after approximately 6 months of parole supervision (Time 2). To establish the predictive validity of the LSI, two outcome variables were collected. The first included a 10-point rating of the offenders' compliance with their supervision, completed by the parole officer or ComCor case manager at 6 and 12 month follow-up periods. The second outcome variable was whether or not offenders were reincarcerated within a year of release to the community.

Findings

The primary function of the present study was to evaluate the predictive validity of the LSI with community based offenders in Colorado. Parolees and ComCor residents did not appear to differ in significant ways from each other, thereby supporting the belief that these groups represent the same population. Several important trends emerged from this study:

- **The LSI was predictive of outcome with parolees but not community corrections offenders.** The validity results of the LSI with ComCor offenders were contrary to the

study hypotheses. In fact, the findings were particularly curious because the two groups were similar and the LSI was valid with parolees. Differing test versions between agencies might account for the discordant validity findings. Moreover, exploratory analyses indicated that scores of ComCor assessments varied as a function of the individual conducting the assessment. Yet, there was no evidence to suggest that these fluctuations should relate to case managers' caseloads. In fact, a ComCor official reported that the agency attempted to neutralize any risk differences in the staff's caseloads. Interestingly, the LSI validity estimates were negatively related to the amount of scoring variability between assessors.

- **Time 2 LSI reassessments generally did not improve prediction of 12 month outcomes over the initial assessments.** Also contrary to the study hypotheses was the finding that reassessments were no more predictive of outcome than initial assessments. Although there was no significant change in parolees' risk scores across assessment periods, the initial assessment was a stronger recidivism predictor than the reassessment. At ComCor, there was a significant change in scores over time, but those changes did not produce significantly improved validity estimates. Unfortunately, this study was not able to support the dynamic nature of the LSI. On the other hand, neither could it dispute its dynamic attributes because scoring variations among assessors may have instead accounted for the low validity of reassessments.

- **The LSI was more predictive of recidivism than the Wisconsin for parolees.** While the Wisconsin reassessment displayed some improved validity over the initial assessment, the LSI remained the stronger predictor of risk at both assessment periods. A breakdown of initial classification levels based on the Wisconsin indicated that 98% of offenders scored in the maximum supervision range. This would suggest that the Wisconsin system is strongly influenced by raters and may lead to over-classification.

Implications & Recommendations

The LSI was found to be the best risk predictor for parolees when compared to the Wisconsin. These findings replicate research conducted with Canadian populations. Although the LSI was not found to be predictive with halfway house offenders in this study, its use with this population should not be discounted. Exploratory analyses revealed that the low predictive power may have resulted from individuals' assessment styles.

The LSI is not entirely objective. Truly, its strengths seem to lie in quantifying subjective material. Neither subjective clinical decisions nor objective, static ratings have led to more accurate predictions than the LSI. Yet, this poses a problem when the instrument is administered by individuals lacking assessment or clinical training. The most predictive assessments in this study were conducted by TASC case managers, all who were seeking an advanced degree in a counseling field. Hence, it is probable that previous clinical training impacts the validity of the assessments.

These findings taken together highlight the need for rigorous quality assurance and training. It is not the purpose of this research to suggest that only individuals with advanced counseling training be the sole assessors. However, the criminal justice system must consider the resources

required for individuals with less assessment training to conduct the LSI. A thorough training and quality control process needs to be implemented across all jurisdictions. Even for individuals with advanced training, it is recommended that continuous quality control be provided as numerous scoring errors were evident even with TASC assessments.

Several areas of research could expand the knowledge base of the standardized offender assessment. First, further examination of the LSI's inter-rater reliability is recommended. Second, the use of the LSI with community corrections offenders needs to be explored more fully. Thirdly, the utility of the LSI and the other battery instruments in matching identified substance abusers to treatment needs to be researched. Finally, any consideration by the Department of Corrections to implement the LSI with parolees should include the task of developing norms. It would be unwise to accept norms developed for Canadian offenders, a population likely to differ significantly from Colorado felons. Norms need to be tailored for the population involved and should take into consideration the philosophy of the agency using them.

The pendulant shift within the criminal justice system from a punitive focus to a rehabilitative one is reflected in the increasing reliance on offender assessments. Assessment of offenders frequently involves the evaluation of risk level and identification of needs areas. Risk and needs assessments serve not only to maintain greater public safety, but also to facilitate the rehabilitation of offenders through improved case planning. Colorado has witnessed an increased emphasis on offender assessments during this decade.

Standardized Assessment of Colorado Felons

Until recently, agencies within Colorado's criminal justice system identified and assessed offender substance abuse needs differently. As a result of House Bill 1173 passing into law under Colorado Revised Statute 16-11.5 in 1991, a standardized assessment and treatment process was constructed. All jurisdictions composing the criminal justice system collaborated in its development.

The assessment battery was constructed to optimize resources while ultimately matching offenders' treatment needs to available services. Thus, a triage model was established so that the entire battery is not administered to every offender. To match offenders to treatment services, the battery measures two dimensions: criminal risk and substance abuse needs. Consequently, offenders who pose the greatest risk and suffer the most serious substance abuse problems receive more intensive treatment services than those with less serious risk and substance abuse problems.

Five instruments were initially selected to compose the battery. All felons, except Class I offenders, require assessment on the Alcohol Dependence Scale (ADS; Skinner & Horn, 1984), the Drug Abuse Screening Test (DAST; Skinner, 1982), and the Level of Supervision Inventory (LSI; Andrews, 1982). The LSI measures an offender's criminal risk for re-offense. The ADS and DAST screen for substance abuse problems; scores of 5 or greater on either scale indicate the need for further assessment. The Adult Substance Use Survey (ASUS; Wanberg, 1992) and the Substance Use History Matrix (SUHM; Bogue & Timken, 1993) are administered to offenders when the screening tests indicate a problem with chemical dependency.

These assessments are used to assign offenders to one of the seven classifications constituting the treatment system in Colorado. They range in intensity from no treatment to therapeutic community placement. This treatment classification system was based on the patient placement criteria of the American Society of Addiction Medicine (1996). The assessment scores are key for matching offenders' needs with the appropriate treatment level. Resultant scores are entered into a formula from which a treatment level is derived. The assessor may recommend the offender to another treatment modality through an override with collaborating clinical rationale.

The challenging implementation of the standardized offender assessment and treatment system is a continual evolution. The theoretical foundation of the current system contributes to its face validity. Yet, to establish the validity of the entire system, it is necessary to analyze the instruments individually and then in combination with each other. The science of psychometrics demands further review of the system by evaluating the instruments' reliability and validity. This

study seeks to evaluate the LSI as a risk assessment instrument for a Colorado community-based offender population.

Assessment of Risk

Correctional institutions rely upon risk assessments to predict the likelihood of a criminal committing more crimes upon release. Risk assessments promise to have the most utility in making decisions regarding the level of correctional supervision assigned to convicted offenders. Although the judicial process is responsible for the sentencing of offenders, the onus lies with corrections officials to determine each offender's appropriate supervision level. Errors are costly in terms of public safety, institutional image, and recidivism.

Historically, risk assessments were contingent upon the professional judgment of corrections personnel. Subjective assessments involve gathering data and interpreting it in a meaningful manner (Bonta, 1993). Unfortunately, subjective assessments are susceptible to a multitude of biases, including the assessor's professional and personal experiences, types of questions asked, initial impressions, level of rapport, and theoretical perspective (Groth-Marnat, 1990). Furthermore, such assessment are difficult to uphold legally.

More recently, the increased emphasis on making accurate risk appraisals shifted the focus from subjective to objective assessments. Objective assessments are actuarial in nature and rely chiefly on factors found to be empirically related to recidivism. The actuarial approach is not subject to the same inequities as the subjective approach, and it has substantial support as the better approach statistically (Goldberg, 1965; Meehl, 1954, 1965; Wiggins, 1973). Yet, there are limitations associated with objective assessments; they can not probe certain areas, they are reliant upon static, historical data, and they can not quantify important idiosyncratic events (Groth-Marnat, 1990). Moreover, objective assessments provide little direction for treatment (Bonta, 1993).

Both subjective and objective assessments are practiced today in corrections. However, the ideal assessment process couples subjective information with objective data. Together, both types of data can be used to assess risk of re-offense and treatment needs. Bonta (1993) speculated that needs are not fundamentally different from risk; rather, needs are linked to criminal behavior. Furthermore, he surmised that risk levels can be lessened by addressing those needs. In this spirit, the LSI was designed as a risk and needs tool.

Level of Supervision Inventory

The LSI was developed in the late 1970s in Canada through the collaboration of probation officers, correctional managers, practitioners, and researchers (Andrews, 1982). The Ministry of Correctional Services in Ontario implemented the LSI as the classification instrument for probationers and parolees (Andrews). Now the LSI is being used in various correctional settings across the United States.

The LSI is administered through a semi-structured interview. Information is obtained from the offender and is verified through official records and other sources of information whenever

possible. The LSI is comprised of 54 items across ten different subcomponents. Each item is scored either 0 or 1, where a point indicates that an item is true. After each item is scored, the points are totaled to obtain a composite risk score. Higher scores are indicative of greater recidivism risk.

The LSI was designed to be a comprehensive risk and needs assessment rather than a short checklist of the most predictive risk items (Andrews, 1982). To this end, the LSI can aid officers who must make decisions regarding interventions to improve offenders' outcomes. The following ten LSI subcomponents illustrate its broad spectrum: criminal history, education/employment, financial, marital/family, accommodations, leisure/recreation, companions, alcohol/drug problems, emotional/personal, and attitudes/orientation. Although change in some areas may impact recidivism risk more than others, Andrews recommends using the subtotals for developing case plans to impact the offender's functioning.

An instrument which embodies both risk and needs assessment must consider those items which best predict risk and survey various needs domains. Static items are often the most powerful predictors of risk, as past behavior is the best predictor of future behavior. On the other hand, needs assessments warrant evaluation of changing, or dynamic, realms. The LSI has both static and dynamic items. Thirteen of the 54 items are considered highly sensitive to change in offenders. These items are rated on a scale from 0 to 3 (in addition to the item score). The 13 ratings are then totaled to obtain a rater score. Higher scores are suggestive of more pro-social influences in an offender's life. The rater score may provide a more dynamic measure of offenders' risk over multiple assessments than the risk score itself. Unfortunately, no research has evaluated the reliability and validity of the rater score.

Psychometric Characteristics of the LSI

Reliability. Internal consistency is a fundamental estimate of an instrument's reliability. Internal consistency describes the degree of correlation between the items on a scale. Several studies have investigated the internal consistency of the LSI composite score. The alpha coefficient was .72 for Canadian probationers (Andrews, 1982), .81 for Canadian federal inmates (Loza & Simourd, 1994), .80 for Colorado inmates and .78 for Colorado parolees (Arens, Durham, O'Keefe, Klebe, & Olene, 1996). These relatively high estimates not only suggest an internally consistent inventory, but also imply that the factors contributing to recidivism risk, as measured by the LSI, are related to each other.

The most important aspect of reliability for the LSI is inter-rater reliability. Institutions endorsing the LSI should not have to worry about whether different assessors will rate the same offender differently. The strength of a risk instrument used for classifying offenders to a supervision level lies in its ability to produce stable ratings, regardless of the assessor. In the initial LSI validation study, Andrews (1982) examined the inter-rater reliability of the LSI. The correlation coefficients ranged from .94 with different assessors and no time difference to .80 with different assessors and a duration of two or more months between assessments. While these estimates are sufficiently high, it should be noted that no other studies were uncovered in the literature which attempted to replicate these findings.

Underlying structure. The ten subcomponents of the LSI are based on a logical grouping of items rather than a statistical basis. The arrangement of items into subgroups aids administration of the instrument and lends itself to assessing needs. Several studies have examined the statistical foundation of the ten subscales. In the first study conducted with Canadian probationers, subtotal intercorrelations ranged from .04 to .43 and subtotal-total correlations ranged from .37 to .79 (Andrews, 1982). A study conducted with Canadian federal inmates found subscale intercorrelations ranging from .02 to .51, subtotal-total correlations of .45 to .77, and Cronbach's alpha for each subscale of .84 to .86 (Loza & Simourd, 1994). Another study examined Colorado inmates and parolees (Arens et al., 1996). For the inmate population, intercorrelations between the subscales ranged from -.13 to .38, subtotal-total correlations ranged from .31 to .70, and Cronbach's alpha for each subscale ranged from .20 to .73. The parole population had subcomponent intercorrelations of .04 to .30, subtotal-total correlations of .30 to .65, and internal consistency estimates of the ten subscales ranging from .08 to .72. Notably, the most inconsistent results across these four groups were the substantially higher subscale internal consistency estimates of the Canadian federal inmate population over the other groups. These findings alone neither support nor negate the statistical existence of the subscales.

The underlying structure of the LSI was examined in two studies using factor analysis of all 54 items. Two factors emerged with Canadian federal inmates: "Criminal Lifestyle" which accounted for 27% of the variance and "Emotional/Personal Problems" which accounted for 23% of the variance (Loza & Simourd, 1994). However, this study had only 3 cases per item as opposed to the recommended 10 cases per item for this type of analysis. The Colorado study met this statistical assumption, yet found no factors in a factor analysis of incarcerated offenders' item scores (Arens et al., 1996).

Two studies used a factor analysis on the subscale scores to probe the underlying dimensions. The first study, conducted with Canadian probationers (Andrews & Robinson, 1984), offered a three factor solution with the first factor accounting for 75% of the variance. Subscales that loaded on each factor included: (1) companions, leisure/recreation, and attitudes, (2) education/employment, financial, accommodation, and marital/family, and (3) alcohol/drugs, emotional/personal, and criminal history. The second study involved two separate analyses with inmates and parolees (Arens et al., 1996). The analysis with the inmate population identified two factors: (1) "Needs" with loadings from the education/employment, financial, marital/family, accommodation, leisure/recreation, alcohol/drugs, and emotional/personal subscales, and (2) "Risk" with dominant subscales including criminal history, companions, and attitude/orientation. The parolee analysis revealed a three factor solution. Subscales that loaded on each factor included: (1) family/marital, leisure/recreation, and attitude/orientation, (2) criminal history and alcohol/drugs, and (3) education/employment, financial, accommodation, and companions.

There is no support for any underlying dimensions of the LSI. Three studies, involving five separate analyses, aimed at establishing the underlying structure produced no concordant findings. Additionally, the relatively high internal consistency of the LSI suggest that it is homogeneous in its assessment of risk. It can be concluded that the LSI is a unidimensional scale; the greatest confidence can be placed in the composite risk score.

Validity. Validity is an indicator of whether a test actually measures what it is supposed to measure. An instrument must first be found reliable, as there can be no validity without reliability. Because the reliability of the subcomponents was not substantiated, the present study only reviewed the validity of the composite score. Numerous LSI validation studies have been conducted over the past two decades. The research has primarily been conducted with Canadian offender groups.

The convergent validity of the LSI to other instruments measuring similar constructs was evaluated. In a meta-analysis summarizing the validity of various risk predictors, the LSI was found to be the best predictor of risk when compared to other risk assessment instruments, antisocial personality scales, static characteristics, and dynamic factors (Gendreau, Little, & Goggin, 1996). Nevertheless, these findings should be interpreted lightly as there are numerous problems associated with meta-analytic techniques (see Nunnally & Bernstein, 1994). Loza and Simourd (1994) correlated the LSI to several validated scales for a Canadian federal inmate population. The LSI had high correlations with the Psychopathy Checklist-Revised total score (r = .78; Hare, 1991). The LSI had a moderate negative correlation with the General Statistical Information on Recidivism Scale where lower scores indicate greater recidivism risk (r = -.62; Nuffield, 1982) and low positive correlations with the Millon Clinical Multiaxial Inventory (Millon, 1982) antisocial personality (r = .14), passive-aggressive (r = .38), alcohol dependence (r = .20), and drug dependence (r = .38) scales.

The LSI was shown to successfully predict outcomes for probationers. The LSI predicted success of supervision (r = .35), in-program outcome status (r = .47), and in-program recidivism (r = .38) in the initial LSI validation study (Andrews, 1982). A longer term follow-up of this validation sample (Andrews & Robinson, 1984), averaging more than two years after assessment, revealed a strong correlation of the LSI to recidivism (r = .43) and incarceration (r = .37). Interestingly, this study found that the magnitude of the correlations between LSI scores and outcome was greater for female offenders than male offenders. In a cross-validation sample, Andrews and Robinson (1984) found a strong relationship of LSI levels to in-program recidivism rates (r = .39) and a moderate relationship to self-reported, undetected reoffenses (r = .22).

Similar to probationers, the LSI was found to be predict outcomes for halfway house participants. The instrument predicted successful completion of halfway house programs and reincarceration following release from halfway house facilities (Bonta & Motiuk, 1985, 1987, 1990). In contrast to the substantial support for the LSI with Canadian halfway house offenders, a study of Colorado offenders sentenced to a community corrections facility revealed that the LSI was not predictive of either program termination status (r = .10) or institutional misconduct (r = .16; Philbrick, Gat, & Guisti, 1993).

With the rising interest in applying the LSI to different offender populations, research of the LSI has extended beyond probationers and halfway house offenders to inmate populations. LSI scores were found to be predictive of institutional misconducts, parole violations, and reincarceration rates for the year following release (Bonta & Motiuk, 1992).

These findings lend evidentiary support for the use of the LSI as a risk prediction tool. However, the limited scope of the research samples presents a drawback. First, no research studies were

found in the literature which evaluated American offender populations. Secondly, no adequate research studies evaluated the LSI's utility with community based offenders following release from prison (i.e., parolees and transitional halfway house residents). Finally, Canadian inmate samples were taken from the provincial system except for one study which used inmates in the federal system (Loza & Simourd, 1994). The provincial system houses offenders sentenced for less than two years while the federal system houses offenders who must serve two or more years. Thus, there is an important need for Colorado Department of Corrections to research the use of the LSI with parole and transitional community corrections offenders.

Purpose of Study

The present research evaluated the validity of the LSI as a risk assessment instrument with community based offenders in Colorado. The sample consisted of transitional community corrections and parole offenders. As currently applied within the Colorado criminal justice system, the LSI is a component of the standardized substance abuse evaluation. Nonetheless, the LSI was designed as a risk assessment instrument *not* restricted to substance abusing populations. Thus, this research considered the utility of the LSI in predicting recidivism rather than identifying and treating substance abusing offenders. The findings of this study will assist the Colorado Department of Corrections in making a decision regarding full implementation of the LSI as a classification system with community based offenders.

It was hypothesized that high scores on the LSI risk scale would be associated with poor outcomes for parole and community corrections offenders. The LSI rater score was also evaluated as a risk predictor; it was postulated that lower scores on this scale would correlate with poorer outcomes. Although both the risk and rater scores of the LSI were analyzed, the risk score was expected to be the better predictor of the two because it is more objective and includes more items. It was also theorized that reassessments would improve risk predictions over the initial assessments. Finally, it was hypothesized that the LSI would provide more accurate risk predictions with parolees than the currently employed Wisconsin classification system.

Method

Participants

The participants in this study included felony offenders under community supervision in Colorado Springs. This sample consisted of 172 parolees admitted to the southeast parole office between November, 1994, and December, 1995. In addition to the parolee sample, 85 offenders admitted from May, 1995, through December, 1995, to ComCor, a community corrections center located in Colorado Springs, were included. The final participant count was 257. Some participants were under the supervision of both Parole and ComCor within the scope of this study. For the purposes of this research, offenders were considered participants only for the agency to which they were first admitted.

The participants in this study may not be representative of all offenders released to Parole and ComCor. A total of 37 offenders was excluded from the study for the following reasons. Three ComCor residents and twenty parolees were not assessed on the LSI. The ComCor residents were arrested or discharged from the facility before assessment was possible. Of the 20 parolees not assessed, 10 offenders absconded or were arrested prior to assessment. The remaining 10 parolees were not assessed for various reasons such as a regional transfer or lack of referral to the case management agency conducting the initial LSI assessments. Furthermore, 10 parolees who did not receive an assessment within 45 days of release to the community were excluded from the study. One offender died during the course of the study and was therefore excluded from the analyses.

While screening the data, three outliers were discovered and consequently eliminated from the study. The outliers became evident when examining the bivariate relationship between risk scores and outcome variables. Outliers beyond 2.9 standard deviations on the risk score were eliminated from the study because they unduly impacted the validity analyses. Two outliers were parolees and one was a ComCor resident. All outliers were cases of offenders with low LSI risk scores (and correspondingly high LSI rater scores) combined with poor outcomes.

Age, ethnicity, gender, and marital status were examined for both participant groups. The demographic characteristics for the parolee sample were as follows: the mean age of participants was 33.2 (n = 172); males comprised 86% of the sample (n = 148), females 14% (n = 24); 46% were Caucasian (n = 79), 34% African American (n = 59), 17% Hispanic (n = 30), and 2% of another racial background (n = 4); and 47% were single (n = 80), 22% married (n = 37), 9% common law married (n = 15), and 22% separated, divorced, or widowed (n = 38). Marital status data was missing for two parolees. The demographic characteristics for ComCor residents were as follows: the mean age was 31.9 (n = 85), 86% were males (n = 73) whereas 14% were females (n = 12); 56% were Caucasian (n = 47), 25% African American (n = 21), 17% Hispanic (n = 14), and 2% Native American (n = 2); and 41% were single (n = 34), 31% married (n = 26), 7% common law married (n = 6), and 21% separated, divorced, or widowed (n = 18). Ethnicity and marital status information were missing for one offender.

The demographic characteristics of parole and ComCor were compared to determine if differences existed between the groups. Only offenders of Caucasian, African American, or

Hispanic descent were included in the ethnicity analysis in order to meet the assumptions of the chi-square test. There were no significant differences between groups for ethnicity, χ^2 (2, N = 250) = 2.74, p = .26, gender, χ^2 (1, N = 257) = .00, p = .97, marital status, χ^2 (3, N = 254) = 2.67, p = .45, or age, t (255) = -1.23, p = .22.

Measures

The ADS is a 25-item multiple choice questionnaire which screens for alcohol abuse and dependence (Skinner & Horn, 1984). Test takers self report their use of alcohol and related symptoms over the previous 12 months. The ADS was derived from the Alcohol Use Inventory (Wanberg, Horn, & Foster, 1977) which measures the following four dimensions: (1) loss of behavioral control, (2) psycho-perceptual withdrawal symptoms, (3) psycho-physical withdrawal, and (4) obsessive-compulsive drinking. The ADS was found to have high internal consistency (Skinner & Allen, 1982), high concurrent validity with the Michigan Alcoholism Screening Test (Selzer, 1971; Ross, Gavin, & Skinner, 1990), and low concurrent validity with cognitive functioning, psychiatric symptoms, and physical symptoms (Kivlahan, Sher, & Donovan, 1989).

The DAST is a self report survey which screens for drug abuse (Skinner, 1982). There are 20 items to which individuals respond "yes" or "no" about their use during the last 12 months. The DAST yields an index of the severity of drug problems. The DAST was found to have high internal consistency (Skinner) and moderate to high concurrent validity with the five Diagnostic and Statistical Manual III (DSM-III; American Psychiatric Association, 1980) criteria for drug abuse/dependence (Gavin, Ross, & Skinner, 1989).

The ASUS is a self-report instrument which quantifies individuals' alcohol and other drug use over a lifetime (Wanberg, 1994). The instrument measures five domains: (1) involvement in ten categories of drugs, (2) degree of disruption resulting from use of drugs, (3) antisocial attitudes and behavior, (4) emotional or mood adjustment difficulties, and (5) defensiveness and resistance to self-disclosure (Wanberg, 1997). Internal consistency estimates for a criminal justice population for each of the five subscales was high (Wanberg, 1997). The scales had moderate to high concurrent validity estimates with prior inpatient treatment, low estimates with prior outpatient treatment, and moderate to high estimates with the ADS and DAST (Wanberg, 1997).

The Wisconsin classification system (Baird, 1981), currently employed with Colorado parolees, incorporates separate risk and needs scales to classify offenders to a supervision level. The risk scale evaluates offenders' likelihood of recidivism. There are two versions of the risk scale; one for the initial assessment and one for reassessments. On the initial risk assessment, offenders are evaluated on criminal history, drug and alcohol problems, employment, number of address changes, and attitude. The reassessment evaluates the same factors; however, it also assesses offenders' current living situation, social identification, response to court or parole board conditions, and use of community resources. The Wisconsin needs component assesses the following 12 areas: academic/vocational skills, employment, financial management, marital/family relationships, companions, emotional stability, alcohol usage, other drug usage, mental ability, health, sexual behavior, and agent's impression of client's needs. Offenders are

classified into minimum, medium, or maximum categorizes on both dimensions. Whichever dimension is higher dictates the supervision level, although officer overrides are permissible.

Procedures

All participants under parole supervision were referred to the Treatment Accountability for Safer Communities (TASC) program, a case management agency closely linked with the parole office, for the initial LSI assessment (Time 1). Approximately 6 months following the Time 1 assessment, parole officers completed a LSI reassessment for participants still under parole supervision (Time 2). The Wisconsin risk assessment is routinely completed by parole officers for all parolees. The instrument, used for classification purposes, is administered within 30 days of release to parole and readministered every 6 months, although reassessments may occur earlier as needed. Therefore, parolees received LSI and Wisconsin assessments upon release to the community and after approximately six months of community supervision.

ComCor participants are regularly assessed on the LSI by their case manager upon admission to the facility. The case managers use the LSI-Revised (Andrews & Bonta, 1995) form in lieu of the Colorado version. This version is easier to use which results in fewer scoring errors. However, it is accompanied by an interview guide that promotes a structured, rather than semi-structured, interview format. ComCor case managers conduct LSI reassessments at 3 month and 6 month intervals, dependent upon offenders' LSI scores. LSI scores of 29 or greater require a 3 month reassessment whereas scores less than 29 require a 6 month reassessment. Consequently, all ComCor subjects had a LSI assessment at Time 1, but not all of them had a 6 month reassessment (Time 2) even if they were still residing at ComCor.

Time constraints impacted the assessments included in this study. As previously stated, 10 offenders were excluded from the parolee sample for untimely Time 1 LSI assessments. Initial Wisconsin assessments were included in the analyses only if completed within 45 days of the offender's release to the community; however, participants were not excluded from the study for missing or tardy Wisconsin assessments. Only reassessments conducted 6 months after the initial LSI, with a margin of 45 days, were included. Again, subjects were not eliminated from the study if their reassessments did not meet this time criteria, but the reassessments constituted missing data.

Two outcome measures were collected to establish predictive validity. The first measure consisted of a rating from zero to nine of the participants' compliance with supervision over the previous 6 month period. Table 1 presents definitions for the supervision compliance ratings. Parole officers and ComCor case managers rated participants at 6 and 12 month follow-up periods. The second outcome variable was a dichotomous measure of recidivism, or return to prison. For the purposes of this study, recidivism was considered to be regression to prison within one year of the intake assessments, including any arrests or escapes within a year that led to a regression.

Table 1. Supervision Compliance Ratings

Rating	Definition

0	100% compliance; no violations
1	Technically some minor noncompliance; no formal actions
2	Formal violation; no sanctions or regression
3	Formal violation; sanctions but no regression
4	Formal violation; sanctions with revocation proceedings but no regression
5	Formal violation; regression and/or revocation of supervision status
6	Absconsion or escape but reappeared within 48 hours
7	Absconsion or escape with no return within 48 hours
8	Charged with a new offense / misdemeanor
9	Charged with a new offense / felony

Participant information, other than assessment data and supervisory ratings, was collected from the Department of Corrections Information System, a computerized database management system.

Quality Control Procedures

Although more objective than early risk assessments, the LSI has some subjective features in administration and scoring. Consequently, quality assurance was conducted in this study to ensure the integrity of the data. Two main areas were involved in the quality control process. First, rater boxes must be congruent with the item scores. Thirteen items on the LSI have rater boxes requiring a rating of prosocial behavior. Items that are scored a point should have a rating of 0 or 1 while items that are not scored a point should have a rating of 2 or 3. For example, an item that scored a point but had a rating of 2 constituted a rater box incongruency. Secondly, there are nine syllogistic rules where the scoring of one item is contingent upon the score of another item (see Table 2).

Table 2. LSI Syllogistic Rules

1. If #2 is scored, then #1 must be scored
2. If #3 is scored, then #2 must be scored
3. #6 must be scored for all incarcerated offenders
4. If #11 is scored, then #s 18, 19, and 20 must have rater box values of 0
5. If #34 is scored, then #33 must be scored
6. If #35 is scored, then #36 must be scored
7. If #39 is scored, then #37 must be scored
8. If #40 is scored, then #38 must be scored
9. If #47 is scored, then #46 must be scored

In the present study, continuous quality assurance was conducted for parolee assessments. This involved returning LSI assessments to assessors for corrections when missing item scores, missing rater box values, rater box incongruencies, and violations of the syllogistic rules were found. However, for assessments collected archivally, including those conducted prior to the inception of this study in June 1995, it was not feasible to return assessments for correction. Nonetheless, efforts were made to minimize the number of errors on these assessments. All of these forms had at least one missing rater box for which it was impossible to obtain ratings. There were some incongruencies between rater box scores and item scores. These pertained only

to ratings done by one case manager who used reverse scoring on the rater boxes. Consequently, those rater box scores were recoded to provide more accurate data. Moreover, there were some syllogistic rule violations pertaining to rules 4, 6, 7, and 8 as listed in Table 2. The TASC case managers kept interview notes; it was possible to use these to correct most of the scoring errors. Corrections to the assessments were made only when notes were clear enough to make these changes with confidence. For example, for an offender who was scored on #40 (current drug problem) but not #38 (drug problem, ever), the score on #38 was changed if the LSI interview notes clearly stated the existence of a drug problem. These quality assurance procedures successfully reduced the number of scoring errors even though it did not altogether eliminate them.

The quality control process was limited with ComCor assessments. Because of the test version used at ComCor, few scoring errors resulted. Most of the scoring problems included errors of omission, and in these cases, the assessments were returned for completion.

Results

Risk Assessment Profiles

LSI scores were examined across groups to delineate risk levels of community based offenders. Both risk and rater scores from the LSI were examined. For parole subjects, Wisconsin scores were also reviewed. Table 3 presents descriptive statistics for assessments conducted at Time 1 and Time 2. The number of offenders receiving a Time 2 reassessment was lower than those assessed upon intake due to recidivism, transfers, and discharges from community supervision. Moreover, any assessments with a missing rater box value were excluded from all analyses involving the rater score.

Table 3. Risk Scores

	Parolee Sample			ComCor Sample		
	M	*SD*	*n*	*M*	*SD*	*n*
Time 1						
LSI Risk Score	23.7	8.0	172	23.9	6.5	85
LSI Rater Score	20.6	7.4	79	17.8	5.4	85
Wisconsin Risk	23.5	5.0	160	--	--	--
Wisconsin Needs	28.5	6.0	160	--	--	--
Time 2						
LSI Risk Score	23.7	9.1	105	17.2	6.4	34
LSI Rater Score	21.0	6.8	102	27.8	4.4	34
Wisconsin Risk	13.6	4.7	77	--	--	--
Wisconsin Needs	21.3	9.0	77	--	--	--

Several analyses were conducted to make comparisons between the two groups. LSI risk scores were similar across groups at Time 1; however, a *t* test was performed to determine whether the groups differed on reassessment risk scores. ComCor residents had significantly lower risk scores at Time 2 than parolees, $t(80.26) = -4.64$, $p = .00$. Comparisons across groups on the LSI rater score revealed that ComCor offenders had significantly lower scores than parolees at Time 1, $t(142.59) = -2.73$, $p = .01$, and significantly higher scores than parolees at Time 2, $t(88.46) = 6.70$, $p = .00$.

To determine whether either offender population changed over time, matched sample *t* tests were conducted with the LSI. Parolees had no significant change in risk scores, $t(104) = -1.86$, $p = .07$, or rater scores, $t(47) = 1.63$, $p = .11$, over the 6 month period. On the other hand, ComCor residents had a significant decline in risk scores, $t(33) = 8.88$, $p = .00$, and a significant increase in rater scores, $t(33) = -14.63$, $p = .00$, over the same 6 month period. Additional analyses were conducted to discern changes in parolees' Wisconsin risk and needs scores. Parolees had significantly lower risk scores, $t(73) = 16.60$, $p = .00$, and needs scores, $t(73) = 8.11$, $p = .00$, at Time 2 than Time 1.

Substance Abuse Assessment Profiles

Scores on the ADS, DAST, and five ASUS subscales were obtained at Time 1 for all participants. Table 4 presents the mean, median, standard deviation, sample size, and possible range of each instrument for both groups.

Table 4. Substance Abuse Scores

	Parolee Sample				ComCor Sample				Possible
	M	*Mdn*	*SD*	*n*	*M*	*Mdn*	*SD*	*n*	Range
ADS	1.4	0	3.2	157	2.7	0.0	4.2	84	0-44
DAST	3.7	2	4.3	159	4.8	3.0	4.6	85	0-20
ASUS Involvement	8.8	7	7.0	162	8.5	8.5	6.3	82	0-40
ASUS Disruption	12.9	7	14.9	162	11.3	5.0	13.7	82	0-80
ASUS Social	8.6	8	4.4	162	7.7	7.5	3.7	82	0-30
ASUS Emotional	4.0	3	3.8	161	2.8	2.0	2.7	82	0-27
ASUS Defensive	7.6	7	3.1	161	7.7	7.5	2.8	82	0-15

Comparative analyses were conducted to determine whether differences existed between groups on any substance abuse measure. Significant differences were found between groups on the ADS, $t(135.67) = 2.42$, p = .02, and the ASUS emotional subscale, $t(218.45) = -2.78$, $p = .01$. Nonetheless, these differences were slight; there was a mean difference of 1.3 for the ADS and 1.2 for the ASUS emotional subscale. No differences were found between the groups for the DAST, $t(242) = 1.88$, $p = .06$, or the ASUS involvement, $t(242) = -.31$, $p = .76$, disruption, $t(242) = -.81$, $p = .42$, social, $t(242) = -1.62$, $p = .11$, and defensive subscales, $t(241) = .32$, $p = .75$.

Predictive Validity Estimates

Outcome data was analyzed for each group. Forty two percent of ComCor offenders and 36% of parolees were reincarcerated within a year of release. The other outcome variables, supervision ratings (as defined in Table 1), were analyzed for each group at two different periods. Figures 1 and 2 display these outcome ratings.

Predictive validity of the LSI and Wisconsin instruments was examined using three outcome variables. For the LSI, both risk and rater scores were analyzed. Pearson's correlations of the test scores to the outcome measures were analyzed by agency for each test administration (see Table 5).

Figure 1. Parole Supervision Outcome Ratings

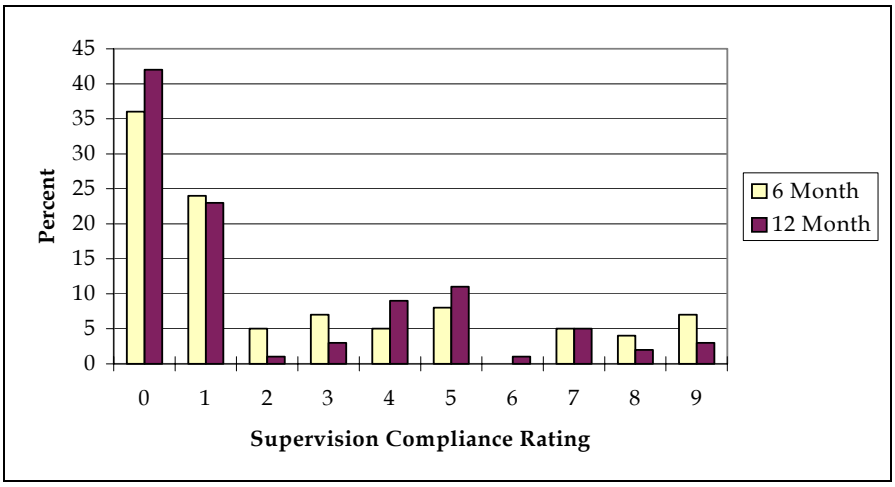


Figure 2. ComCor Supervision Outcome Ratings

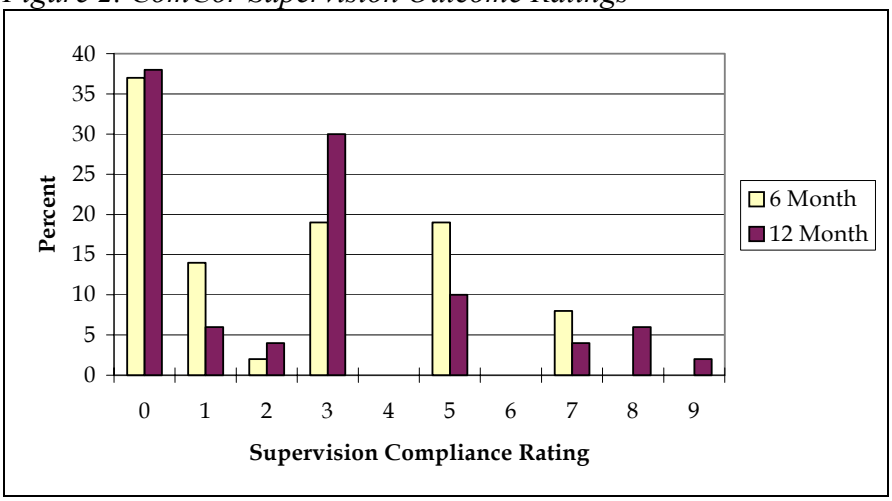


Table 5. Pearson's Correlations of Risk Scores to Outcome Measures

	Parole Participants					
	Time 1 Rating		Time 2 Rating		12 Month Recidivism	
	r	*n*	*r*	*n*	*r*	*n*
Time 1						
LSI Risk Score	.37**	169	.37**	107	.31**	172
LSI Rater Score	-.40**	76	-.36*	50	-.36**	79
Wisconsin Risk	.12	157	.06	99	.09	160
Wisconsin Needs	.21**	157	.01	99	.10	160
Time 2						
LSI Risk Score	--	--	.29**	103	.22*	105
LSI Rater Score	--	--	-.18	100	-.18	102
Wisconsin Risk	--	--	.17	75	.11	77
Wisconsin Needs	--	--	.27*	75	.12	77

(table continues)

Table 5. Pearson's Correlations of Risk Scores to Outcome Measures (continued)

	ComCor Participants

	Time 1 Rating		Time 2 Rating		12 Month Recidivism	
	r	*n*	*r*	*n*	*r*	*n*
Time 1						
LSI Risk Score	.19	84	-.02	50	.08	85
LSI Rater Score	-.20	84	.06	50	-.15	85
Time 2						
LSI Risk Score	--	--	.30	33	.11	34
LSI Rater Score	--	--	-.17	33	-.23	34

* $p < .05$.
** $p < .01$.

To explore whether the LSI could differentiate between recidivists and non-recidivists, mean risk scores were compared for each group. Offenders were classified as recidivists and non-recidivists according to the 12 month dichotomous outcome variable. Within each agency, mean LSI scores at both administration periods were analyzed. Figure 3 presents the mean scores for parolees while Figure 4 presents the mean scores for ComCor residents.

A series of *t* tests was conducted to determine where the mean differences lie. Within the parole sample, recidivists had significantly higher LSI risk scores than non-recidivists at Time 1, $t(170) = -4.32$, $p = .00$, and Time 2, $t(103) = -2.33$, $p = .02$. When examining the LSI rater score with parolees, the recidivist group had significantly lower scores than the non-recidivist group at Time 1, $t(77) = 3.40$, $p = .00$, but not at Time 2, $t(100) = 1.78$, $p = .08$. With ComCor participants, there were no significant differences between recidivists and non-recidivists on the LSI risk score at Time 1, $t(83) = -.70$, $p = .49$, or Time 2, $t(32) = -.62$, $p = .54$, or on the LSI rater score at Time 1, $t(83) = 1.33$, $p = .19$, or Time 2, $t(32) = 1.36$, $p = .18$. Finally, Wisconsin risk and needs scores were examined differentially for recidivists and non-recidivists (see Figure 5). There were no significant differences between the groups on the risk measure at Time 1, $t(158) = -1.11$, $p = .27$, or Time 2, $t(75) = -.94$, $p = .35$, or on the needs measure at Time 1, $t(158) = -1.23$, $p = .22$, or Time 2, $t(75) = -1.01$, $p = .32$.

Because the two offender groups appeared to be similar across numerous measures, further analyses were conducted to examine why the LSI was not predictive with ComCor offenders as it was with parolees. ComCor case managers reported that they found the LSI to be confusing and difficult to administer. Thus, it was hypothesized that scores varied as a function of the assessor. To test this postulate, mean LSI risk scores were examined by assessor for those who conducted more than five assessments within an administration period. Figures 6 and 7 display box plots of the distributions of LSI scores.

Four analysis of variance (ANOVA) tests were conducted to check for scoring variations among assessors. The first two statistical tests involved parolee assessments. Four TASC assessors conducted the bulk of the initial LSI assessments. The ANOVA revealed no significant differences in mean LSI risk scores across assessors, $F(3, N = 166) = .74$, $p = .53$. At Time 2, assessments from six parole officers qualified for the analyses. A significant difference was found between assessors on mean LSI scores, $F(5, N = 97) = 2.91$, $p = .02$. With ComCor assessments, four assessors entered into the Time 1 analysis and three assessors entered into the Time 2 analysis. Mean LSI scores varied as a function of the assessor conducting the test at Time

1, $F(3, N = 78) = 6.86$, $p = .00$, and Time 2, $F(2, N = 29) = 12.31$, $p = .00$. In each case with significant differences, one assessor tended to rate offenders higher than two other assessors.

Figure 3. Mean LSI Risk and Rater Scores for Parole Participants

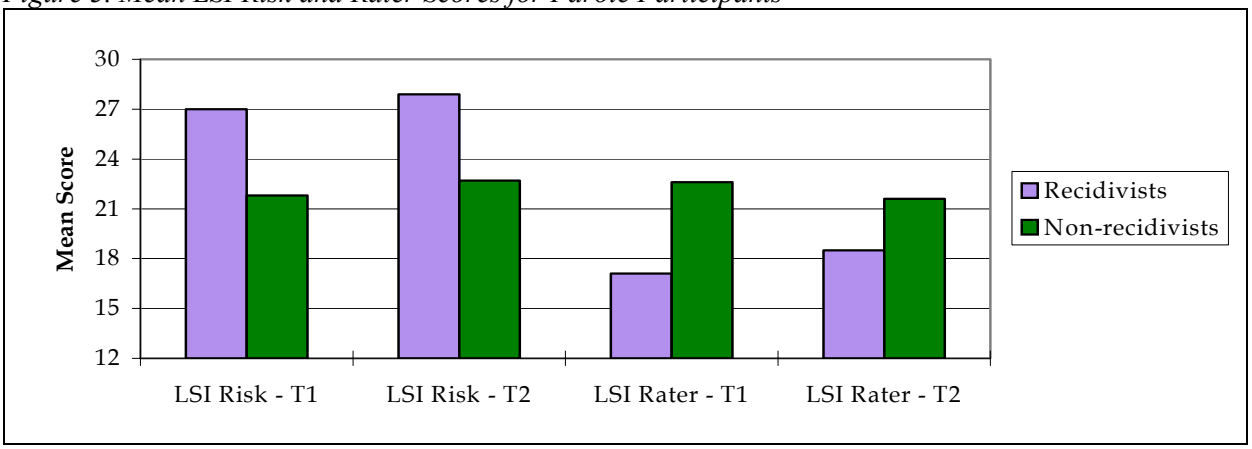


Figure 4. Mean LSI Risk and Rater Scores for ComCor Participants

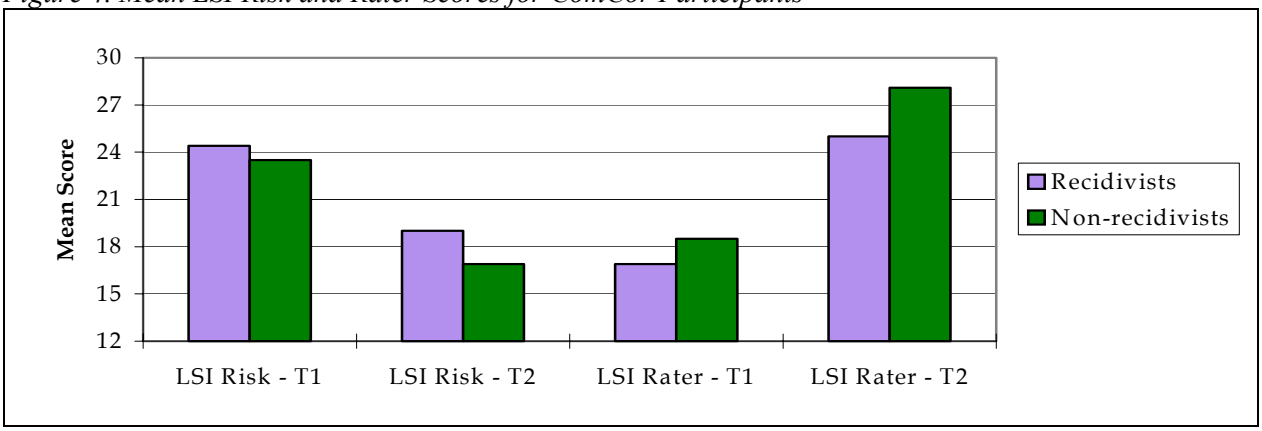


Figure 5. Mean Wisconsin Risk and Needs Scores for Parole Participants

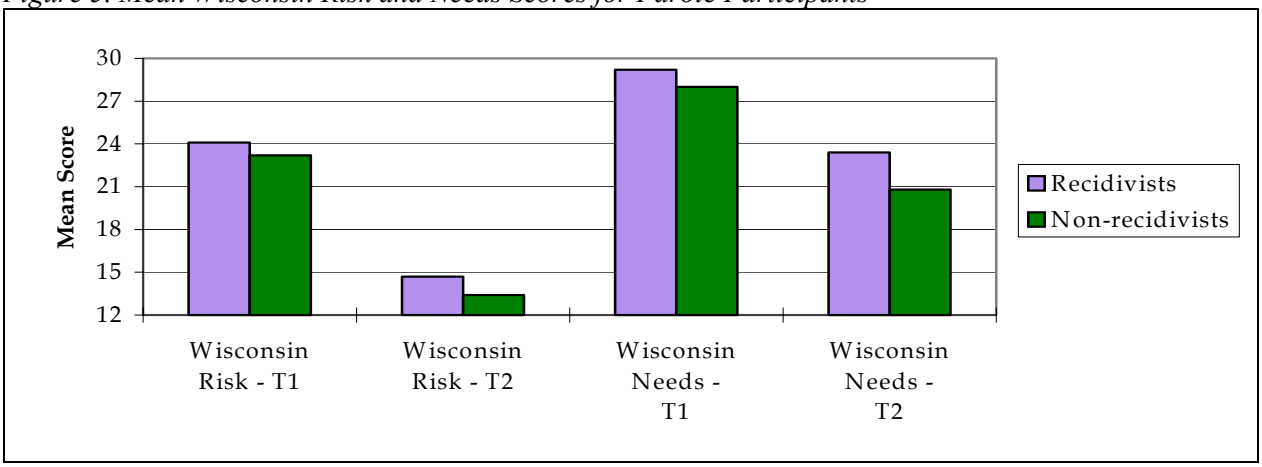


Figure 6. LSI Scores by Assessor at Parole

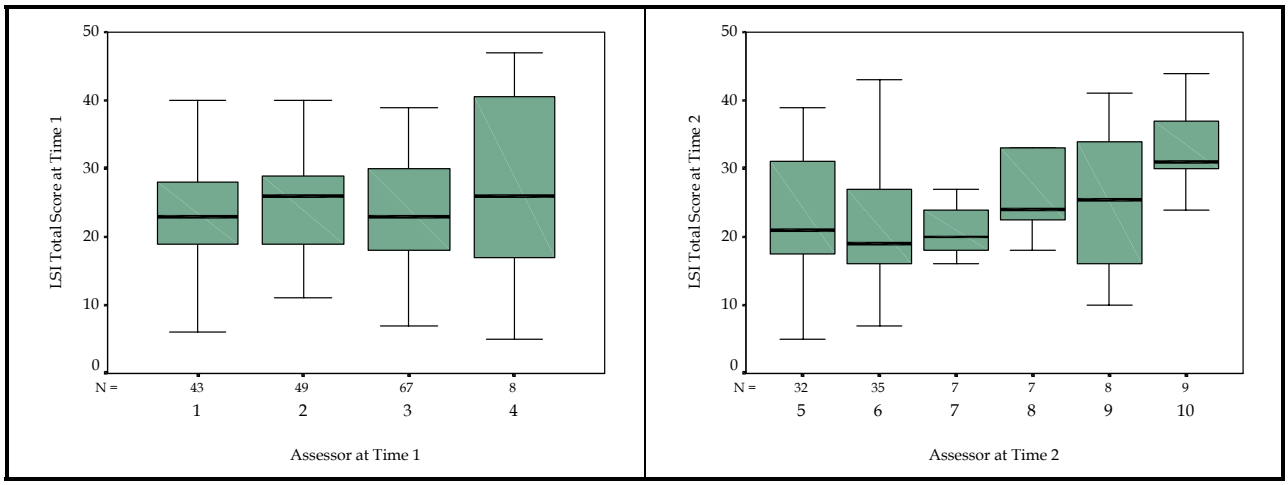
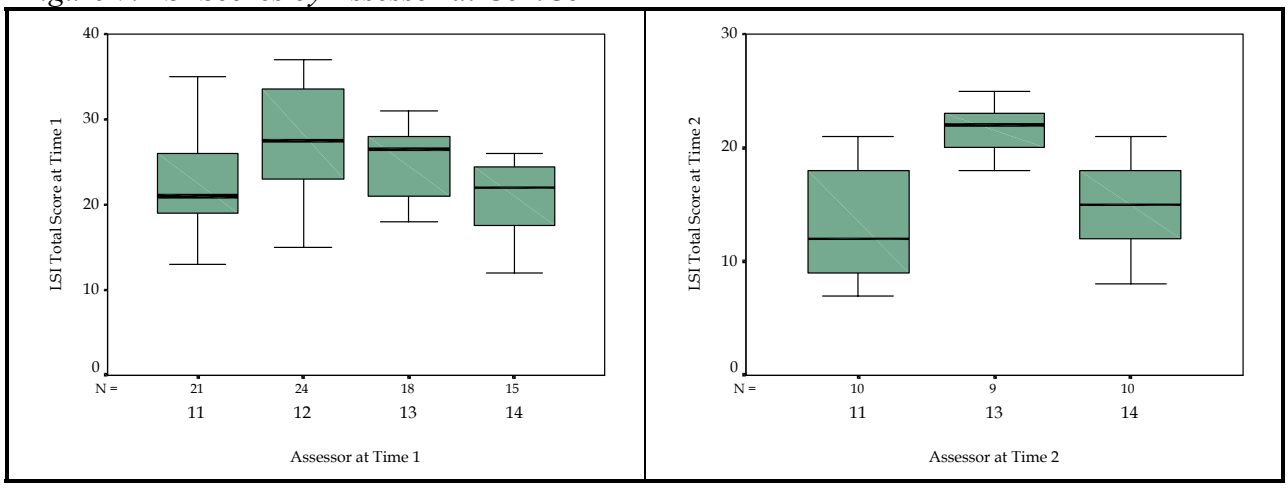


Figure 7. LSI Scores by Assessor at ComCor



Classification Levels

The LSI has not yet been normed for an American population; however, Canadian norms do exist. For the purposes of this study, LSI risk scores were classified according to Canadian norms for parolees (see Figure 8) and ComCor offenders (see Figure 9). Supervision classification levels for parolees, based on the Wisconsin, are presented in Figure 10. Wisconsin risk levels were based on actual offender scores rather than officer overrides.

Figure 8. Supervision Levels per LSI for Parolees

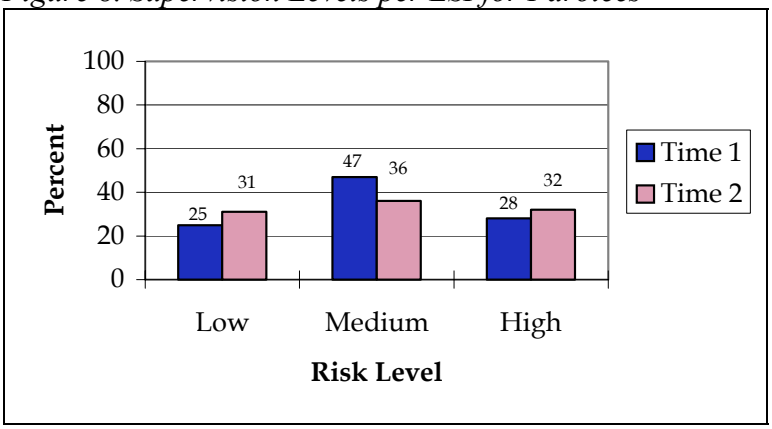


Figure 9. Supervision Levels per LSI for ComCor Residents

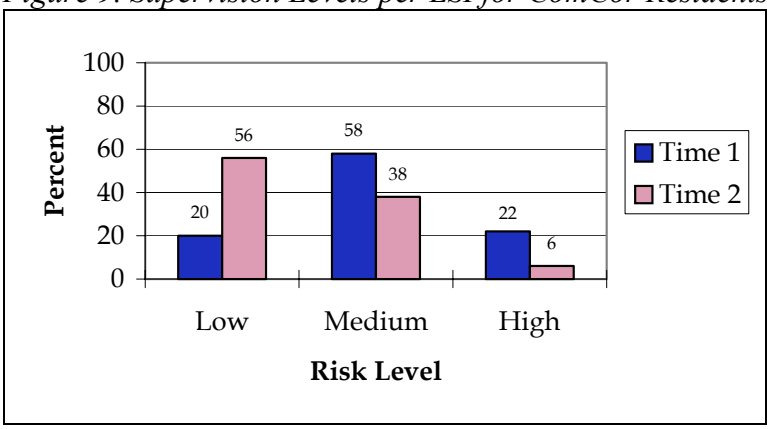
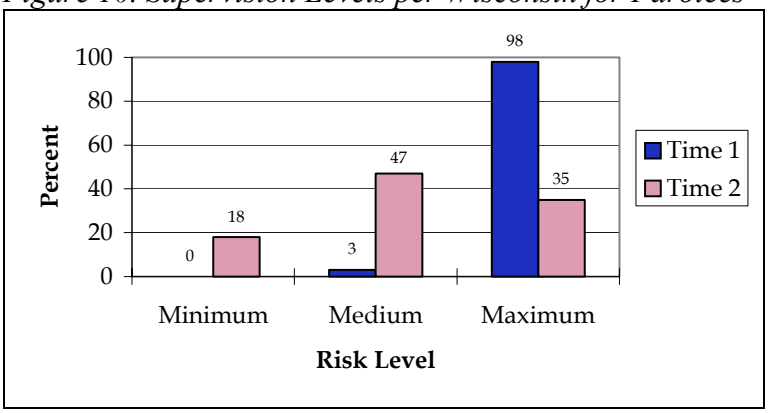


Figure 10. Supervision Levels per Wisconsin for Parolees



Discussion

The primary function of the present study was to evaluate the predictive validity of the LSI with community based offenders in Colorado. Two groups were used to denote community based offenders: parolees supervised in the southeast parole region and transitional community corrections offenders at ComCor. These populations did not appear to differ in significant ways from each other, thereby supporting the belief that these groups represent the same population.

Although this study strove to obtain a representative sampling of community based offenders, some offenders were excluded for lack of assessment. In general, it appeared that excluded offenders consisted of the most severe offenders who absconded or were arrested shortly after release and the most successful offenders who transferred or successfully discharged almost immediately from the supervising agency. Therefore, these research findings may not generalize to highly successful or highly unsuccessful offenders. Additionally, three offenders who presented well on the LSI but actually reoffended after release were excluded. Consequently, these findings may not apply to offenders with similar profiles (e.g., psychopaths).

In estimating the predictive validity of the LSI, several interesting trends emerged. First, the LSI was more predictive of outcome with parolees than community corrections offenders. Second, Time 2 LSI reassessments generally did not improve prediction of 12 month outcomes over the initial assessments. Third, parolees' initial LSI rater score was a strong risk predictor; however, this finding was not consistent across both groups or administration periods. Fourth, the LSI was more predictive of recidivism than the Wisconsin for parole offenders. Finally, declines in parolees' Wisconsin scores over a 6 month period were associated with a slight improvement in risk prediction.

The validity results of the LSI with ComCor offenders were contrary to the study hypotheses. In fact, the findings were particularly curious because the two groups were similar and the LSI was valid with parolees. Different test versions between agencies might account for the discordant validity findings. Moreover, exploratory analyses indicated that scores of ComCor assessments varied as a function of the individual conducting the assessment. Yet, there was no evidence to suggest that these fluctuations should relate to case managers' caseloads. In fact, a ComCor official reported that the agency attempted to neutralize any risk differences in the staff's caseloads. Interestingly, the strength of the LSI validity estimates were negatively related to the amount of scoring variability between assessors. Unfortunately, it was impossible to determine the impact of the test versions on raters' scoring.

Also contrary to the study hypotheses was the finding that reassessments were no more predictive of outcome than initial assessments. Although there was no significant change in parolees' risk scores across assessment periods, the initial assessment was a stronger recidivism predictor than the reassessment. At ComCor, there was a significant change in scores over time, but those changes did not produce significantly improved validity estimates. Unfortunately, this study was not able to support the dynamic nature of the LSI. On the other hand, neither could it dispute its dynamic attributes because scoring variations among assessors may have instead accounted for the low validity of reassessments.

Parolee assessments at Time 1 resulted in high validity estimates for the LSI rater score. However, this high correlation did not hold constant across groups or administration periods. In fact, the predictability of the LSI rater score seemed to decline as scoring variability increased. This finding makes inherent sense because the rater score is even more susceptible to rater bias than the risk score. Thus, the LSI rater score may have potential as a strong risk predictor, but appears less robust than the risk score.

The LSI was more predictive of recidivism than the Wisconsin. While the Wisconsin reassessment displayed some improved validity over the initial assessment, the LSI remained the stronger predictor of risk at both assessment periods. A breakdown of initial classification levels based on the Wisconsin indicated that 98% of offenders scored in the maximum supervision range. This would suggest that the Wisconsin system is strongly influenced by raters and may lead to over-classification.

Implications

The LSI was found to be the best risk predictor for parolees when compared to the Wisconsin. These findings replicate research conducted with Canadian populations. Although the LSI was not found to be predictive with halfway house offenders in this study, its use with this population should not be discounted. Exploratory analyses revealed that the low predictive power may have resulted from individuals' assessment styles.

The LSI is not entirely objective. Truly, its strengths seem to lie in quantifying subjective material. Neither subjective clinical decisions nor objective, static ratings have led to more accurate predictions than the LSI. Yet, this poses a problem when the instrument is administered by individuals lacking assessment or clinical training. The most predictive assessments in this study were conducted by TASC case managers, all who were seeking an advanced degree in a counseling field. Hence, it is probable that previous clinical training impacts the validity of the assessments.

These findings taken together highlight the need for rigorous quality assurance and training. It is not the purpose of this research to suggest that only individuals with advanced counseling training be the sole assessors. However, the criminal justice system must consider the resources required for individuals with less assessment training to conduct the LSI. A thorough training and quality control process needs to be implemented across all jurisdictions. Even for individuals with advanced training, it is recommended that continuous quality control be provided as numerous scoring errors were evident even with TASC assessments.

The LSI is purported to be a risk *and* needs assessment. This study only evaluated the LSI as a risk tool; however, other research reviewed herein yielded no support for the statistical foundation of the instrument's subscales. In fact, no underlying structure was consistently upheld by research. It may be that the LSI provides a better needs assessment than a largely subjective needs assessment or an actuarial risk assessment. Nonetheless, caution should be exercised when using the LSI as a needs assessment. Without a reliable underlying structure, it would be unwise to develop offender case plans merely on an LSI profile; the interview process, rather than the instrument, is important for developing case plans. Furthermore, the same individuals who find

the instrument difficult to administer are also likely to be unable to interpret it in a meaningful manner.

Recommendations for Future Research

Foremost, these research findings stress the relevance of rigorous training and quality assurance with the LSI. It is a complex instrument, particularly for individuals with limited assessment training. The high frequency of scoring errors is a concern as is the scoring variation among assessors. These concerns raise doubts regarding the reliability of the instrument despite one research study which found high inter-rater reliability (Andrews, 1982). Future research should examine inter-rater reliability through random assignment of offenders to assessors or using different assessors to administer the LSI twice to the same offenders.

This study surfaced an important question: Do different LSI test versions impact its validity? Future research needs to compare the validity of both test versions. Thorough investigation of both versions at other community corrections centers can establish (1) whether the LSI is valid with community corrections offenders and (2) whether there is a difference in validity across test versions.

The present study did not account for all offenders released to the community. To achieve a greater breadth of study participants in future research, assessments should be completed just prior to discharge from prison. Furthermore, the LSI may not predict risk accurately for certain offenders such as psychopaths. A risk instrument designed specifically for psychopaths, such as the Hare Psychopathy Checklist-Revised (Hare, 1991), is recommended for review as a supplemental risk assessment instrument.

The LSI is used as part of the substance abuse evaluation in Colorado. This study was able to confirm the capability of the LSI in assessing the risk dimension of the standardized offender assessment. Yet, this study did not evaluate the utility of the LSI or the other battery instruments in matching identified substance abusers to treatment. Additional research can benefit the field by answering the question of whether offenders are appropriately placed into treatment programs.

The findings herein support the LSI over the Wisconsin in predicting recidivism risk with parolees. Any considerations by DOC to implement the LSI with parolees should include the task of developing norms. It would be unwise to accept norms developed for Canadian offenders, a population likely to differ significantly from Colorado felons. Norms need to be tailored for the population involved and should take into consideration the philosophy of the agency using them.

References

American Psychiatric Association (1980). *Diagnosis and Statistical Manual of Mental Disorders (DSM-III)*. Washington, DC: Author.

American Society of Addiction Medicine (1996). *Patient Placement Criteria for the Treatment of Substance-Related Disorders (Second Edition)*. Annapolis Junction, MD: Author.

Andrews, D. A. (1982). *The Level of Supervision Inventory (LSI): The first follow-up*. Toronto: Ontario Ministry of Correctional Services.

Andrews, D. A. & Bonta, J. L. (1995). *LSI-R: The Level of Service Inventory - Revised*. Toronto: Multi-Health Systems.

Andrews, D. A., & Robinson, D. (1984). *The Level of Supervision Inventory: Second Report*. Toronto: Ontario Ministry of Correctional Services.

Arens, S. A., Durham, B., O'Keefe, M., Klebe, K., & Olene, S. (1996). *Psychometric properties of Colorado substance abuse assessment instruments*. Unpublished manuscript.

Baird, C. S. (1981). Probation and parole classification: The Wisconsin model. *Corrections Today*, *43*, 36-41.

Bogue, B. & Timken, D. (1993). Substance Use History Matrix. (JDF Publication No. 328P). Denver: Colorado Judicial Branch.

Bonta, J. (1993). *Risk-Needs Assessment and Treatment*. Paper prepared for the International Association of Residential and Community Alternatives.

Bonta, J. & Motiuk, L. L. (1985). Utilization of an interview-based classification instrument: A study of correctional halfway houses. *Criminal Justice and Behavior, 12*, 333-352.

Bonta, J. & Motiuk, L. L. (1987). The diversion of incarcerated offenders to correctional halfway houses. *Journal of Research in Crime and Delinquency, 24*, 302-323.

Bonta, J. & Motiuk, L. L. (1990). Classification to correctional halfway houses: A quasi-experimental evaluation. *Criminology, 28*, 497-506.

Bonta, J. & Motiuk, L. L. (1992). Inmate classification. *Journal of Criminal Justice, 20*, 341-351.

Goldberg, L. R. (1965). Diagnosticians versus diagnostic signs: The diagnosis of psychosis versus neurosis from the MMPI. *Psychological Monographs*, 79, No. 602.

Groth-Marnat, G. (1990). *Handbook of Psychological Assessment, Second Edition.* John Wiley & Sons: New York.

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

Kivlahan, D. R., Sher, K. J., & Donovan, D. M. (1989). The Alcohol Dependence Scale: A validation study among inpatient alcoholics. *Journal of Studies on Alcohol, 50*, 170-175.

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.

Meehl, P. E. (1954). *Clinical versus statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.

Meehl, P. E. (1965). Seer over sign: The first good example. *Journal of Experimental Research in Personality*, 1, 27-32.

Millon, T. (1982). *The Millon Clinical Multiaxial Inventory Manual* (2nd ed.). Minneapolis: National Computer Systems.

Nuffield, J. (1982). *Parole Decision Making in Canada*. Ottawa: Solicitor General of Canada.

Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). New York: McGraw-Hill, Inc.

Philbrick, P., Gat, I., & Guisti, L. (1993). *A community corrections study: The LSI as a predictor of program success?* Unpublished manuscript.

Ross, H. E., Gavin, D. R., & Skinner, H. A. (1990). Diagnostic validity of the MAST and the alcohol dependence scale in the assessment of DSM-III alcohol disorders. *Journal of Studies on Alcohol, 51*, 506-513.

Selzer, M. L. (1971). The Michigan Alcoholism Screening Test: The quest for a new diagnostic instrument. *American Journal of Psychiatry, 127*, 1653-1658.

Skinner, H. A. (1982). The Drug Abuse Screening Test. *Addictive Behaviors, 7*, 363-371.

Skinner, H. A. & Allen, B. A. (1982). Alcohol Dependence Syndrome: Measurement and validation. *Journal of Abnormal Psychology, 91*, 199-209.

Skinner, H. A. & Horn, J. L. (1984). *Alcohol Dependence Scale (ADS): User's Guide*. Toronto: Addiction Research Foundation.

Wanberg, K. W. (1994). *The Adult Substance Use Survey - ASUS*. Arvada, CO: Center for Addictions Research and Evaluation.

Wanberg, K. W. (1997). *User's Guide to the Adult Substance Use Survey - ASUS*. Arvada, CO: Center for Addictions Research and Evaluation.

Wanberg, K. W., Horn, J. L., & Foster, F. M. (1977). A differential assessment model of alcoholism: The scales of the Alcohol Use Inventory. *Journal of Studies on Alcohol, 38*, 512-543.

Wiggins, J. S. (1973). *Personality and Prediction: Principles of Personality Assessment*. Reading, MA: Addison-Wesley.