Classifying Risk Development and Predicting Parolee Recidivism with Growth Mixture Models

Andy Hochstetler[1] · David J. Peters[1] · Matt DeLisi[1]

Abstract Using two data sets, containing 582 total cases, this study investigates whether classifying offenders on trajectories of risk scores helps predict parolee recidivism. One data set has 4 years of risk scores and another has three. Both data sets contain control variables measuring released inmates' characteristics. The dependent variable measures arrest or return to prison over a 2-year span. A growth mixture model, classifies offenders into three classes, a stable and high trajectory group, a group with a high but declining risk trajectory, and a small, low-risk group with little change. Trajectory class membership correlates with recidivism in both data sets. Supplementary analyses show that assigned classes are better predictors of recidivism than last risk scores or simple change scores. Discussion centers on the appeal and relevance of trajectories of risk, as opposed to static measures, for predicting offender misconduct and other outcomes.

Keywords Recidivism · Trajectory · Risk assessment · Corrections

There may be no more practical area of criminology than risk assessment of known offenders. It potentially informs policy makers and officials on how to behave rationally and assists in logical choices, such as selective incapacitation, custody-levels, and treatment plans that can be very consequential. Most contemporary corrections systems implement formal risk assessments. There are an array of risk assessment schedules available for general offenders as well as for various types (Singh, Grann, & Fazel, 2011). For example, the Violence Risk Appraisal Guide and the Sex Offender Risk Appraisal Guide have been used in many places and evaluated many times. Generally, these risk assessments have performed well when predicting future offenses and also have been improved and made more efficient over the years (Rice, Harris & Lang, 2013). The Level of Service Inventory –Revised (LSI-R) is the most frequently used

✉ Andy Hochstetler
   hochstet@iastate.edu

[1]   Department of Sociology, Iowa State University, 203 East Hall, Ames, IA 50011-1071, USA

risk assessment for general offenders and it has been repeatedly examined and generally supported in replications; it too has improved over time (Hsu, Caputi, & Byrne 2011). Most implementations of risk scores use a single, static risk measure, although the measurements are of dynamic conditions and often are repeated to update scores and account for changes.

This study extends interest in risk assessment in corrections by focusing on dynamic risk/needs factors in a unique way. Most assessments of offender risk use scores from one point in time, while a very few use two scores to measure change. Rather than using a single time-1 indicator of risk to predict subsequent recidivism or simple change, we examine trajectories of risk. We draw on two data sets. One set is comprised of offenders with three LSI-R scores and the other of those with four scores. The trajectories provide a measure of dynamic risk development. Resulting scores track developmental trajectories, including improvement, and allow placement of offenders in trajectory groups to determine if group membership influences recidivism. We predict recidivism using assigned trajectories under different model specifications and assess, in a supplementary exercise, whether the use of dynamic classifications in assessing offender risk adds information not available in the last scores or in simple change scores reflecting only the first and last data points.

Literature Review

This is not the place to engage philosophical debates about or to explain the reasons for studying actuarial risk and risk assessment in criminology and the justice system (see Mauratto & Moffat, 2006; Silver, 1998). Regardless of how these debates turn, practitioners and social scientists are aimed toward greater use of formal risk assessment scores. Improvements in data quality, storage and retrieval contribute to increasing reliance on formal risk assessments, and improvement in predictive ability is almost inevitable as what is learned is incorporated into development of measures. With better data and replication, measurement properties and prediction should improve.

There are many instruments and checklists designed specifically for subpopulations of offenders including substance users, psychiatric populations and even parolees in a given state. On the whole, empirical evaluation shows that these formal risk assessment devices do much better than chance or professional judgment when it comes to identifying those clients who fail upon release (Andrews, Bonta, & Wormith, 2006; Ægisdóttir et al., 2006; Smith & Cullen, 2013; Smith, Cullen, & Latessa, 2009). As already mentioned, the Level of Service Inventory-Revised (LSI-R) is one of the most widely used assessments for evaluating the risk, needs and responsivity of general offender populations (Andrews & Bonta, 1995; Gendreau, Goggin, & Smith, 2002; Rocque & Plummer-Beale, 2014; Smith & Cullen, 2013). Most items in the LSI-R scale focus on indications of whether offenders are cut-off from conventional structures of work and social life and embedded in criminal behavior patterns and harmful habits. The designers constructed the measure to provide a dynamic score of offender risk and to highlight areas of need. It contains 54 items in 10 subscales indicating criminal history, education and employment, finances, family/marital, accommodations, leisure and recreation, companions, alcohol/drug abuse, emotional/personal, and attitudes/orientation. The dimensions are not equally weighted, with criminal history,

employment/education, and alcohol and drug problems scoring heavier because more items are devoted to these constructs. While some dimensions and items are cumulative, with little possibility of undoing past harm (such as criminal history and certain items on health consequences of drug use), others are more fluid and reversible. The scale's creators distinguish between *risks*, that are cumulative or static dimensions of scores, and *needs* that are reversible dimensions of scores (Andrews & Bonta, 2006). The predictive ability of the scale among those who have been incarcerated lends evidence to the theoretical and empirical importance of stable components, particularly criminal history, but also to the capacity for change among even high-risk groups (Labrecque, Smith, Lovins, & Latessa, 2014; Lussier & Gress, 2014; Vose, Smith, & Cullen, 2013).

Subscales are used to identify needs of offenders and to classify into treatments as well as to provide focus for areas of improvement, but risk assessment for general purposes, such as evaluating outcomes and predicting recidivism, usually is dependent on the total score. Typically, investigators divide offenders into risk categories that are normed to their position in the criminal justice system (i.e., probationer, parolee) and evaluations of classification effectiveness focus on how membership predicts imprisonment or recidivism. Psychometric properties of the LSI-R are sound, but the effect on recidivism varies widely from study to study and across ethnic and gender groups (Reisig, Holtfretter, & Morash, 2006; Schlager & Simourd, 2007; Simourd, 2006). Investigators have tested it among many types of offenders and in varied cultural contexts (Folsom & Atkinson, 2007; Hsu et al., 2011; Manchak, Skeem, Douglass, & Siransosian, 2009; Zhang & Liu, 2014; Simourd, 2004; Smith & Cullen, 2013; Wilson & Guiterrez, 2014), and findings on the whole are supportive. The measure ranges in predictive ability from explaining anywhere from a relatively modest 20 % of the variation in recidivism to a robust 40 % (Andrews et al., 2011). In evidence of its consistent strength, a 2002 meta-analysis of 30 studies showed that scores were a powerful predictor of both general and violent recidivism (Gendreau et al., 2002). Smith, Cullen, and Latessa's (2009) review of 27 reports on the predictive validity of the LSI-R among female offenders provided further support.

The LSI-R is based on a social learning model of crime. Indeed, its developers emphasized early on that their approach was developmental and that the measure should be administered repeatedly to capture change. This builds on the theoretical contention that criminality is learned and that criminal behaviors occur when crime leads to perceptually positive outcomes for the offender; when offenders are positioned to receive relative rewards for continuing with crime as opposed to turning to conventional pursuits, it is entrenched. As an offender's lifestyle, attitudes, and rationalizations move toward those more favorable to committing crime, the likelihood of continuing crime and related pursuits increases. By contrast, indicators of conventional pursuits should dynamically shift offenders away from criminal thinking and risk (Andrews & Bonta, 2003; Blanchette & Brown, 2006).

While most recent tests of the LSI-R's performance use static LSI-R scores to predict misconduct or eventual recidivism, there is scattered evidence that the LSI-R captures some of the dynamism in offenders' lives. For example, change in risk scores is an important predictor of recidivism (Tillyer & Vose, 2011; Vose, Lowenkamp, Smith, & Cullen, 2009; Vose et al., 2013). In one study of 2849 offenders who had taken the LSI-R at least twice, a 10 % drop in score for high risk offenders led to a 6 % reduction in

recidivism. However, a 10 % drop in recidivism for low risk offenders resulted in only a 1 % drop in recidivism. The effect of a 10 % drop was most pronounced for probationers as opposed to parolees (Vose, 2008; Vose et al., 2013). Conclusions were that programs should be targeted to reducing risk of high risk offenders and that the LSI-R should be administered multiple times as a way to monitor the rehabilitative process (Vose et al., 2013). If an offender's score declines, a lower level of supervision or discharge may be merited (Prell, 2009; Vose, 2008). The policy implications of these conclusions about change are significant. Indeed, one of the attractions of this measure is that, while incorporating the empirical reality that the harms of a sustained criminal career are cumulative and the, often unacknowledged, likelihood that part of criminal propensity may be stable, it also measures significant intra-individual changes that may be treatable. However, to date, examination of change in risk scores and the implications of change for predicting crime has been rudimentary at best, typically relying on the last score of record or simple difference scores between the first and last data points. There is almost no methodological attention to the possibility that offending trajectories may be shared among like offenders or to the exogenous predictors of change. Reliance on simple change scores also may obscure multiple trajectories of change with unique effects on recidivism.

Sample and Methods

This study draws on data from a Midwestern state. The data were collected for administrative purposes and provided by the state Department of Correction (DOC). The database includes all persons having contact with the corrections system in the state, linking together records from the courts, prisons, probation and parole. This corpus includes conviction information, risk assessments, treatment participation, psychological assessments, employment and various demographic measures. Here, the primary independent variables are based on a risk assessment instrument, the Level of Service Inventory-Revised. All who enter the Department of Correction in the state are required to be scored and to be reassessed as they proceed through their correctional program at regular intervals. There are also unscheduled assessments brought on by new events or on future contacts with the DOC. The DOC administers the LSI-R to everyone entering the corrections system upon classification, although many offenders already have taken the assessment on probation; scores are updated at least annually while on parole. The score is used in treatment assignment and other case decisions throughout the DOC, which includes probation and parole in the state.

For the current analysis, we selected all offenders in the database paroled in calendar year 2010. We excluded offenders who expired their sentence (no longer under supervision on release), those sent to other prisons and institutions upon release (such as federal custody or civil institutions, those paroled to other states (not under the state's supervision that provided data) and those paroled for operating under the influence sentences (specialized and shorter sentences). Of the offenders ($n = 3071$) released from state prisons in 2010, 1199 expired their sentence and 602 were not under state supervision. From the 1270 remaining parolees who were released and resided in the state and who had the potential, therefore, to commit recorded offenses in the community, we selected those who had either 3 or 4 scores of the LSI-R on record. These

groups, containing 226 released offenders with four waves and 356 offenders with three waves, are compared to the 1270 frame, the number of offenders with generally complete data in Table 1.

Statistical comparisons and tests of significance in Table 1 are of offenders in the group of interest to all others, so that those with three LSI-Rs are tested against those with four or any other number and those with four LSI-Rs are compared to those with 3 or any other number. As Table 1 shows, the smaller four wave sample differed significantly from others (release class) in two respects of the nine characteristics compared. Sample members served longer sentences and were more likely to have committed a violent crime in their last offense, explaining why they had multiple risk scores on record. Those sampled were not statistically different on age, sex, marital status, recidivism rate, education, or initial LSI-R score. The three wave group differed only on recidivism - one of the nine variables compared. In sum, while both samples are in several respects representative of the larger population of parolees released in 2010 and residing in the state, those who have taken the LSI-R multiple times are significantly different from the general population on a few variables. While we would not suggest that differences are ignorable entirely, each group average looks a great deal like the entire release class average on most variables. Nevertheless, a cautionary interpretation of findings may be warranted, as offenders who take very few LSI-Rs almost by definition have served shorter sentences or had fewer contacts with corrections than those who take many.

While much of the analysis consists of assigning offenders to classes based on LSI-R scores, the dependent variable and test of the method is recidivism within 2 years of release date. Recidivism is defined as either being arrested or returned to prison within a 2-year window (0 = no recidivism; 1 = recidivism). The database links with court records to identify when a parolee is arrested and also links with prison records to identify when a parolee is returned to prison. Most arrests eventually lead back to prison, so many recidivists have two instances of failure. A return to prison in the absence of arrest typically indicates a parole violation; and

Table 1 Descriptive statistics for two LSI samples, and non-sampled cases from parole class frame with tests for differences

	Mean 4 LSIs	Mean 3 LSIs	Mean Non-sample	Mean Entire Class
Recidivism	41 %	43 %*	36 %	39 %
Age	37	37	38	38
Race	22 %	25 %	27 %	26 %
Sex (1 = male)	84 %	87 %	85 %	85 %
Marital Single	81 %	85 %	84 %	83 %
Education level	4.2	4.2	4.2	4.2
Violent Crime last	23 %**	19 %	16 %	18 %
Months served	31***	23	23	25
1stLSI	32	32	32	32
	$n = 356$	$n = 688$	$n = 1270$	$n = 226$

*Chi-square or *t*-test comparing sample to all others in the frame

the decision to send parolees back to prison is made by parole officers and the court. This definition of recidivism includes offenders who are arrested for minor as well as serious infractions and it also includes offenders whose parole violations represent the full range of severity. We erred on the side of an inclusive and sensitive definition of recidivism because of the need for a measure of failure without a very large sample, and because practically speaking any arrest or imprisonment represents a violation of the conditions of parole and is important to the public and officials. Using this definition of recidivism, about two-fifths of parolees (38.6 %) were rearrested or readmitted to prison within 2-years. The recidivism rate for those in the four-wave sample was 41 % and it was 36 % for those not sampled. In the three wave sample, it was slightly higher (43 %) than in the four wave data. Table 1 presents descriptive statistics for variables included in the analysis. It also presents descriptive information for all offenders released in 2010 meeting all criteria for inclusion in the study ($n = 1270$) except for the requirement to have 3–4 LSI-R scores on record in the non-sample column. Tests of statistically significant differences are of those with the designated number of LSI-R (either 3 or 4) scores to the remainder of the sample.

Table 2 presents the mean development of LSI-R scores over time for the three and four wave sample. It shows that the scores are high on the first measure, increase very slightly on the second to peak at 32, and then drop slightly on the third and fourth administration with a final mean score of 29.56 in the four LSI-R data. In other words, the scores drop 8.6 % from the peak at the second scoring on average in the four wave data.

For the three wave data, the drop from first to last score represents a 6 % decline. Overall, the average LSI-R scores are similar and declining with similar slope in both samples. The magnitude of the average decline in the entire sample is not large. The declines in both sets may be due to the fact that offender risk decreases when offenders are institutionalized as there are limited opportunities for crime and drug use, and there are informal incentives for administrators to show offender improvement before releasing to parole, and because some measures in the score, such as work difficulties or criminal friends, are difficult to assess in the prison context. However, findings presented in the following analysis show that trajectory differences are substantively meaningful, lending predictive validation to the observed score changes.

Table 2 Mean LSI-R scores in two samples		Mean 4 LSIs	Mean 3 LSIs
	1.	32.33	31.88
		(7.59)	(7.43)
	2.	32.36	31.53
		(7.85)	(8.05)
	3.	29.87	30.12
		(7.61)	(8.20)
	4.	29.56	
		(8.05)	

We model the development of risk scores over time, to determine if there is a single or multiple developmental trajectories identified in the data and then proceeded to examine whether they are meaningful as predictors. The LSI-R total score, feasibly ranging from 1 to 54, is used as the indicator of offender class in the model. Four waves are used to form a latent class variable. This variable represents a single measure reflecting how these scores cluster together.

Other independent control variables used to better predict offender class and to determine their influence on recidivism include: race (0 = nonwhite; 1 = white), sex (female = 0; male = 1), marriage status (1 = single and 0 = other) , violent criminal offender in the last offense (0 = other; 1 = violent crime),age in years, number of months served in the last sentence, and the highest education level achieved (1 = grade 1–5; 2 = 6–9; 3 = 10–12; 4 = HSD or GED; 5 = some coll.; 6 = post high-school vocational or Associate's Degree; 7 = Bachelor's or higher).

The general methodological goal of this analysis is to explore whether data driven classes can be identified from trajectories of LSI-R scores over time and whether those classes then can be used to predict recidivism. The strategy entails several steps. First, we examine a growth mixture model (GMM) comprised of multiple LSI-R scores over time with recidivism as a categorical distal outcome. Figure 1 presents the model graphically.

Growth Mixture Models

The objective of this step is to measure information about inter-individual differences in intra-individual change in scores over time. GMM relaxes the assumption that all cases form a single population with common parameters. It also does not assume that covariates affect growth factors across types of individuals in the same way; rather it acknowledges that heterogeneity of growth trajectories exist (Jung & Wickrama, 2008). The technique is appropriate for capturing information about inter-individual differences in intra-individual change, taking into account unobserved heterogeneity, or different groups, within a larger population (Jung & Wickrama, 2008).

The model forms latent trajectory classes in the form of categorical variables that allow for different groups of individual growth trajectories to vary around different means, resulting in separate growth models for each latent class (Asparouhov & Muthén, 2008). Mixture modeling refers to modeling with categorical latent variables that represent subpopulations where population membership is not known a priori but is
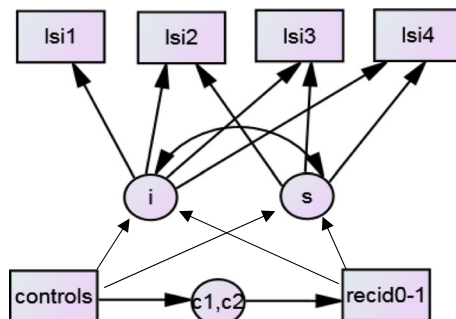
Fig. 1 Growth mixture model with LSI class and 2-year recidivism distal outcome

inferred from the data (Asparouhov & Muthén, 2008). The growth is modeled as linear. Classes are categorical latent variables. In growth mixture modeling, within class variation of individuals is allowed for identified latent trajectory classes.

In the absence of guiding theory, investigators specify the number of classes and test for the best model before selecting the number of classes, using the reduced model without controls or a distal outcome. The classes are formed by the intercept and slope "clusters." Addition of control variables acknowledges and controls for exogenous influences on class assignment and often improves the ability of the model to assign and distinguish class membership as reflected in the entropy statistic. Investigators regress distal outcome (recidivism) on the categorical latent variable using logistic regression to see whether class membership, which is assigned by the model, affects it.

One problem with this approach is that classification of offenders into "clusters" may be fuzzy or imprecise. Cases are fit to classes based on probability and entropy may be low. However, this is not necessarily devastating to the model as assigned class may still be a significant predictor of the outcome, and this lends evidence validating classes by suggesting that the assigned classes are significant predictors despite the imprecision of case assignment. Also, lower entropy can produce good parameter estimates.

In a supplementary analysis, we use assigned classes and probability of class assignment to test whether class adds information that is not available in the last LSI-R score. We do so by exporting data for simple logistic regression of recidivism on the covariates. This allows us to regress recidivism on the exported class assignment. The purpose of this is to show whether the assigned class provides predictive ability above and beyond what is provided by the last LSI-R score, which is the conventional way of assigning offenders' risk levels. Comparing results allows us to assess whether the latent class assignments perform comparably to the last LSI-R score of record, a question of some policy, theoretical, and practical significance. We also examine whether change in LSI-R performs better than assigned class with a logistic regression of recidivism on the last LSI-R score and the first (lagged to wave 1) LSI-R score.

One means of examining whether confidence in assignment affects results is to export the probability of membership for each case and use the assigned probabilities in subsequent analysis to determine how sensitive results are to imprecision in assignment. Having examined whether assigned class predicts recidivism as well as the last LSI-R score or change between the first and last scores, we proceed to examine how sensitive these results are to confidence in class assignment, by repeating this approach with only high confidence of assignment cases. Using the exported probabilities of class assignment from the growth mixture model, we repeat the two logistic regression analyses. Results for this model are based on those cases where we can be confident of class assignment (probability > .80). Recall that all steps in the analysis, including those for high confidence cases, were repeated on two separate sets of data, a set containing offenders who had exactly four LSI-R socres on record and a set that had exactly three LSI-R scores on record.

Results

Determining the number of classes depends on a combination of factors in addition to fit indices, including "one's research question, parsimony, theoretical justification, and

interpretability" (Jung & Wickrama, 2008). We first examined a Latent Class Growth model with a single class with no within class variance. Fixing the within class variance to zero for exploratory purposes leads to clear identification of classes. The results showed that the three-class solution fit better than a one or two-class solution and examination of plots revealed no reason to model class-specific variances for the classes (1 class Sample Size Adjusted Bayesian Information Criterion = 6352.464, 2 class = 6125.607, 3 class = 6057.744). In the three wave analysis, the same was found (1 class Sample Size Adjusted Bayesian Information Criterion = 6352.464, 2 class = 6125.607, 3 class = 6057.744). Results from this exercise indicate that multiple classes were present in the data; the three class solution fit best and also revealed a consistently low LSI-R class that prior research on the score would indicate is likely to be substantively important.

In fitting the growth mixture model, we also compared one, two, and three class solutions. From the single class model in the four LSI-R data, the most relevant statistic is that the mean intercept for the entire sample's four LSI-R scores is 32.245 with a declining slope of −.976. This demonstrates that on the whole prisoner and parolee risk is high and remains so even though there is a decline over the four LSI-R records. Figures 2 and 3 present the estimated single class trajectories.

The GMM model confirmed the decision that three-groups offered the best fit, for the 4 LSI-R data. We used a number of indices to reach this conclusion including BIC (Bayesian Information Criterion), sample size adjusted BIC, and the Boot strap likelihood ratio test (Sample Size adjusted BIC for 3 class solution = 7239.505 entropy = .802; Sample size adjusted Bayesian Information Criteria for 2 class solution = 7256.004 entropy = .643; Sample size adjusted BIC for the 1 class solution = 7277.266). The posterior probabilities for class one were 82 % and 89 % for class 2, and 96 % for class 3.

In the three wave data, results were similar (Sample Size adjusted BIC for 3 class solution = 8261.785 8947.752 entropy = .742; Sample size adjusted Bayesian Information Criteria for 2 class solution = 8956.162.7410 entropy = .614; Sample size adjusted BIC for the 1 class solution = 8982.667). In the three class model, posterior probabilities for class membership indicate that for Class 1 there was an 89 %
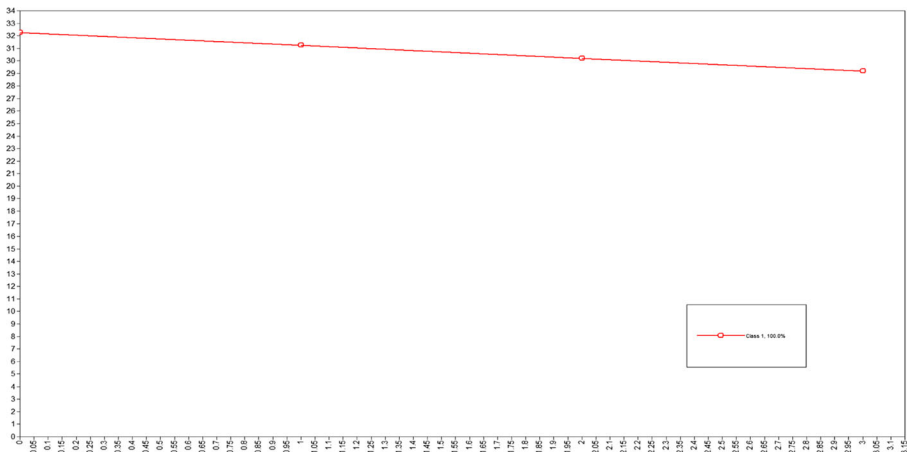

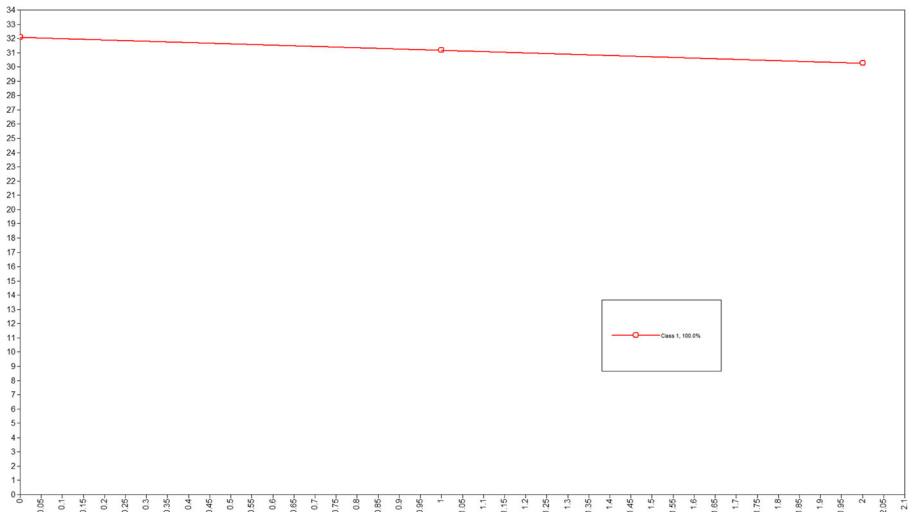Fig. 2 Estimated means for 4 LSI sample

Fig. 3 Estimated means for 3 LSI Sample

probability of membership, a 90 % probability for Class 2, and an 87 % probability for Class 3. Graphical analysis revealed that in both data sets there was a small group of offenders with consistently low LSI-R scores present in the three group solution that were not identified in the two group solution. The three class model is superior. The program for class assignment exported probabilities from a three-class model containing the full set of exogenous, independent variables.

Figures 4 and 5 present the estimated slopes and intercepts for the three classes in each data set. For the 4 wave data, Class 1 has an estimated intercept of 29.913, a relatively low LSI-R score. The slope for Class 1 exhibits a decrease (−.441). Class 2 exhibits a much higher intercept (42.951) and a more steeply decreasing slope of LSI-R scores over time (−7.845). Class 3 also has a high LSI-R intercept 38.058 but exhibits
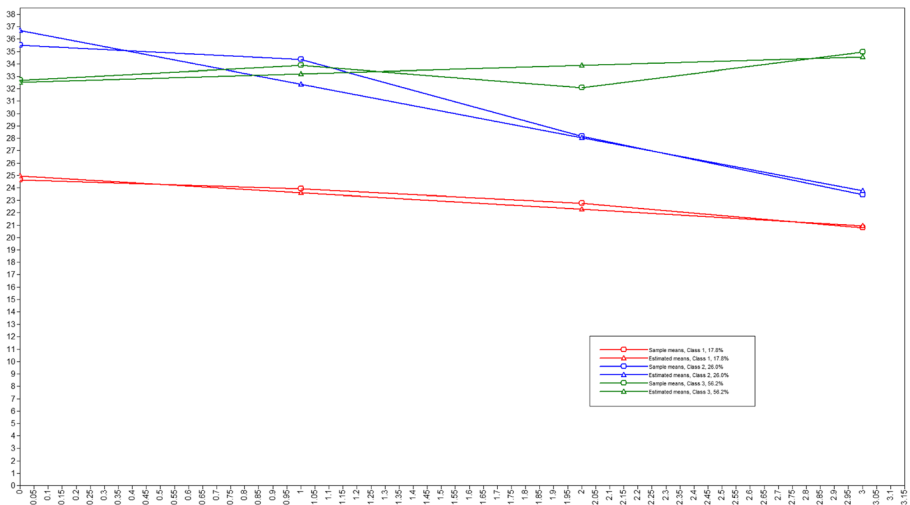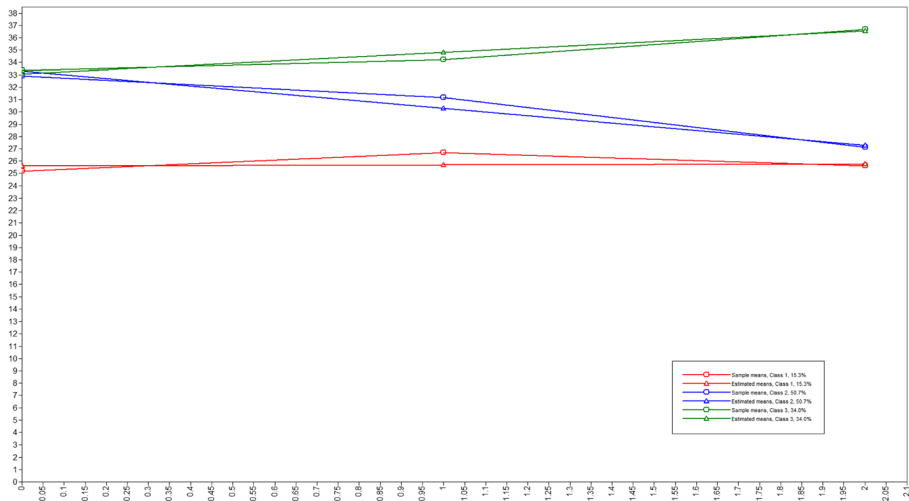

Fig. 4 Sample and estimated means for 3 class solution, 4 LSI sample

Fig. 5 Sample and estimated means for 3 class solution, 3 LSI sample

little decline over time in slope (−2.477). Of the total number of offenders ($n = 226$), the model assigned 126 (56 %) to class 3, 58 (26 %) to class 2, and 42 (19 %) to class 1.

The three LSI-R score data reveals a similar pattern. Class 1 has a low LSI-R score intercept (28.23) and a very small slope (.0621). Class 2 has an intercept of 43.082 and steeply declining slope (−5.96). Class 3 has a high intercept 43.342 and a slightly declining slope (−.5112). Of the 356 offenders, 54 (15 %) are in the class 1, low LSI-R group. There are 116(33 %) in the consistently high class 3 group, and 186 (50 %) in the class 3 high but declining risk group.

Table 3 presents results from the growth mixture model with the distal outcome recidivism, with focus on exogenous predictors of class. The entropy statistic for classification quality is better as it approaches 1.00 and while there are no firm

Table 3 Descriptive statistics and tests for differences for three classes with Class 1 as the reference group

	4 LSIS ($n = 226$)			3 LSIs ($n = 356$)		
	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
Recidivism	.00	.31***	.61***	.09	.29***	.90***
Age	42.46	36.14	34.33***	40.05	38.23	34.64***
Race	0.0	.31***	.25***	.24	.18**	.36
Sex (1 = male)	.71	.88***	.85*	.37	1.00***	.91
Marital single	.83	.79	.80	.80	.85**	.85
Education level	4.12	4.41	4.16	4.33	4.19	4.15
Violent crime last	.24	.36**	.16	.17	.19	.18
Months served	29.88	32.12	30.90	24.69	21.52	19.44
First LSI	23.69	36.14*	32.83	25.11	32.97***	33.29***
N=	42	58	126	54	186	116

**p < .05, ***p < .005

conventions for cut-off, there is agreement that caution is warranted for many applications and exporting classes for further analysis when entropy is below .80; .80 represents high entropy and a high degree of correct classification. The four LSI-R data meet the .80 cut-off, but the three LSI-R data have an entropy of .74. However, the assignment probabilities for each class exceed .82 in all classes in both data sets. Here, membership significantly predicted recidivism, however, and this evidences that assignment is meaningful even when entropy is not high.

Table 4. Presents results of the GMM with the recidivism distal outcome. In the four LSI-R data, none of the covariates significantly affected the intercept or slope. Age also was the only covariate that was significantly related to the latent class in the logistic regression portion of the model when controlling for other variables with Class 1 coded as the in-category outcome (b = .093, p = .025), and with Class 2 as the in category outcome (b = .057, p = .012). The most important effect in the study is that of class 2 vs. class 3 on recidivism, the distal outcome. Recall that class 1 represents the low and slightly declining LSI-R group. In this group, none of the 42 offenders recidivated. In latent class 2, the high and declining group, .31 of the 58 offenders recidivated. In class 3, the consistently high group, .609 of the 126 offenders recidivated. Those in class 2 were significantly (74 %) less likely to reoffend as those in class 3, despite a similar wave 1 LSI-R (p = .030). Odds cannot be compared to class 1 as none recidivated in this group. There is a large difference between the other two classes on recidivism despite the fact that they started their LSI-R trajectories with similar initial average scores.

In the larger three LSI-R data, there were 356 offenders and 152 recidivists. In this data set, being male and minority affected both the slope and intercept. Being male decreased the LSI-R intercept (b = −11.363, p = .000), whereas being minority increased it (b = 3.235, p = .003). The higher the intercept, the steeper the slope in this set (correlation = 8.178, p = .027). Being male and minority predicted class membership (class 1 in-category, male = b = −2.76, p = .000; class 2 as in-category minority = b = −.894, p = .047). Again, the most important effect in this study is that class category predicts recidivism. Class 1 represents the low and slightly declining LSI-R group, and had a 9 % recidivism rate. Class 2 is high and declining; it had a 29 % recidivism rage., Class 3 begins high with a relatively steeply declining slope, and it had a 90 % recidivism rate. There is no significant difference between Class 1 and 2 or Class 1 and 3 identified in the whole model, but this is likely due to the small numbers in class 1; members of class 1 are 97 % less likely to reoffend than Class 3 and 63 % less likely to reoffend than class 2. The most noteworthy finding is that there is a significant (p = .054) difference between class 2 and class 3 members with those who start high and with a significantly declining slopes having odds of recidivism 93 % less likely than those who remain high.

In sum, we find that where the distal outcome is recidivism, being in the consistently high group predicted higher recidivism than being in the high intercept but declining category. Both high intercept classes had higher recidivism than the consistently low group. The distal outcome validates the classes.

In supplementary analyses, we compare results using raw LSI-R scores to GMM assigned classes as a simple test of whether assigned classes add information that is not contained in the last LSI-R data point. We compare results using raw LSI-R scores to assigned class as a simple test of whether assigned classes add information that is not contained in the last LSI-R data point. Then, we compare assigned class to LSI-R change

Table 4 Growth Mixture Model with Distal Outcome Recidivism by Class

4 LSIs	Estimate	S.E.	p value	3 LSIs	Estimate	S.E.	p value
Class 1 (n = 42, 19 %) 0 % recidivism				Class 1 (n = 54, 15 %) 9 % recidivism			
Intercept	*29.913*	*6.30*	*.000*	*Intercept*	28.23	2.90	.000
Slope	*−0.44*	*0.715*	*.541*	*Slope*	0.06	1.68	.970
Class 2 (n = 58, 26 %) 31 % recidivism				Class 2 (n = 186, 52 %) 29 % recidivism			
Intercept	*42.95*	*5.266*	*.000*	*Intercept*	43.08	3.08	.000
Slope	*−7.85*	*1.62*	*.000*	*Slope*	−5.96	1.67	.000
Class 3 (n = 126, 58 %) 61 % recidivism				Class 3 (n = 116, 33 %) 90 % recidivism			
Intercept	*38.06*	*4.27*	*.000*	*Intercept*	40.98	2.90	.000
Slope	*−2.48*	*1.29*	*.055*	*Slope*	−0.51	1.70	.093
Latent Variable (Class#1) on				Latent Variable (Class#1) on			
Age	0.09	.04	.03	Age	0.05	.04	.220
Race (1 = nw)	−0.55	.64	.39	Race(1 = nw)	−0.13	1.02	.898
Sex (1 = male)	−21.22	.00	.99	Sex (1 = male)	−2.76	.42	.000
Marital Single	0.842	.66	.20	Marital Single	−0.26	1.34	.848
Education level	−0.36	.41	.382	Education level	0.03	.61	.959
Violent crime last	0.60	1.25	.633	Violent crime last	−0.29	1.12	.792
Months served	0.00	.01	.909	Months served	0.02	.02	.377
Intercept	−3.37	2.16	.119	Intercept	−1.06	2.09	.610
Latent Variable (Class#2) on				Latent Variable (Class#2) on			
Age	0.06	.02	.01	Age	0.026	.02	.220
Race (1 = nw)	0.52	.68	.45	Race (1 = nw)	−0.89	.45	.047
Sex (1 = male)	0.41	.55	.47	Sex (1 = male)	11.62	6.36	.068
Marital Single	0.23	.52	.65	Marital Single	0.04	.56	.946
Education level	0.25	.30	.42	Education level	−0.00	.21	.994
Violent crime last off.	1.12	.58	.05	Violent crime last off.	0.24	.54	.656
Months served	−0.01	.01	.27	Months served	−0.02	.01	.656
Intercept	−4.73	1.97	.016	Intercept	−11.54	6.66	.083
Odds ratio for Recidivism outcome				Odds ratio for Recidivism outcome			
Class 2–3	*.26*	*0.12*	*.030*	*Class 1–2*	.37	.36	.308
				Class 1–3	.03	.03	.355
				Class 2–3	.07	.03	.054

Model fit 4 LSI: BIC = 7397.97, Sample Size Adjusted BIC = 7239.51, Likelihood ratio Chi Square for binary outcomes 23.32, df = 22 p = .38, entropy = .80, latent class probabilities 1 = .82, 2 = .89, 3 = .96

Model fit 3 LSI: BIC = 9103.20, Sample Size Adjusted BIC = 8947.75, Likelihood ratio Chi Square for binary outcomes 26.39, df = 22 p = .24, entropy = .74, latent class probabilities 1 = .89, 2 = .90, 3 = .87

between the first and last score. We also export class and probabilities from the growth mixture model for further analysis and use them to test the resultant findings for sensitivity to confidence in classification by dropping cases with classification probabilities that indicate greater uncertainty and retaining only cases with high probability (p ≥ .80).

After exporting the data from the GMM for further exploration, we compared results for the last LSI-R score of the four to class dummy variables to explore whether class

assignment using linear growth over time added predictive ability beyond the last measure of the LSI-R score. The last data point has considerable influence on forming the growth line for the classes, so it shares a great deal of covariation on recidivism with trajectory assignment and it should not be examined in the same model. Nevertheless, model comparisons of logistic regression models indicate that class membership outperforms the last LSI-R score and change in LSI-R score as recidivism predictors. The last LSI-R score of the four explains less variation (Nagelkerke $R^2 = 25$ %) than the class dummies (Nagelkerke $R^2 = 34$ %) and classifies relatively poorly (68 % correct-70 % correct). Change in LSI-R first to last compares comparably to the last data point in the 4 LSI-R data (Nagelkerke $R^2 = 25$ %; correct classification = 68 %). When only the 185 high confidence of classification cases are used (≥.80), the class variables outperforms by slightly improving correct classification (Nagelkerke $R^2 = 32$ %; correct classification = 71 %).

Using the exported classes in logistic regression in the three LSI-R score data revealed a similar but stronger difference. With Class 1 as the reference category, the classes alone explained more variation and classified much more effectively (Nagelkerke $R^2 = 52$ %; correct classification = 83 %) than did last LSI-R data point in a separate model (Nagelkerke $R^2 = 15$ %; correct classification = 67 %). They also outperformed change in LSI-R from the first to last score (Nagelkerke $R^2 = 15$ %; correct classification = 67 %). The 288 high confidence of classification cases predicted even better with the three LSI-R data, just as it did for the 4 LSI-R data (Nagelkerke $R^2 = 61$ %; correct classification = 87 %). In both data sets, the assigned classes for LSI-R score development appear to be better predictors of recidivism than the last LSI-R data point and the first to last LSI-R change. We do not report these comparative, supplementary results in tables, as the coefficients add little information that is not found in the reported GMM.

In the logistic regression of recidivism using high confidence in assignment cases from the 4 LSI-R data, those in Class 3, the consistently high class, were more likely to recidivate and the difference was significant (B = 1.40, p = .000). In the three LSI-R data, for high confidence of classification cases, those in Class 2 were more likely to recidivate than those in Class 1 (B = 1.22, p = .053) and those in Class 3 were much more likely to recidivate than those in Class 1 (B = 5.289, p = .000). The most conceptually important distinction is between Class 2, the high but declining Class, and Class 3 the high stable class. Class 3 members were much more likely to recidivate than Class 2 members (B = 5.435, p = .000).

Overall, the evidence leans toward support for predicting recidivism by classifying on the LSI-R trajectory. Yet, there may be reason to be wary of results based on low entropy of classification. At least for practical purposes one would not want to hazard classifying an offender on the margins of the classes or between the two categories into a risk level to be used for substantive decisions. Criminal justice decisions could be made based on trajectories more confidently than for single data points, however, and this is especially true where confidence of classification is high.

Discussion

Taken together findings indicate that LSI-R latent class categories generated by a growth mixture model of risk scores are excellent predictors of recidivism. Across the data sets, those placed in the consistently low LSI-R group have recidivism rates

ranging from 0 to 9 %, those in the beginning high and declining group have recidivism rates ranging from 29 to 32 %, and those in the consistently high group have rates ranging from 61 to 90 %. While these results are from a single release class, and enthusiasm for the finding should be tempered and probably viewed as exploratory, they are more than theoretical in their potential implications and applications. It is no small thing to show that change matters even in a relatively high risk population. There also are several problems with the use of risk scores in prison populations that a more dynamic data collection and interpretation might be useful for overcoming. These include restricted variation and a high degree of stability in risk scores.

On average, released prisoners are at high risk for reoffending (Singh et al., 2011; Stahler et al., 2013; Vaughn, DeLisi, Beaver, Perron, & Abdon, 2012; Wodahl, Boman, & Garland, 2015; Zettler, Morris, Piquero, & Cardwell, 2015). Indeed, their riskiness distinguishes them from other criminal justice populations such as probationers. Consider the current four-LSI data. Using the state's LSI-R cut-offs, only 4 offenders fall in the low risk category (1–13), 59 offenders in the low moderate category (14–23), 90 in the moderate category (24–33), 56 in the moderate to high category (34–40), and 20 in the high category (41+). The mean score at wave 1 is at the top of the moderate range and remains there over time as is seen by the average for the fourth administration of the LSI-R.

Risk scores can be normed for the more risky clientele leaving the prison system, but the likely concentration of scores in a range from moderately high to high reduces predictive capacity in smaller data sets by comparison to samples of equivalent size that contain a broader range of offenders, such as combined parolee and probationer data. Where there is more variation in risk scores, there likely is more predictive capacity. Developmental models of risk add an additional sources of variation and the changes they reflect might be predictive over and above a static point.

Stability in scores may be another practical problem. Large parts of the LSI-R risk scores are nearly intractable (the *risks*), such as the criminal history section and the questions indicating stable psychiatric problems. Other sections also are likely to be highly stable such as substance problems that led to work or school failure due to the long time periods needed to reverse damage. In addition, some dimensions of scores are stable due to the fact that it is difficult for many reentering offenders to build a foundation of stability and conventional life; workplace failures and moves between employers are common and lasting residences are difficult for many to establish. Parole officers often are frustrated by the stability of risk indicators and prefer *need*s indicators, as the former do not allow them discretion to adequately assess changing improvements or signs of trouble for clients. A common complaint is that the LSI-R may not be as dynamic a measure as its creators intended (SUPPRESSED Citation). Our data indicates some stability between waves, but LSI-R scores do not appear to be as stable as these skeptics might assume. The Pearson correlation coefficients in the entire four-LSI sample are illustrative: wave 1 to 2 correlation .60, wave 2 to 3 correlation .61, wave 3 to 4 correlation is .63. This means that LSI score are predictive of subsequent waves, however, there also is considerable unexplained variation across LSI-R scores over time. Knowledge not only of last LSI-R scores, but also of how they are moving over time appears to be valuable information and it should be used, especially given that thresholds of risk often are set on simple categorical scores ranging from one to five that are based on recoding raw LSI scores from a single time point (Singh et al., 2011; Whitacre, 2006).

Conclusion

This study addressed whether data driven classes based on development of risk scores predicts 2-year recidivism in two discrete samples drawn from a single prison release class. Growth mixture modeling of linear development of LSI-R scores, identified three classes of parolees in the data: those who have relatively flat high LSI-R scores, those who have LSI-R scores that diminish significantly, and a consistently low group. Membership in the classes proved significant, with those in the declining group having much lower recidivism than those in the flat and high group, and both high groups recidivating at higher levels than the stable low scoring group.

The administration of LSI-Rs used here occurred in different institutions by different trained personnel. That three developmental groups exist that in both data sets that are predictive of misconduct and that they operate in the intuitive direction; speaks well of the trans-setting reliability and usefulness of the instrument (Rocque & Plummer-Beale, 2014). Data would be much better if all offenders had several LSI-R scores so that results would be general to the entire release class and ideally, for statistical analysis purposes, LSI-R administrations would occur at perfectly spaced intervals. It would be even better if there were near simultaneous administrations for all offenders. This is one of the disadvantages of administrative data compared to researcher-designed cohort research and it introduces limitations for the analysis of panel data. This is not an indictment of the correctional system, as scores were not collected with panel data analysis techniques in mind, but rather to better serve offenders.

Hazards of classification models are that they may have low entropy making confidence of individual assignment low and that they may reify groups that not only have little substantive meaning but also little practical value. We conducted supplementary analysis to evaluate the groups not only by examining their effect on the distal outcome but to examine sensitivity to confidence in assignment. Even with low entropy and all cases included, the classes have predictive value and this justifies storing and keeping record of risk scores over time. It appears that there is valuable information in the longitudinal movement of risk scores. Those interested in treatment should be gratified to see that developmental change matters, as does the group one's scores reflect even when controlling for current risk level.

Another hazard associated with developmental research is whether it is justified practically when cross-sectional models predicting offending can reveal so much. It is, after all, an added analytical burden and may be unduly complex to examine trajectories. In the current case there is reason for examining behavior over time in some form because the dependent variable does not occur at the same point that the independent variables occur. And, the risk measure is designed to be dynamic. Still, it is an open question whether there is significant information in the development of risk scores that is not in the latest score or in a simple change estimate.

This research has limitations. Data represent one type of risk score in one Midwestern state among a sample of persons drawn on an unusual characteristic for a group of parolees; they have exactly three or four LSI-R scores. While descriptive statistics show that this does not make them extremely distinguishable from the larger population of parolees, they differ in some important ways and the

generality of findings to broader groups is questionable. Another important limitation is that the model violates an important statistical assumption of the technique. Scores over time ideally should occur at regular intervals or researchers should be able to account for interval differences, and our data do not allow for that. Violation of this standard is common in corrections research, as change scores often do not account for time. We assert that the necessity of examining classes in risk scores and useful application of this technique outweighs strict adherence to the assumptions of equivalent time intervals. Criminal justice systems use and retain scores across time, and we are unaware of any risk assessment that officials present to all offenders in any system at equal times between waves. In many states, correction officials retain scores over multiple administrations of risk assessments. This is because it is believed that they contain valuable information about offender development, but typically only the last score is used. This research can be viewed as an exploration of whether trajectories contain predictive information even when the data that form them are not ideally structured.

This is a parole only sample. Using multiple data points can be viewed as a way to add variation to prison samples where the last data point will be high for many offenders, especially by comparison to a probation population. The current samples of prison parolees reveal a skewed distribution of risk scores toward the high end, but the four assessments provide more inter-individual variation and movement than a single score to assess risk.

The practical implications of these findings are two-fold. First, the potential presence of developmental indicators of risk should be explored for all risk-assessment instruments in corrections and it is important that such studies be replicated. It is possible that developmental trajectories for some of them contain more predictive information than a single score and officials always need better indicators of risk. In the current data, multiple LSI-R scores for an individual contain more information than a single score and additional predictive value beyond the latest score or even a change score. There is considerable evidence in the three class model in both data sets that some offenders exhibit a steady decline in risk scores and that these are less likely to offend. Second, and if it proves to be true of other measures, researchers of risk scores should focus on means of classifying new offenders into the appropriate risk group as their scores accumulate and perhaps even on building relevant information from developing records into a total risk score by incorporating change into assessment scores. If it turns out that development of risk scores is not predictive, only the last data point should be retained in official records as development of scores adds no value beyond a rough indicator or prompt for the next official assigned to score an offender. If developmental trajectories prove insignificant, officials might focus exclusively on the last administration of risk assessment and delete the record of others so as not to bias new measurements. Risk scores are not fundamental drivers of supervision decisions but evidence based assessments are increasingly important. For example, some parole offices in the state where our data are drawn, use scores to determine initially how often an offender reports in person. While there is much to be said for simple scoring procedures, any improvement in prediction is worth understanding. Our findings, from an admittedly exploratory study, suggest that investigators should use dynamics of risk to predict offender outcomes where possible and in the future. While it is far too early to advocate for such a device, it is easy to imagine an algorithm that assigns individual offenders to trajectories of risk based on their development.

References

Ægisdóttir, S., White, M. J., Spengler, P. S., Mugherman, A. S., Anderson, L. A., Cook, R. S., et al. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *Counseling Psychologist, 34*(3), 341–382.

Andrews, D. A., & Bonta, J. (1995). *Level of service inventory-revised: user's manual*. Toronto: ON: Multi-Health Services.

Andrews, D. A., & Bonta, J. (2003). *The psychology of criminal conduct*. Cincinnati: Anderson Co.

Andrews, D. A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Newark, NJ: LexisNexis/Matthew Bender.

Andrews, D. A., Bonta, J., & Wormith, S. J. (2006). The recent past and near future of risk/need assessment. *Crime & Delinquency, 52*, 7–27.

Andrews, D. A., Bonta, J., Wormith, J. S., Guzzo, L., Brews, A., Rettinger, J., et al. (2011). Sources of variability in estimates of predictive validity: a specification with level of service general risk and need. *Criminal Justice and Behavior, 38*(5), 413–432.

Asparouhov, T., & Muthén, B. (2008). Multilevel mixture models. In G. R. Hancock and K. M. Samuelson (Eds.), *Advances in Latent Variable Mixture Models* (pp. 27–51). Charlotte, NC: Information Age Publishing.

Blanchette, K., & Brown, S. (2006). *The assessment and treatment of women offenders: an integrative perspective*. West Sussex: Wiley.

Folsom, J., & Atkinson, J. (2007). The generalizability of the LSI-R and the Cat to the prediction of recidivism in female offenders. *Criminal Justice and Behavior, 34*(8), 1044–1056.

Gendreau, P., Goggin, C., & Smith, P. (2002). Is the PCL-R really the "unparalleled" measure of offender risk? A lesson in knowledge cumulation? *Criminal Justice and Behavior, 29*(4), 397–426.

Hsu, C., Caputi, P., & Byrne, M. K. (2011). The Level of Service Inventory-Revised (Lsi-R) and Australian offenders: factor structure, sensitivity and specificity. *Criminal Justice and Behavior, 38*(6), 600–618.

Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality and Psychology Compass, 2*(1), 302–317.

Labrecque, R. M., Smith, P., Lovins, B. K., & Latessa, E. J. (2014). The importance of reassessment: How changes in the LSI-R risk score Can improve the prediction of recidivism. *Journal of Offender Rehabilitation, 53*(2), 116–128.

Lussier, P., & Gress, C. L. (2014). Community re-entry and the path toward desistance: a quasi-experimental longitudinal study of dynamic factors and community risk management of adult sex offenders. *Journal of Criminal Justice, 42*(2), 111–122.

Manchak, S. M., Skeem, J. L., Douglass, K. S., & Siransosian, M. (2009). Does gender moderate the predictive utility of the level of service inventory-revised (LSI-R) for serious violent offenders? *Criminal Justice and Behavior, 36*(5), 425–442.

Mauratto, P., & Moffat, K. H. (2006). Assembling risk and the restructuring of penal control. *British Journal of Criminology, 46*(3), 438–454.

Prell, L. (2009). LSI-R scores: change matters. *Iowa Department of Corrections: Data Download, 12*(1), 1–2.

Reisig, M., Holtfretter, K., & Morash, M. (2006). Assessing recidivism risk across female pathways to crime. *Justice Quarterly, 23*(3), 384–405.

Rice, M. E., Harris, G. T., & Lang, C. (2013). Validation of and revision to the VRAG and SORAG: the Violence Risk Appraisal Guide—Revised (VRAG-R). *Psychological Assessment, 25*, 951–965.

Rocque, M., & Plummer-Beale, J. (2014). In the eye of the beholder? An examination of the inter-rater reliability of the LSI-R and YLS/CMI in a correctional agency. *Journal of Criminal Justice, 42*(6), 568–578.

Schlager, M. D., & Simourd, D. J. (2007). Validity of the Level of Service Inventory—Revised among African American and Hispanic male offenders. *Criminal Justice and Behavior, 34*(4), 545–554.

Silver, E. (1998). Actuarial risk assessment: reflections on an emerging social-scientific tool. *Critical Criminology, 9*(1–2), 123–143.

Simourd, D. (2004). Use of dynamic risk/need assessment instruments among long-term incarcerated offenders. *Criminal Justice and Behavior, 31*(4), 306–323.

Simourd, D. (2006). *Validation of risk. Needs assessment in the Pennsylvania department of corrections: final report—grant 15,288*. Harrisburg: Pennsylvania Department of Correction.

Singh, J. P., Grann, M., & Fazel, S. (2011). A comparative study of violence risk assessment tools: a systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review, 31*(3), 499–513.

Smith, P., & Cullen, F. T. (2013). Predictive validity and the impact of change in total LSI-R score on recidivism. *Criminal Justice and Behavior, 40*(12), 1383–1396.

Smith, P., Cullen, F. T., & Latessa, E. J. (2009). Can 14,737 women be wrong? A meta-analysis of the LSI-R and recidivism for female offenders. *Criminology & Public Policy, 8*(1), 183–208.

Stahler, G. J., Mennis, J., Belenko, S., Welsh, W. N., Hiller, M. L., & Zajac, G. (2013). Predicting recidivism for released state prison offenders examining the influence of individual and neighborhood characteristics and spatial contagion on the likelihood of reincarceration. *Criminal Justice and Behavior, 40*(6), 690–711.

Tillyer, M. S., & Vose, B. (2011). Social ecology, individual risk, and recidivism: a multilevel examination of main and moderating influences. *Journal of Criminal Justice, 39*(5), 452–459.

Vaughn, M. G., DeLisi, M., Beaver, K. M., Perron, B. E., & Abdon, A. (2012). Toward a criminal justice epidemiology: behavioral and physical health of probationers and parolees in the United States. *Journal of Criminal Justice, 40*(3), 165–173.

Vose, B. (2008). *Assessing the predictive validity of the Level of Service Inventory Revised: recidivism among Iowa parolees and probationers*. Cincinnati: OH. University of Cincinnati.

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*(4), 459–471.

Vose, B., Smith, P., & Cullen, F. T. (2013). Predictive validity and the impact of change in total LSI-R score on recidivism. *Criminal Justice and Behavior, 40*(12), 1383–1396.

Whitacre, K. (2006). Testing the Level of Service Inventory-Revised (LSI-R) for Racial/Ethnic bias. *Criminal Justice Policy Review, 17*(3), 330–342.

Wilson, H. A., & Guiterrez, L. (2014). Does one size fit all? A meta-analysis examining the predictive ability of the level of service inventory (LSI) with aboriginal offenders. *Criminal Justice and Behavior, 41*(2), 196–219.

Wodahl, E. J., Boman, J. H., & Garland, B. E. (2015). Responding to probation and parole violations: Are jail sanctions more effective than community-based graduated sanctions? *Journal of Criminal Justice, 43*(3), 242–250.

Zettler, H. R., Morris, R. G., Piquero, A. R., & Cardwell, S. M. (2015). Assessing the celerity of arrest on 3-year recidivism patterns in a sample of criminal defendants. *Journal of Criminal Justice, 43*(5), 428–436.

Zhang, J., & Liu, N. (2014). Reliability and validity of the Chinese version of the LSI-R with Probationers. *International Journal of Offender Therapy and Comparative Criminology, 59*(13), 1474–1486.

Andy Hochstetler is Professor of Sociology. He uses multiple methodologies both qualitative and quantitative to study Criminology and Corrections. Much of his work examines offender decision-making and outcomes by incorporating offenders' self-appraisals, estimated prospects for conventional success, and environmental contexts into understanding and predicting recidivism and crime continuation.

David J. Peters is an Associate Professor of Sociology and Extension Rural Sociologist at Iowa State University. He also conducts research on rural poverty and inequality, demographic change, agricultural adoption, and rural criminology. Dr. Peters is published in top journals in rural sociology, agricultural economics, regional science, and criminology; and is lead investigator on grants from USDA and NIJ.

Matt DeLisi is Professor and Coordinator of Criminal Justice Studies and Affiliate with the Center for the Study of Violence at Iowa State University. Professor DeLisi is Editor-in-Chief of the *Journal of Criminal Justice* and the author of more than 300 scholarly publications mostly in the areas of pathological criminality, psychopathy, self-control/self-regulation, inmate behavior, and the molecular/behavioral genetics of antisocial behavior. In 2012, Dr. DeLisi received the Fellow Award from the Academy of Criminal Justice Sciences.