Comparing Static-99 Scores of Incarcerated White, Black, and Latino Sex Offenders

Michele O. Waldron

B.A., Southern New Hampshire University, 2002
M.S., Antioch University New England, 2006

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Psychology in the Department of Clinical Psychology
of Antioch University New England, 2012

Keene, New Hampshire

ANTIOCH
UNIVERSITY
NEW ENGLAND

Department of Clinical Psychology

DISSERTATION COMMITTEE PAGE

The undersigned have examined the dissertation entitled:

COMPARING STATIC-99 SCORES OF
INCARCERATED WHITE, BLACK, AND LATINO SEX OFFENDERS

presented on March 30, 2012

by

Michele O. Waldron

Candidate for the degree of Doctor of Psychology
and hereby certify that it is accepted*.

Dissertation Committee Chairperson:
George Tremblay, PhD

Dissertation Committee members:
James Graves, PhD
Alejandro Leguizamo, PhD

Accepted by the
Department of Clinical Psychology Chairperson

Kathi A. Borden, PhD

on 3/30/12

* Signatures are on file with the Registrar's Office at Antioch University New England.

Acknowledgements

Table of Contents

List of Tables

Abstract

Risk assessments contribute to sentencing and parole decisions, and thus are among the highest stakes assessments in the mental health field. The Static-99 has become a standard element of risk assessments for sex offenders, yet its norms and predictive accuracy have been established primarily with Caucasian samples, while the incarcerated population is disproportionately minority. Scoring of the Static-99 depends heavily on history of criminal offenses; if patterns of offenses differ along ethnic lines, the possibility that offense history should be understood to have ethnically-specific predictive validity (that is, the predictive significance of a given factor differs by ethnicity) becomes more compelling. This study does not address predictive validity directly, but it does examine patterns of scoring on the Static-99 for White, Black, and Latino incarcerated male sex offenders. Static-99 scores from 427 incarcerated male sexual offenders (264 White, 79 Black, 84 Latino) from the Massachusetts Treatment Center revealed that Whites were more likely than Blacks or Latinos to sexually assault male victims. Blacks had higher scores than Whites or Latinos on items related to violence, and were more likely to offend against stranger victims. Statistical significance was not reached for the age, cohabitation, and unrelated victim items. Researchers have recently found that the Static-99 has variable accuracy with offender subgroups, such as non-White offenders, but it remains a better predictor of sexual recidivism than clinical judgment alone. Researchers continue to explore and understand the variables that predict sexual recidivism. Dynamic risk factors and normative groups will be important areas to research to enhance the accuracy of actuarial measures with non-White offenders.

Keywords: Static-99, sex offender, ethnicity

Predicting Recidivism Risk among Sex Offenders

The Static-99, a widely utilized actuarial measure, is often used as part of a sexual offender risk assessment. It was constructed from empirically validated static (unchangeable) risk factors to evaluate sexual or violent risk potential. The Static-99 is a combination of the Rapid Risk Assessment of Sex Offender Recidivism (RRASOR) and the Structured Anchored Clinical Judgment-Minimum (SACJ-Min). To score this measure, a clinician must review an offender's criminal record to produce a total score, which estimates the probability of violent and sexual recidivism (separate base rates are provided in the manual). The success of a risk assessment depends on its accuracy in predicting future offending for all of the groups on which it is used.

Differential validity amongst ethnic groups was not adequately scrutinized in the development of the Static-99. If a risk tool is to be considered cross-culturally valid, it must have the same meaning across ethnic groups (Van de Vijver, 2000). While the current study does not evaluate the predictive validity of the Static-99, questions are raised about method (sample) and item bias in the development of the Static-99. For instance, Hanson and Thornton (2000) used risk factors from a meta-analysis which considered few non-White samples. These authors also did not complete validation studies with non-White participants.

Recently, researchers have found preliminary evidence to suggest that there may be an ethnically driven pattern in scoring, potentially affecting the Static-99's predictive validity. Hanson (personal communication, 6/18/09) reported that the Static-99 predicted sexual recidivism "reasonably well" in Canada, United States, and continental Europe, while it worked "particularly well" in the United Kingdom, Australia, and New Zealand, but was not as accurate in Japan. Further, there was insufficient research to determine its validity in Asia, South

America, and Africa. Researchers in the adult and juvenile systems established preliminary

evidence of ethnic differences in scoring (Forbes, 2007) and predictive validity on actuarial

measures including the Static-99 (Chapman, Desai, Falzer, & Borum, 2006; Langstrom, 2004;

Schwalbe, Frazier, & Day, 2007). Ethnic patterns in scoring would suggest the need for the

development of norms and for future researchers to evaluate the Static-99's predictive accuracy

with diverse groups.

The goal of the present study was to support the development of cross-cultural

assessments by comparing scores on the Static-99 of incarcerated male offenders from three

ethnic groups: White, Black, and Latino. This study sought to replicate previous findings of

differential response patterns between White and Black groups, but also added a group of Latino

offenders, who have yet to be researched in this area. Static-99 scores from 427 incarcerated

male sex offenders were used for analyses. Because recidivism data were not available for this

sample, the predictive validity of the Static-99 cannot be examined.

Literature Review

This section reviews the following areas: (a) current uses for sex offender risk

assessments, (b) What is the Static-99? (c) ethnic patterns in the criminal justice system (d)

current research concerning ethnic differences on risk assessments, (e) ethnic patterns in

Static-99 item scores, (f) the potential for cross-cultural bias, and (g) ethical implications.

Current Uses for Sex Offender Risk Assessments

Current uses for actuarial assessments, such as the Static-99, include (a) sentencing

options, (b) institutional placement, (c) civil commitment, and (d) community notification. All of

these have important long-term consequences and are discussed in further detail below.

Sentencing options. Offenders' sentences are dictated by their perceived risk to the

community, with sentencing guidelines mandating higher minimum sentences for offenders who

are considered more likely to reoffend (Wood, 2006). As part of a capital punishment decision,

some states require risk assessments to determine the individual's danger to society (Claussen-

Schultz, Pearce, & Schopp, 2004). These risk assessments typically involve actuarial tools such

as the Static-99.

Generally, offenders who are determined to be a higher risk are more likely to receive

harsh consequences such as castration or the death penalty. Eight states presently allow either

chemical or surgical castration of convicted sex offenders, a procedure which is meant to lower

the offender's testosterone level and his sexual interest (Spaulding, 1998; Zott, 2008). California

passed a law in 1997 that requires repeat sex offenders and first-time offenders whose victims

are under the age of 13 to undergo chemical castration as a condition of their release or as part of

a treatment program (Connelly & Williamson, 2000; Harrison, 2007).

While Massachusetts does not presently require surgical or chemical castration, chemical

castration drugs are regularly used in community treatment settings. Chemical castration

involves an injection of antiandrogens, which lower and even eliminate free testosterone in the

offender's system. The theory behind the use of these drugs is that a lower or nonexistent

testosterone level decreases deviant sexual fantasies and desires, and eventually results in

impotence. The use of these drugs is controversial (Spaulding, 1998) due to the significant side

effects often associated with long-term use of these pharmacological interventions. These side

effects may include headaches, weight gain, blood clots, depression, insomnia, or difficulty

breathing (Spaulding, 1998). Surgical castration, in which the offender's testicles are removed,

leaves a permanent scar and longstanding stigma, which will render him impotent and prevent

him from engaging in some sexual acts within a healthy, age-appropriate relationship (Harrison,

2007).

Seven states allow the use of the death penalty in cases involving the sexual assault of a child (Zott, 2008). This author also noted that the Supreme Court upheld a 2003 decision where a man was sentenced to death for raping his stepdaughter.

Institutional placement. Currently, all newly convicted inmates in the Massachusetts correctional system are sent to an intake facility where their security level risk is determined. The resulting risk classification determines the degree of security deemed necessary in the institutional placement. The Massachusetts' classification system places inmates with some sexual crimes at a higher risk to harm staff and other inmates. An inaccurate risk determination could present a safety issue for the offender, staff, or other inmates.

Civil commitment. In 1990, Washington State passed the first sex offender civil commitment law after a repeat sex offender abducted and then sexually mutilated a seven-year-old boy. Since this ruling, twenty states, including Massachusetts, have enacted similar laws, permitting offenders to be incarcerated beyond their release date (Levenson & Prescott, 2009). Each state has its own regulations for commitment, but they all have one criterion in common: the risk of perpetrating future sexually harmful behavior (Jackson & Hess, 2007). The Association for the Treatment of Sexual Abusers (ATSA, 2005), a national association that advocates for empirically supported practices in sex offender assessment, treatment, and management, recommends that its members use an actuarial tool to measure risk as part of a comprehensive assessment to evaluate sexual dangerousness. As part of the civil commitment process, designated forensic psychologists offer their expert opinion regarding an offender's potential risk to the community. These experts routinely use the Static-99 as part of risk assessments (M. Henry, personal communication, February 27, 2009).

Community notification. Community notification plays an important role in public safety. The Jacob Wetterling Act requires sexual offenders to register their home and work information with local authorities (Zott, 2008). Meghan's Law, enacted after seven-year-old Meghan Kanka was raped and murdered in 1994, requires that sexual offenders' work and home addresses be made available to the public (Zott, 2008).

The Massachusetts Sex Offender Registry Board (SORB), the regulatory entity charged with managing sex offender information and risk potential to the community, established a classification system to determine which of three risk levels to assign an offender. The board determines the offender's level of risk using the following criteria: (a) the offender's criminal history, (b) the circumstances of the offense, (c) the presence or absence of physical harm caused by the offense, (d) whether the offense involved consensual contact between adults, and (e) other matters that demonstrate whether the offender presents a risk to reoffend (Criminal History Registry Board, 2004). Those considered the highest risk (level three) have their information posted on the Internet. The SORB uses static items (e.g., criminal history) that are also found on the Static-99. In order to make accurate decisions about an offender's risk level, the SORB could benefit from further knowledge about the extent to which static variables may differ across ethnic groups.

 Obviously, all of these uses for risk assessments with sexual offenders carry very substantial and enduring consequences for the offender and the community. This makes the test's accuracy of utmost importance.

What is the Static-99?

The Static-99 is a widely utilized actuarial tool used by psychologists to evaluate an offender's risk of sexual recidivism (Jackson & Hess, 2007; McGrath, Cumming, & Burchard,

2003). Actuarial risk assessments provide an objective approach to evaluating risk compared to

purely clinical judgment, which is subject to personal feelings and prejudices (Craig & Beech,

2010). Actuarial assessments are often comprised of static (unchangeable) risk factors, which can

evaluate long-term risk, but they cannot address changes in risk over time. Current best practice

suggests that an assessment that evaluates risk should include actuarial tools in addition to

clinical judgment (Craig & Beech, 2010).

Actuarial assessment risk factors were developed from two widely cited meta-analyses

(Hanson & Bussiere, 1996; 1998) of violent, sexual, and general recidivism. These studies

identified risk factors that became the basis for most of today's actuarial risk measures with

sexual offenders. These factors were classified as either static or dynamic (changeable). Several

actuarial risk instruments have used Hanson and Bussiere's (1996, 1998) variables to predict

sexual and violent recidivism. These include the Sex Offender Risk Appraisal Guide (SORAG;

Quinsey, Harris, Rice, & Cormier, 1998), Rapid Risk Assessment for Sexual Offense Recidivism

(RRASOR; Hanson, 1997), and Thornton's Structured Anchored Clinical Judgment (SACJ;

Grubin, 1998). The SORAG is a derivative of the Violence Risk Appraisal Guide (VRAG). Both

were designed to assess violence and sexual risk (Quinsey et. al, 1998). The SACJ is rarely used

in the United States but is consistently used in England and Wales. The RRASOR and the SACJ

were developed to be brief actuarial assessments for sexual and violent recidivism (Hanson &

Thornton, 1999). The Static-99 is a combination of the RRASOR and SACJ. According to its

authors, Harris, Phenix, Hanson, and Thornton (2003), the Static-99 predicts sexual offense and

violence recidivism with "moderate accuracy."

Harris et al. (2003) reported that the Static-99 can be used with "adult males who have

been convicted or charged of at least one sexual offense against a child or nonconsenting adult

and can include first time offenders" (p. 5). An offender's criminal record and the historical

variables contained within is the primary source of information on which the Static-99 will be

rated. When the evaluator has scored each item (items are further described in the Method

section), the scores of these items are added together, creating a total numerical score that then

categorizes an offender's risk of reoffense (e.g., low, moderate, high). It is important to note that

a distinction is not made between violent and sexual risk when calculating the total score, but

base rates are available for each category. Although the Static-99 incorporates items that are

designed to speak to risk of violence and general recidivism, the total score is used to help

predict the risk an individual poses of recidivating sexually. Norms are not available and

comparisons are made with a standardization sample of adult, male offenders.

Ethnic Patterns in the Criminal Justice System

The terms race and ethnicity are often used interchangeably. Differences in how

researchers operationalize race and ethnicity make it difficult to synthesize findings across the

reviewed literature. Many of the participants in this study listed their ethnic origin; therefore,

ethnicity is used rather than race. There is little research about the comparability of sexual versus

violent risk factors among ethnic groups in the countries using the Static-99. Patterns in the

criminal justice system suggest that African Americans are arrested at disproportionate rates for

violent offenses when compared to other ethnic groups (Blumstein, 2009) and Whites are

disproportionally represented among sexual offenders (West & Templer, 1994; Wheeler &

George, 2005). According to the US Department of Justice (U.S. DOJ; 2007), in 2001 sexual

offenders represented less than 5% of the total correctional population in the US, and, while

ethnic minorities comprised 64% of incarcerated adults in the prison system, they accounted for

a much smaller proportion of sex offenses. Of all sexual offenders incarcerated for rape, 52%

were Caucasian, 44% were African-American, and 4% were from other races. Of those

incarcerated for sexual assault, 74% were Caucasian, 23% African-American, and 3% from other

races. The ethnic distribution of sex offenders versus general criminal offenders in

Massachusetts mirrors the national statistics above (L. Sampson, personal communication,

2/11/08).

If African Americans are more likely to be arrested and convicted of general criminal

offenses, then they will be more likely to score higher on Static-99 items related to criminogenic

factors. There are multiple reasons for these differences that go beyond the scope of this study,

but they have the potential to impact the severity of Static-99 scores since an individual's

criminal history (including arrests and convictions) is used in scoring. Also, given that the Static-

99 items do not distinguish between violent and sexual recidivism risk, it is unclear which items

in a high score contribute to an offender's risk to violently versus sexually reoffend.

Current Research Concerning Ethnic Differences on Risk Assessments

Few studies have explored ethnic patterns in scoring and predictive validity among

various ethnic groups using actual measures. Predictive validity refers to whether a given test

score indicates the same probability or risk potential across ethnic groups (Kazdin, 2003). The

following section outlines the current research on juvenile and adult risk assessment measures

that were evaluated with various ethnic groups. A discussion of available research comparing

scores of different ethnic groups on the Static-99 items follows.

Juvenile risk assessments. In response to the Juvenile Justice and Delinquency

Prevention Act of 1974, the juvenile system conducted several studies over the past 35 years

regarding the impact of ethnicity on the legal system. This act has been amended twice since its

enactment in order to address the disproportionate amount of minority offenders in the juvenile

justice system (Chapman, Desai, Falzer, Borum, 2006). There appears to be a wide range of thought about the impact of ethnicity on risk to recidivate violently or sexually. The Static-99 cannot be used to evaluate risk potential with juveniles, but an offender's actions as a juvenile are included in scoring.

Of the four studies cited, two (Schwabe et al., 2007; Schwalbe et al., 2004) reported higher recidivism rates among Blacks than Whites, whereas the other two studies (Chapman et al., 2006; Schwalbe et al., 2006) found the reverse. There are indications that Blacks commit more violent offenses from an earlier age than White and Latino groups (Chapman et al., 2006; Schwalbe et al., 2007), which would elevate scores on three items on the Static-99. This increase could cause an inflated total score in comparison to other ethnic groups. Researchers questioned whether risk factors were homogenous with all populations and suggested that more research is needed in this area (Schwabe et al., 2007; Schwalbe et al., 2006; & Schwalbe et al., 2004). Further research should be conducted to determine whether they should be part of the risk assessment or be explored as possible "nuisance variables" (Van de Vijver, 2000).

Adult Risk Assessments. This section describes the five available studies. Three studies (Langstrom, 2004; Smallbone, 2011 ; Varela, Boccaccini, Gonzalez, Murrie, & Caperton, 2011) evaluated the predictive accuracy of the Static-99 with different ethnic groups. Two other studies (Forbes, 2007; Sudo, Sato, Obata, & Yamagami, 2006) compared how various ethnic groups scored on Static-99 items, but they did not have recidivism data to comment on predictive validity.

Sjostedt and Langstrom (2001) cross-validated the use of the Static-99 and Rapid Risk Assessment for Sexual Offense Recidivism (RRASOR) with a sample of 1,400 Swedish prisoners. They noted that these tools were primarily validated on Canadian and US samples and

may provide differences in predictive accuracy with their Swedish sample. In this study the

Static-99 demonstrated moderate accuracy in predicting violent recidivism (AUC = .74) and

sexual recidivism (AUC = .76). These authors also completed item analyses to determine if the

items contributed to the tool's predictive accuracy. Four Static-99 items yielded odds ratios

greater than 3 for the probability of any violent recidivism (included sexual offenses): scoring

"2" (*OR* = 4.19) or "3" (*OR* = 5.35) on prior sexual offenses, having four or more prior

sentencing dates (*OR* = 4.00), prior nonsexual violence (*OR* = 3.93), and having any stranger

victims (*OR* = 3.27). Sjostedt and Langstrom also found the item index nonsexual violence

moderately predictive, and having any male victims not significantly predictive in their study.

These researchers suggested the need for future researcher to compare the predictive accuracy of

the Static-99 with other ethnic groups.

Subsequently, Langstrom (2004) administered the Static-99 and RRASOR to members of

three ethnic groups: 1,085 Nordic, 49 European, and 128 African Asian. The predictive validity

of the Static-99 with these three groups was assessed using the Receiving Operator Statistic,

which determines the Area Under the Curve (AUC). The Static-99 was moderately accurate with

the Nordic (.76) and European groups (.79); however, it did not differentiate between African

Asian recidivists and nonrecidivists (.50). Langstrom compared characteristics and items among

the groups, but did not offer recidivism data for each item. The African Asian group was more

likely to have sexually offended against a stranger or a nonrelative victim. This group was

predominantly comprised of young, single males (related to Static-99 item regarding living with

a partner for at least two years), and obtained higher total scores on the Static-99 when compared

to the European and Nordic groups. The ethnic groups in this study are not easily generalized to

US ethnic groups because Sweden's demographic composition is different from that of the

United States. However, Langstrom was the first to suggest that the Static-99 may not have the same predictive accuracy for all ethnic groups.

Another study by Varela et al. (2011) evaluated the predictive accuracy of 1,911 Black, White, and Hispanic offenders on the Static-99 and Static-99R. Using the area under the curve (AUC) statistic to determine predictive accuracy, it found no scoring differences across the groups (White = .62; Black = .63; Hispanic = .58). They did not report item scores.

A more recent study by Smallbone (2011) compared the Static-99 scores and recidivism rates for 399 indigeneous and nonindigineous sex offenders in Australia. Indigineous offenders scored significantly higher, twice as likely to to be arrested for a sexual offense, and significantly more likely to be arrested for violent and general offenses. Compared to nonindigineous offenders, the Static- 99's predictive accuracy was slightly lower with indigineous offenders (ROC = .76 vs .82). Smallbone suggested that the Static-99 had high sensitivity and low specificity. The author further said that false positive predictions were five times higher than true positive predictions.

Forbes (2007) conducted research with 1,612 White and Black convicted sex offenders using three popular actuarial assessments: Static-99, RRASOR, and MnSOST-R to review potential ethnic patterns in scoring frequency. Forbes reviewed item and total scores given to Black and White offenders on all three instruments and found that Black offenders scored higher on total and violence item scores than White offenders. Forbes found that Blacks were more likely to score on items related to general criminality and Whites were more likely to score on items related to sexual deviance. On item scores, Blacks' sexual offenses were more likely to involve threats or use of force, stranger victims, unrelated victims, and female victims. In contrast, White offenders were more likely to be married, have no history of violence, have

longer histories of sexual offending, and have victims who were known or related to them and were male. When Forbes conducted item-total correlations, he found differences, but "none of them were strong enough to suggest that any item contributed to total scores more than others" (p. 66). He further noted that the internal consistency of the measures was poor and suggested that it may have been due to each measure viewing the construct of risk differently.

Another study by Sudo et al. (2006) compared a small group of Japanese offenders (n = 45) to "Western Offenders" using the Static-99. The authors compared a high-risk group (those scoring 6 and above; n = 9) and a low-risk group (those scoring under 6; n = 36) with other demographic characteristics. The authors reported that few studies had compared the characteristics of Japanese sexual offenders to "Western sex offenders." Sudo et al. suggested that their high-risk group was comparable to Western sexual offenders because they had a history of delinquency. Characteristics of this group included acting alone and being under the influence of substances.

Findings from the above studies demonstrate preliminary evidence of ethnic patterns in scoring on the Static-99 and other actuarial measures. The current study was expected to replicate Forbes's (2007) recent conclusion that Blacks are more likely to score on items related to criminogenic factors (i.e., prior sentencing dates, stranger victims), whereas Whites are more likely to score on items related to sexual deviance (Hanson & Thornton, 2000). No research was found in the adult literature regarding sexual offender risk potential with Latino offenders. Latinos have not been included in any studies, despite the fact that they represent an increasing portion of the US population (Anastasi & Urbina, 1997; Groth-Marnat, 2003).

Ethnic Patterns in Static-99 Item Scores

The following sections present the limited literature concerning possible ethnic

differences in scoring on various Static-99 items.

Sexual offense history items. The Static-99 has two items that address the offender's

sexual offense history: prior sex offenses (charges and convictions), and convictions for

noncontact sex offenses. Rearrest rates for Blacks and Whites are about the same (5.6% and

5.3%), and higher than those for Latinos (4.1%; US DOJ, 2003), whose rates may be

underreported because they are often deported upon their release, after which their offense

histories are no longer reflected in our national databases. Blacks and Whites may score

differently on the prior offense item since both charges and convictions are considered in

scoring. Blacks are arrested for sexual offenses at a higher rate than Whites, but Whites are

convicted more frequently and serve longer sentences (FBI, 2003; Maxwell, Robinson, & Post,

2003). For noncontact sexual offenses, Wolak, Mitchell, and Finkelhor (2003) found that Whites

are overrepresented among internet sexual offenders (92% white), child pornography offenders

(noncontact offense; 89%), and crimes for sex transportation (noncontact offense; 70%).

Violence and criminal history items. The Static-99 has three items that pertain to the

offender's criminal history and to nonsexual violence: (a) "prior sentencing dates," (b) "prior

nonsexual violence," and (c) "index nonsexual violence." There is a general consensus in the

literature that Blacks and Latinos are more likely to have longer criminal histories than Whites

(Demuth & Steffensmeier, 2004; Forbes, 2007; US DOJ). Similarly, the research indicates that

Blacks (23.2%) and Latinos (22.3%) are more likely to be arrested for a violent offense than

Whites (19.4%; Demuth & Steffensmeier, 2004; Forbes, 2007; Snyder, 2003; Walsh, Swogger,

& Kosson, 2004).

Victim characteristic items. The Static-99 has three items about victims: (a) "any male

victims," (b) "any stranger victims," and (c) "any unrelated victims." The limited available

literature concluded that White perpetrators are more likely than Blacks to target male victims (Forbes, 2007; Frazier, 1993; US DOJ, 1997), and that Blacks are more likely to offend against unrelated and stranger victims, while Whites more often offended against acquaintances or related victims (Forbes, 2007; Langstrom, 2004; Muir & MacLeod, 2003). No research regarding victim characteristics was available on Latino offenders.

Age item. Literature (Wolak, Mitchell, & Finkelhor, 2003; US DOJ, 2007) indicates that White offenders tend to be older than Black offenders at the time of their first offense. However, a small study (n = 94) by Heilbrun, Emory, and Cross (1979) found no differences in age between White (M = 29.38) and Black (M = 28.22) rapists. Again, no research was found regarding age patterns among Latino offenders.

"Ever lived with . . ." item. For this item, Forbes (2007) assumed that marriage implied cohabitation. Forbes (2007) found that Whites were more likely than Blacks to have been married, but this finding did not specify the length of the marriage (the Static-99 item requires at least two years). No research was found regarding whether Latino offenders cohabitate with their partners.

Total scores. Forbes (2007) found that Blacks scored higher on all three actuarial measures used in his study, and that Blacks scored more often on the items related to violence when compared to Whites. Forbes also found that there were ethnic differences in item-total correlations, but none of the comparisons were significant enough to suggest that any one item contributed more to the total score. Langstrom (2004) also found that the Static-99 did not accurately distinguish recidivists from nonrecidivists with his group of African Asian offenders. Forbes indicated that Blacks displayed a pattern of general criminality, whereas Whites showed more sexual deviance. Sexually deviant items (particularly since one item can add up to three

points) would contribute more points to the total score than the general criminality items. If

criminality and sexual behavior, in fact, carry equal validity for predicting subsequent reoffense,

this pattern could result in overestimation of risk for Whites, compared with Blacks. The field

would benefit from future research in this area. No research was available regarding how Latino

offenders would score on the Static-99. Given the United States' large Latino population, this is

an important group to include in the present study.

The Potential for Cross-Cultural Bias

Interpretation of measures for minority versus majority groups has been an ongoing and

controversial topic (Groth-Marnat, 2003; Padilla & Borsato, 2008; Padilla & Medina, 1996). One

way to increase the differential validity and competence of an assessment is to reduce the

potential for test bias (Van de Vijver & Poortinga, 1997). Test bias in cross-cultural assessments

refers to "a lack of psychological meaning of test scores across cultural groups" (Van de Vijver,

2000, p. 88). In other words, a test score may not have the same significance for individuals from

different ethnic groups. Van de Vijver identified sources of bias in multicultural assessments to

include method and item bias. As noted above, researchers have already found evidence of

differences in item and total scores (Forbes, 2007) and predictive validity (Langstrom, 2004;

Schwalbe, Frazier, & Day, 2007; Chapman et al., 2006; Smallbone, 2011) among various ethnic

groups on the Static-99 and other actuarial risk assessment measures. Sections below discuss

potential sources of bias in the Static-99's development.

Method bias. Van de Vijver (2000) described method bias as "the presence of nuisance

variables due to method-related factors" (p. 92). A type of method bias, sample bias, occurs

when interpretation of an individual's test scores is compromised because the test's norm group

has different characteristics (e.g., race, socioeconomic status, and gender) than the person tested

(Van de Vijver & Poortinga, 1997). For example, a study by Sudo, Sato, Obata, and Yamagami (2006) reported potential sample bias in their study of Japanese offender scores on the Static-99 when they found a significant number of offenders who were mentally retarded in one of their groups. The authors noted that bias may have occurred in their sample because the test had never been normed for mentally retarded individuals. Norms have not been created for the Static-99 with different ethnic groups and scores are compared with the standardization sample. When an offender does not match the demographics of the standardization sample, then his scores may not have the same meaning since "nuisance variables" are more likely to be present.

There were two notable examples of potential sample bias in the development of the Static-99. The first example was in the initial identification of risk factors for recidivism. As stated above, Hanson and Bussiere (1996, 1998) conducted pivotal meta-analyses (n=28,972) that provided the sex offender field with static and dynamic factors, which differentiated sexual and violent offenders who recidivated from those who did not. Hanson and Bussiere (1998) focused on the size of the correlation with recidivism data using median r values and r+ (weighted average) along with Q, which also measures variability across studies. These authors noted that correlations greater than .30 would be large and suggest recidivism rate differences of 30%, with .20 being moderate and .10 small. "Minority status" was a significant predictor for violent recidivism (Q = 11.21, r+ = .23) and general recidivism (Q = 49.35, r+ = .14; with outlier). Hanson and Bussiere (1998) indicated further that the majority of the studies used (27 out of 28) in this meta-analysis were conducted with "predominantly Caucasian" participants. There are many non-White offenders who do not fit into this standardization sample.

Sample bias has continued beyond the initial identification of risk factors for the Static-99, in that estimates of its effectiveness have yet to be explicitly compared across various ethnic

groups. The following explanation regarding the use of the Static-99 with different ethnic groups was offered in its Coding Rules Revised:

> Most members of the original samples from which recidivism estimates were obtained were white. However, race has not been found to be a significant predictor of sexual offense recidivism. It is possible that race interacts with Static-99 scores, but such interactions between race and actuarial rates are rare. It has been shown that the SIR Scale works well for Aboriginal offenders as it does for non-aboriginal offenders. The LSI-R has been shown to work as well for non-White offenders as it does for white offenders and as well for aboriginal offenders as it does for non-aboriginal offenders. In Canada there is some evidence that STATIC-99 works as well for Aboriginal sexual offenders as it does for whites. At this time, there is no reason to believe that the STATIC-99 is culturally specific. (Harris et al., 2003, p. 7)

The standardization sample was comprised of primarily White offenders, and indeed, US Department of Justice (2007) statistics indicate that sexual offenders are primarily White. Even if there are fewer non-White offenders available for research to explore the Static-99's performance, there is more research available today than there was eight years ago, when the coding rules were published. Recent studies (Langstrom, 2004; Smallbone, 2011; Varela, et al., 2011;) have demonstrated that the Static-99 is less accurate when predicting sexual recidivism with non-White offenders.

Item bias. Hoffer et al. (2005), described item bias as "when subjects with the same underlying psychological construct (e.g., power motivation) from different (cultural) groups react diversely to a given item (e.g., picture card)" (p. 690). Two examples of how different cultural beliefs and legal practices could impact scores on Static-99 items are presented below.

In the first example, a study of Hispanic sexual offenders in California by De Apodaca, Schultz, Anderson, and McLennan (2005) found that Hispanic men were arrested at a disproportionate rate for engaging in sexual acts with individuals under 14 years old. These authors further suggested that, while engaging in sexual acts with an adolescent is considered illegal and morally reproachable in the United States, it not considered deviant or illegal in all Latino cultures. These authors questioned whether the offender's behaviors were sexually deviant (or a risk), since he is following his own culturally acceptable standards. Because this behavior would not be considered illegal in his home culture, but it is illegal in the United States, his Static-99 score would vary depending on the country in which he committed the behavior.

In the second example, Sudo et al., (2006) reported potential bias related to Japanese offenders' scores on the Static-99. The authors indicated that sexual offenses were unlikely to be reported in Japan, and some of their laws made it more difficult for a sexual offense to be classified. Sudo et al. stated that in Japan a sexual assault against a male is not considered a sexual offense unless reported by the victim. Further, when a child (under 18) is sexually assaulted, coercion needs to be proven by the court. Without a finding of coercion, the charge is pled down to "violation of acts regarding juvenile custody," which is not considered a sexual offense (Sudo et al., 2006, p. 152). These laws, the authors reported, make it very difficult to accurately assess an offender's sexual behavior and also may underestimate his risk on items such as prior sex offenses, noncontact offenses, and male victims.

The prior sexual offense item on the Static-99 is scored from zero to three rather than zero or one like the other items. If an ethnic group were more likely to score on this item, their total score would be significantly higher due to the weight of this item in comparison to the other items. When scores were compared, researchers offered mixed results about which ethnic group

may be more likely to have a longer sexual offense history (Forbes, 2007; U.S. DOJ, 2003; Wolak et al., 2003).

In conclusion, the Static-99's development did not include extensive research regarding its comparability across ethnic groups and could potentially raise ethical concerns about the use of the Static-99 with groups for which it was not normed (Sue, 1999).

Ethical Implications

Unfortunately, criminal justice systems provide treatment and assessment that often fails to consider individual or ethnic differences. A paucity of available ethnic literature in the sex offender field is a contributor, but not an excuse for assumptions of generality (Sue, 1999). The American Psychological Association's Ethical Standards (2003) states:

> Psychologists administer, adapt, score, interpret, or use assessment techniques, interviews, tests, or instruments in a manner and for purposes that are appropriate in light of research on or evidence of the usefulness and proper application of the techniques. . . . Psychologists use assessment instruments whose validity and reliability have been established for use with members of the population tested. When such validity or reliability has not been established, psychologists describe the strengths and limitations of test results and interpretation. (p.12)

The creators of the Static-99 reported that there is a lack of evidence to suggest that the measure would be inappropriate to use with non-White offenders; however, as in any clinical evaluation, psychologists should report any potential limitations when interpreting test results. For example, a psychologist who is conducting an evaluation on an offender who does not possess similar characteristics of the Static-99's standardization sample should report this limitation. Uninformed decisions have the potential to cause harm to clients, which also violates

ethical standards of practice (American Psychological Association, 2003). The Association for

the Treatment of Sexual Abusers' Practice Guidelines (2005) suggested that its members

familiarize themselves with a measure's strengths and limitations to avoid formulating

conclusions that may go beyond the scope of the measure.

Summary

The Static-99, an actuarial tool designed to estimate risk potential for sexual and violent

recidivism, contributes to decisions with potentially irreversible effects for an offender, making

its accuracy of utmost importance. In the United States, Whites are disproportionally arrested for

sexual offenses (West & Templer, 1994; Wheeler & George, 2005), while far more Blacks are

arrested for violent crimes (Blumstein, 2009). An offender's criminal history, particularly sexual

and violent offenses, is used to score the items on the Static-99.

Researchers have found preliminary evidence of ethnic differences in item scores

(Forbes, 2007) and predictive inaccuracy (Langstrom, 2004), but more replication studies are

needed. In general, Forbes found that Blacks are more likely to score on items related to criminal

offending, while Whites are more likely to score on items related to sexual deviance. However,

Blacks are arrested for sexual offenses at a higher rate than Whites, but Whites are convicted

more often and serve longer sentences (FBI Crime Reports, 2003; Maxwell, Robinson, & Post,

2003). Among sexual offenders, Blacks and Latinos often have a longer criminal history than

Whites, which elevates their scores on the Static-99 (Demuth & Steffensmeier, 2004; US DOJ,

2003; Forbes, 2007). In terms of their choice of victims, Blacks were more likely than Whites to

offend against unrelated and/or unknown targets, while Whites were more likely than Blacks to

offend against males (Forbes, 2007; Frazier, 1993; US DOJ, 1997). At present, whether there are

ethnic differences in the predictive validity of the Static-99 remains controversial, with some

evidence that it is less accurate in predicting recidivism for ethnic minority offenders than for

Europeans or other Caucasians (Langstrom, 2004), and other data indicating no differences in the

predictive accuracy of total scores with White, Black, and Hispanic offenders (Varela et al.,

2011).

Possible sources of sample and item bias in the development of the Static-99 were

presented to offer potential reasons why various ethnic groups may score differently. US

offenders do not demographically match the White, Canadian offenders or Aboriginal samples

on whom the measure was standardized. The potential for item bias occurs in the variation of

legal practices and cultural beliefs towards sexual offenses. Differences in state and national

laws—to say nothing of variations in enforcement—ensure that arrests and convictions will have

different meaning in different places.

All of these sources of bias are relevant to the ethics of predicting recidivism. The

American Psychological Association (2003) and the Association for the Treatment of Sexual

Abusers (2005) recommend clinicians are aware of test limitations (and strengths) when

administering tools on offenders whose characteristics are not similar to the normative sample.

These limitations should also be discussed when interpreting results, particularly since these

results can lead to irreversible consequences for offenders.

Method

Archival data from a sample of 427 intake assessments of incarcerated male sex

offenders from the Massachusetts Treatment Center were partitioned into three groups: White,

Black and Latino. The data were analyzed using the *Statistical Package for the Social Sciences*

16.0 (SPSS). Analysis of variance (ANOVA) was used to compare means for different ethnic

groups on the Static-99, followed by Tukey's Honestly Significance Difference (HSD) Test to

identify which of the three group means differed from each other. Chi-square tests of association were conducted to determine which items were associated with the difference in total scores among the groups. Further, the Gardner Method (Macdonald & Gardner, 2000) was used for post hoc analysis on items where significant associations were found, to determine which groups contributed to the significant association. The following section discusses: (a) research questions and hypotheses, (b) participants, (c) statistical power, and (d) analysis.

Research Questions and Hypotheses

The following research questions and hypotheses were formulated based on the literature review, which indicated that Black sexual offenders tended to have a more anti-social offense pattern and White sexual offenders were more sexually deviant (Forbes, 2007). This study sought to replicate these findings and incorporate other research as noted in the above section, Ethnic Patterns of Static-99 Items:

1. What are the descriptive characteristics (e.g., mean, median, mode, standard deviation) of the Static-99 Total Scores of White, Black, and Latino male sex offenders from the Massachusetts Treatment Center?

2. Are there significant differences in Static-99 Total Scores among White, Black, and Latino male sex offenders in the Massachusetts Treatment Center? It was hypothesized that there would be a statistically significant main effect of ethnicity on Static-99 Total Scores, with the Black group and the Latino group exhibiting higher Static-99 Total Scores than the White group. Furthermore, the Black group would have a higher total score than the Latino group.

3. Are there differences in offender age across the three ethnic groups? It was hypothesized that the White group would have a significantly greater proportion of participants who are aged twenty-five and older than the Black and Latino groups, but the Latino group

would have more participants than the Black group.

    4. Do the three ethnic groups differ in the proportion who received a score of yes to the question on whether they report having "ever lived with a lover for at least two years?" It was hypothesized that the White group would have a significantly greater proportion of participants who have lived with a partner than the Black and Latino groups. There would be no significant difference for Black and Latino groups on this item.

    5. Do the three ethnic groups differ in the proportion who have any convictions for index nonsexual violence? It was hypothesized that the Black and Latino groups would have a significantly greater proportion of participants who had a conviction for an index nonsexual violent offense than the White group. The Black group would be higher than the Latino group on this item.

    6. Do the three ethnic groups differ in the proportion who have any convictions for prior nonsexual violence? It was hypothesized that the Black and Latino groups would have a significantly greater proportion of participants who have had a conviction for prior nonsexual violent offenses than the White group. The Black group would be higher than the Latino group.

    7. Do the three ethnic groups differ in the proportion who have charges or convictions for prior sex offenses? It was hypothesized that the Black group would have a significantly greater proportion of participants with prior sexual offenses than the White and Latino groups. The Latino group would have significantly more participants with prior sexual offenses than the White group.

    8. Do the three ethnic groups differ in the proportion who have three or less sentencing dates? It was hypothesized that the Black and Latino groups would have a significantly greater proportion of participants who had prior sentencing dates than the White group. The Black group

would be higher than the Latino group.

9. Do the three ethnic groups differ in the proportion who have convictions for noncontact sexual offenses? It was hypothesized that the White group would have a significantly greater proportion of participants who had convictions for a noncontact sexual offense than the Black and Latino groups.

10. Do the three ethnic groups differ in the proportion who have unrelated victims? It was hypothesized that the Black group would have a significantly greater proportion of participants who had any unrelated victims than the White and Latino groups. There would be no significant difference between the White and the Latino groups on this item.

11. Do the three ethnic groups differ in the proportion who have any stranger victims? It was hypothesized that the Black and Latino groups would have a significantly greater proportion of participants who have had any stranger victims than the White group.

12. Do the three ethnic groups differ in the proportion who have any male victims? It was hypothesized that the White group would have a significantly greater proportion of participants who have had any male victims than the Black and Latino groups. There would be no significant difference between the Black and Latino groups on this item.

Participants

Participants for this study were 427 White, Black, and Latino incarcerated male sex offenders from the Massachusetts Treatment Center who were assessed as part of their entry into a sex offender treatment program between July 2002 and December 2008. The Massachusetts Treatment Center is a medium security facility that consists of state and civilly committed sexual offenders. The present sample consisted of state sentenced inmates.

The terms race and ethnicity are often used interchangeably. Many of the participants in

this study listed their ethnic origin; therefore, ethnicity is used rather than race. Differences in

how researchers operationalize race and ethnicity make it difficult to synthesize findings across

the reviewed literature. In order to gain enough statistical power for comparison, three groups

commonly found in the United States (White, Black, and Latino) were included.

Participants were excluded from the study if their ethnicity could not be classified into

one of the three groups. Examples of ethnicities that did not fit into the groups are African, Afro-

Caribbean, or Asian. The three ethnicities sampled for this study represent the three largest

ethnic groups within the Massachusetts prison system.

Measure

Static-99 items and scoring. The Static-99 is comprised of ten items that were chosen as

the best predictors of sexual recidivism based on Hanson and Bussiere's meta-analyses (1996,

1998). See Appendix to review the Static-99 coding sheet. The following section details the

items, along with their scoring criteria:

1. "Age." Individuals who are 18-24.99 are given a score of one. Individuals who are

   aged 25 or older are given a score of zero.

2. "Ever lived with." Individuals who have ever lived with a partner for at least two

   years were given a score of zero; otherwise they were given a score of one.

3. "Index nonsexual violence." An index offense is the crime that the individual was most

   recently convicted of, which often is the reason for referral. If there is a violent charge

   as part of their presenting offense (e.g., assault and battery, murder), the individual is

   given a score of one. If not, the individual is given a score of zero.

4. "Prior nonsexual violence." A prior offense includes all offenses the individual was

   convicted of before the index offense. If there is a conviction for a violent offense

that was not sexual, then the individual receives a score of one; otherwise, he is given

a score of zero.

5. "Prior sex offenses." The scoring system is based on the number of charges and/or

convictions for sexual offenses (e.g., rape, indecent assault and battery on a child

under 14), ranging from 0 for no charges or convictions, to 3 for six or more charges

and/or four or more convictions.

6. "Prior sentencing dates." This score is computed from the number of sentencing dates

for any offense, regardless of whether the person was convicted. Three or fewer

sentencing dates result in a score of zero, whereas the presence of four or more

sentencing dates is coded one.

7. "Noncontact sexual offenses." If the individual was ever convicted of a noncontact

sexual offense (i.e., indecent exposure or possession of child pornography), he is given

a score of one. If he has never been convicted of a noncontact sexual offense, then he

is given a score of zero.

8. "Any unrelated victims." If the individual committed his sexual offense against

someone with whom he knew, but was not related to, he is given a score of one. He is

given a score of zero if the victims did not include an unrelated victim.

9. "Any stranger victims." A score of one is given for those individuals who commit their

offense against someone with whom the offenders did not know; otherwise, a score of

zero is given.

10. "Any male victims." A score of one is given for those individuals who commit their

offenses against male victims. Individuals whose offense is against females are given

a score of zero.

Item scores are summed for a total score, and the individual is then assigned to a risk category. Table 1 displays the score and risk categories. Harris et al. (2003) indicated that there was not an increase in recidivism rates between scores of six and twelve. They suggested that the lack of increase in these rates may have been attributed to diminishing sample size, and suggested, "The more risk factors, the more risk" (p. 3). They also noted that the confidence in an offender's high risk score increases with each additional score over six.

Table 1

Risk Category Based on Score

Score	Risk Category
0, 1	Low
2, 3	Moderate-Low
4, 5	Moderate-High
6+	High

Psychometrics. The predictive accuracy of the Static-99 is reflected in AUC analyses as explained above. Sjostedt and Langstrom (2001) summarized the Static-99's predictive accuracy (ability to differentiate between recidivists and nonrecidivists) to be between .37 and .96. These authors also found inter-rater reliability to be .90 using Cohen's *k* and .83 with *k* not included. Hanson and Thornton (1999) determined the internal consistency to be .80 in their manual, but only one researcher (Forbes, 2007) reported alpha levels in the eight years since the Static-99's development. Forbes found the internal consistency to be poor (α = Black .411 and White .530). Anderson and Hanson (2010) suggested that inter-rater reliability, rather than internal consistency, was a more appropriate indicator of reliability for the Static-99 because the items were chosen on empirical, rather than theoretical grounds, and thus would not be expected to show internal consistency.

Statistical Power

The statistical power of a study represents the probability that an existing effect or relationship will be detected as statistically significant (Cohen, 1992; Kazdin, 2003). The most complicated and largest analysis in this study is the 3 x 4 contingency table described in Hypothesis 6, which was evaluated using Pearson's Chi-square statistic. The effect size is typically estimated from similar existing studies; the two extant studies (Langstrom, 2004;

Forbes, 2007) that also compared Static-99 scores across ethnic groups yielded medium sized

effects. For a medium effect ($w = .30$) and an alpha level of .05, Cohen (1992) specifies that a

sample of 151 would be required to reach a statistical power of .80 for a 3 x 4 contingency table

($df = 6$). With 427 participants in this study, all tests were adequately powered.

Procedure

There were over 550 male sex offender intake assessments available from the

Massachusetts Treatment Center. Each assessment contained several variables that outline the

offender's background and governing offense. Intake assessments are typically completed when

an offender is transferred into the Massachusetts Treatment Center. After ethnicities were

reviewed, 427 were left that fit into one of the three groups. In this sample, the Static-99 was

scored from information obtained through a clinical interview and a review of the participant's

criminal history. To enhance the reliability of scoring, Static-99 scores for each inmate were

independently derived by two coders. These scores were then compared, the discrepancies were

discussed, and a final consensus score was documented in a Microsoft Access database.

The current study consisted of a secondary analysis from a larger study (Leguizamo,

Carrasco, & Peltzman, 2008). The inmates' names and Department of Correction identification

numbers were removed prior to creating the database for this study.

Data Analysis

The data were analyzed according to a three-step process. First, one-way ANOVA

compared Static-99 total scores across the three ethnic groups, followed by Tukey's HSD Test

for post hoc pair-wise comparisons. Second, due to the categorical nature of items, chi-squares

were used to investigate if there were statistically significant associations ($p < .05$) between

ethnic groups and item scores. Finally, if group differences were detected in item-level

responses, then the Gardner method was employed to determine which of the three groups differed from each other. This approach made a Bonferroni-type correction to account for the multiple chi-square tests and the portions of the sample being used at one time, in order to reduce the probability of a Type I error (Macdonald & Gardner, 2000). Glass and Hopkins (1996) suggested that the pair-wise comparison is sufficient noting that:

> the contrast-based error rate is advantageous for most applications since it is consistent with the rationale that researchers employ for almost all other hypotheses that they test and it does not suffer from the conservativeness (and loss of power) of the family error rate. (p. 450)

For the present study, the level of significance used for the Gardner Method was $p > .008$ to account for the multiple post hoc comparisons (.05/6) being calculated for each item. Cronbach's alpha was also used to determine the internal consistency of the items for each ethnic group and the total sample. An alpha of .70 or higher is a commonly accepted threshold for asserting that the items on a scale tap into a unitary underlying construct.

Results

This section first presents the sample characteristics of the three groups as noted in Table 2. Results are then reported by research question, including brief commentary concerning whether results were consistent with the associated hypothesis.

Sample Characteristics

The ethnic distribution of the sample was consistent with the Massachusetts Treatment Center population, which as of July 2009 was 64% White, and 14% African American, 14% Hispanic, and 8% other ethnic groups. Ethnic differences in offense patterns will be discussed as part of the exploration of Static-99 scores.

Table 2

Sample Characteristics: Age, Marital Status, SES, and Offense History

	Black n = 79 (18.5%)	White n = 264 (61.8%)	Latino n = 84 (19.7%)	Total n = 427 (100%)
Mean age	38	42	36	41
SD	9.4	12.6	11	11.8
Marital Status				
Single	42 (53.2%)	114 (43.2%)	40 (47.6%)	
Married	13 (16.5%)	49 (18.6%)	20 (23.8%)	
Separated	4 (5.1%)	7 (2.3%)	11 (13.1%)	
Divorced	7 (8.8%)	73 (27.6%)	8 (9.5%)	
Widowed	2 (1.3%)	4 (1.5%)	0 (0%)	
Engaged	7 (8.8%)	11 (4.2%)	1 (1.1%)	
Socio-Economic Status				
Low	23 (29.1%)	57 (21.6%)	38 (45.2%)	
Low Middle	14 (17.7%)	38 (14.4%)	8 (9.5%)	
Middle	11 (13.9%)	67 (25.4%)	11 (13.1%)	
Upper	0 (0%)	11 (4.2%)	0 (0%)	
Offense History				
Age at first arraignment	17 (5.9)	22 (12)	23 (11.6)	
Age at last arraignment	32 (8.5)	37 (12.6)	32 (11)	
Total Arraignments	13.6 (9.5)	9.7 (8.5)	8.4 (8.6)	
Total Charges	28.3 (17.3)	22.6 (18.6)	19.1 (15.6)	
Total Convictions	14.7 (9.7)	12.8 (11.1)	10.6 (8.1)	
Substance Abuse				
Charges	2.4 (3.5)	1.7 (3.2)	2.2 (4.1)	
Convictions	1.5 (3.4)	.9 (2.1)	1 (2.2)	
Property				
Charges	8.0 (9)	4.8 (7.8)	3.4 (4.9)	
Convictions	3.7 (4.6)	2.5 (4.9)	1.6 (2.7)	
Person (nonsexual)				
Charges	7.0 (6.4)	3.5 (5.1)	3.6 (5.4)	
Convictions	3.6 (4.4)	1.8 (3)	1.8 (2.5)	
Sexual				
Charges	5.3 (4.4)	7.9 (8.7)	5.6 (3.9)	
Convictions	3.7 (3.3)	5.9 (6.3)	4.3 (3)	

Note. Where percentages do not total 100%, it is due to missing or unavailable data.

Table 3

Reliability Analysis (Cronbach's alpha) of the Static-99 by Ethnic Group and Total Sample

	α
Black	.510
Latino	.459
White	.450
Total	.469

Internal Consistency of Static 99 Items

Cronbach's alpha for the full Static-99 is reported in Table 3, both for the sample as a whole and separately by ethnic group. Although not elaborated in the table, item-level analyses did not reveal any clear patterns of poorly performing items.

Ethnic Patterns in Static-99 Scores

Are there significant differences in Static-99 Total Score among White, Black, and Latino male sex offenders in the Massachusetts Treatment Center? One-way ANOVA revealed statistically significant differences among the three means; $F (2, 224) = 6.1$, $p < .01$ (see Table 4). Follow-up pairwise comparisons indicated that Blacks ($M = 4.4$; $SD = 2.3$) scored higher than Whites ($M = 3.7$; SD 2.2), consistent with our prediction. Contrary to expectation, Latinos did not differ significantly from Whites ($M = 3.2$; $SD = 2.1$).

Table 4

Mean Total Scores by Ethnic Group

	M	SD	*F*
Black	4.4_a	2.3	
Latino	3.2_b	2.1	
White	3.7_b	2.2	
Total	3.7	2.2	6.1*

Note. Differing subscripts indicate significantly different means.

*p< .05

Are there differences in offender age category across the three ethnic groups? Age is treated as a categorical value because Hanson (2001; 2002), a creator of the Static-99, found that offenders under 25 years old were more likely to recidivate compared to offenders over 25 years old. Chi-square analysis did not reveal statistically significant associations between offender age categories (0-24.99 and 25 and older) and ethnicity (See Table 5).

Table 5

Item 1 Age

Ethnicity	≤ 25	≥ 25
	n (%)	n (%)
Black (n = 79)	75 (94.9)	4 (5.1)
Latino (n = 84)	73 (86.9)	11 (13.1)
White (n = 264)	248 (93.9)	16 (6.1)
Total	396 (92.7)	31 (7.3)

Do the three ethnic groups differ in the proportion who report having "ever lived with a lover for at least two years" (item two of the Static-99)? Chi-square analysis revealed no statistically significant associations between ethnicity and cohabitation: ($\chi^2 = .5$; $df = 2$; $p = .778$; see Table 6), which was not an expected result.

Table 6

Item 2 Ever Lived with a Partner for at least two years

Ethnicity	Cohabitate > 2years	Did not live with Partner for 2 years
	n (%)	n (%)
Black (n = 79)	56 (71)	23 (29)
Latino (n = 84)	63 (75)	21 (25)
White (n = 264)	188 (71.2)	76 (28.8)
Total	307 (71.9)	120 (28.1)

Do the three ethnic groups differ in the proportion who have any convictions for index nonsexual violence (item three of the Static-99)? Chi-square analysis revealed statistically significant associations between ethnicity and convictions for index nonsexual violence ($\chi^2 = 12.8$; $df = 2$; $p = .002$: see Table 7). Post hoc analyses (The Gardner Method) indicated that Blacks (46.8%) scored higher than Whites (26.9%) and Latinos (25%), which was consistent with our prediction.

Table 7

Item 3 Index Nonsexual Violence

Ethnicity	No Index Nonsexual Violence	Index Nonsexual Violence
	n (%)	n (%)
Black (n = 79)	42 (53.2)	37_a (46.8)*
Latino (n = 84)	63 (75)	21_b (25)
White (n = 264)	193 (73.1)	71_b (26.9)
Total	298 (69.8)	129 (30.2)

Note: Differing subscripts indicate significant different proportions.

*p< .01

Do the three ethnic groups differ in the proportion who have any convictions for prior nonsexual violence (item four of the Static-99)? Chi-square analyses revealed statistically significant associations between ethnicity and convictions for nonsexual violence (χ^2 = 15.2; *df* = 2; *p* = .000; see Table 8). Post hoc analyses using the Gardner Method demonstrated that Blacks (65.8%) scored higher than Latinos (45.2% and Whites (41%) as expected.

Table 8

Item 4 Prior Nonsexual Violence

Ethnicity	No History of Violence	History of Violence
	n (%)	n (%)
Black (n = 79)	27 (34.2)	52_a (65.8)*
Latino (n = 84)	46 (54.8)	38_b (45.2)
White (n =264)	156 (59)	108_b (41)
Total	229 (53.6)	198 (46.4)

Note. Differing subscripts indicate significant different proportions.

*p< .01

Do the three ethnic groups differ in the proportion who have charges or convictions for prior sex offenses (item five of the Static-99)? Chi-square analyses revealed no statistically significant associations between ethnicity and frequency of prior sexual offenses (χ^2 = 6.2; *df* = 6; *p* = .403; see Table 9), which was not an expected result.

Table 9
Item 5 Prior Sex Offenses

Ethnicity	Charges	None	1-2	3-5	6+
	Convictions	None	1	2-3	4+
	Item Score	0	1	2	3
		n (%)	n (%)	n (%)	n (%)
Black (n = 79)		42 (53.2)	14 (17.8)	14 (17.8)	9 (11.2)
Latino (n = 84)		52 (61.9)	17 (20.2)	9 (10.7)	6 (7.2)
White (n = 264)		151 (57.2)	42 (15.9)	32 (12.1)	39 (14.8)
Total		245 (57.4)	73 (17.1)	55 (12.9)	54 (12.6)

Do the three ethnic groups differ in the proportion who have three or fewer sentencing dates (item six of the Static-99)? Chi-square analysis revealed statistically significant associations between ethnicity and having four or more (vs. 3 or less) sentencing dates ($\chi^2 = 13.5$; $df = 2$; $p = .001$; see Table 10). Post hoc analyses using the Gardner Method found a larger proportion of Black offenders (59.5%) had longer offense histories as opposed to White (40.5%) and Latino (32.1%) offenders. Contrary to expectation, Latinos did not have a longer offense history when compared to Whites.

Table 10
Item 6 Prior Sentencing Dates

Ethnicity	≤ 3	≥ 4
	n (%)	n (%)
Black (n = 79)	32 (13)	47_a (26)*
Latino (n = 84)	57 (23.1)	27_b (15)
White (n = 264)	157 (63.9)	107_b (59)
Total	246 (57.6)	181 (42.4)

Note. Differing subscripts indicate significant different proportions.

*p< .01

Do the three ethnic groups differ in the proportion who have convictions for noncontact sexual offenses (item seven of the Static-99)? Chi-square analyses revealed no statistically significant associations between convictions for noncontact sexual offenses and ethnicity ($\chi^2 = 4.1$; $df = 2$; $p = .131$; see Table 11), which was not an expected result.

Table 11

Item 7 Any convictions for Noncontact Sexual Offenses

Ethnicity	Convicted of Noncontact Offense	No History of Noncontact Offense
	n (%)	n (%)
Black (n = 79)	69 (20.4)	10 (11.4)
Latino (n = 84)	67 (19.8)	17 (19.3)
White (n = 264)	203 (59.8)	61 (69.3)
Total	339 (79.4)	88 (20.6)

Do the three ethnic groups differ in the proportion who have unrelated victims (item eight of the Static-99)? Chi-square analyses did not reveal statistically significant associations between ethnicity and unrelated victims ($\chi^2 = 8.4$; $df = 2$; $p = .015$; see Table 12), which was not an expected result.

Table 12

Item 8 Any Unrelated Victims

Ethnicity	No Unrelated Victims n (%)	Unrelated Victims n (%)
Black (n = 79)	13 (10.4)	66 (21.9)
Latino (n = 84)	30 (24)	54 (17.9)
White (n = 264)	82 (65.6)	182 (60.2)
Total	125 (29.3)	302 (70.7)

Do the three ethnic groups differ in the proportion who have any stranger victims (item nine of the Static-99)? Chi-square analyses revealed statistically significant associations between ethnicity and those with stranger victims ($\chi^2 = 19.8$; $df = 2$; $p = .000$; see Table 13). Post hoc analysis using the Gardner Method found that, as expected, Blacks were more likely (43%) to have stranger victims than Whites (18.9%) and Latinos (21.4%). Again, no significant differences were found between Whites and Latinos ($p = .616$).

Table 13

Item 9 Any Stranger Victims

Ethnicity	No Stranger Victims	Stranger Victims
	n (%)	n (%)
Black (n = 79)	45 (13.8)	34_a (33.3)*
Latino (n = 84)	66 (20.3)	18_b (17.7)
White (n = 264)	214 (65.9)	50_b (49)
Total	325 (76.1)	102 (238.9)

Note. Differing subscripts indicate significant different proportions.

*p< .01

Do the three ethnic groups differ in the proportion who have any male victims (item ten of the Static-99)? Chi-square analyses revealed statistically significant associations between ethnicity and offending against male victims ($\chi^2 = 25.1$; $df = 2$; $p = .000$; see Table 14). Post hoc analyses using the Gardner Method found that Whites (83.5%) were more likely than Blacks (7.2%) or Latinos (9.3%) to offend against male victims, as expected. Blacks and Latinos in this sample revealed no significant differences in the likelihood of male victims.

Table 14

Item 10 Any Male Victims

Ethnicity	No Male Victims	Male Victims
	n (%)	n (%)
Black (n = 79)	72 (21.8)	7_b (7.2)
Latino (n = 84)	75 (22.7)	9_b (9.3)
White (n = 264)	183 (55.5)	81_a (83.5)*
Total	330 (77.3)	97 (22.7)

Note. Differing subscripts indicate significant different proportions.
*p< .01

In summary, five (index nonsexual violence, prior nonsexual violence, 3 or more

sentencing dates, male victims, and stranger victims) out of the ten items were found to show

statistically significant ethnic differences at the *p*< .001 level. These results, as they relate to the

hypotheses, are discussed in the following section

Discussion

Sex offender risk assessments provide the criminal justice system with information to assist with decisions involving confinement, probation, parole, security level, and treatment. These assessments assist with an estimated prediction of the offender's risk to the community and also direct treatment providers toward risk factors to address in treatment. The purpose of this study was to detect ethnic patterns in scoring among Black, White, and Latino groups of incarcerated sexual offenders. The general hypothesis was that Blacks and Latinos would score more often on Static-99 items than Whites. This hypothesis was partially supported, with Blacks scoring higher more often than Whites, but Latinos frequently scored lower than Whites, which was not expected. This section discusses the following areas: (a) summary of results, (b) comparison to prior research, (c) implications of findings, (d) study limitations, and (e) closing remarks.

Summary of Results

This study compared ethnic group scoring frequencies on Static-99 items. Five out of the ten items yielded significant differences in scoring patterns, and the total score also revealed an elevated mean score for Blacks when compared to the White and Latino groups. Blacks scored more often on index nonsexual violence, prior nonsexual violence, three or more sentencing dates, and stranger victims. Whites scored more frequently on the item reflecting sexual assault of a male victim. Latinos did not yield statistically significant differences on any items when compared with the Black and White groups, which was not an expected result. These findings will now be compared to previous research.

Comparison to Prior Research

The current study cannot provide evidence of ethnic differences in the predictive validity

of the Static-99 because no recidivism outcomes were available. We can, though, examine the

extent to which the Static-99 items on which we found the greatest ethnic differences correspond

to the Static-99 items that Sjostedt and Langstrom (2001) identified as most predictive of violent

recidivism. Even perfect correspondence would not directly bear on differential validity across

the ethnic groups because it remains possible that ethnic differences in scores on highly

predictive items accurately reflect ethnic differences in risk of recidivism. Likewise, a complete

absence of correspondence would leave open the possibility that similar scores (across ethnic

groups) on the items Sjostedt and Langstrom found highly predictive for a mostly White sample

are differentially predictive for different ethnic groups. Still, it may be that ethnic differences in

the predictive validity of the Static-99 would most likely be associated with ethnic differences in

the scores on highly predictive items, among a sample of convicted offenders. Acknowledging

the speculative nature of such an inquiry, we proceed because the alternative is to intentionally

ignore the relevance of Sjostedt and Langstrom's unique data. Table 15 below summarizes the

comparisons in this section.

In Sjostedt and Langstrom's (2001) study, four Static-99 items yielded odds ratios greater

than 3 for the probability of any violent recidivism: having four or more prior sentencing dates,

prior nonsexual violence, having prior sexual offenses, and having any stranger victims. In our

sample, we found statistically significant ethnic differences on all but the first of these items,

with Blacks scoring higher than Whites and Latinos in all cases. Of the remaining two items on

which we found ethnic differences in our sample, Sjostedt and Langstrom (2001) found index

nonsexual violence moderately predictive, and having any male victims not significantly

predictive in their study.

Langstrom (2004) and Varela et al. (2011) evaluated the predictive accuracy of the Static-

99's total scores with various ethnic groups. Langstrom found that scores for the African Asian

group were significantly lower in predictive accuracy when compared to the other two white

groups, but Varela et al. found that the Static-99 predicted violent or sexual recidivism with mild

accuracy for Black (AUC = .65) and White (AUC = .63). However, these authors found that

Hispanics tended to have lower total scores when compared to the other two groups, which was

consistent with the current study group frequencies on Static-99 items. A comparison of

nonindigenous and indigenous offenders demonstrated moderate accuracy for both groups, albeit

lower for indigenous offenders (ROC = .76 vs .82; Smallbone, 2011).

Turning to ethnic differences in item scoring frequency, Varela et al. (2011) offered no

item-level comparisons. Langstrom (2004) did not provide data for two of the statistically

significant items from the current study (prior sentencing dates and index nonsexual violence).

Having prior convictions of a sexual offense did not show ethnic differences in the current study

or Langstroms's (2004), but did in Forbes' (2007) study. The current study found that Whites

were more likely to have a male victim, but this was not consistent with Langstrom's (2004) or

Forbes' (2007) findings. Langstrom (2004) and Forbes found "ever lived with a partner" and

"unrelated victims" significant, but this was not consistent with our study. Only Langstrom

(2004) found the age item to be of significance. Given that the Static-99 was normed primarily

on White offenders, it would be hazardous to assume that Blacks' higher scores reflect

proportionally higher risk of reoffense.

Table 15

Studies Comparing Static-99 item Scoring Frequency with Different Ethnic Groups

Item	Langstrom	Forbes	Current Study
Age	-	n.s	n.s.
Ever Lived With	AA > E, N	B > W	n.s.
Index Nonsexual Violence	-	B > W	B > W, L
Prior Nonsexual Violence	n.s.	B > W	B > W, L
Prior Sex Offenses	n.s.	B > W	B > W, L
Prior Sentencing Dates	-	B > W	B > W, L
Noncontact Offenses	n.s.	n.s.	n.s.
Unrelated Victims	AA > E, N	B > W	n.s.
Stranger Victims	AA > E, N	B > W	B > W, L
Male Victims	n.s.	W > B	W > B, L

Note. n.s. - Not significant, W - White, B - Black, L - Latino, AA - African Asian, E - European, N – Nordic

Implications of Findings

One cannot ignore the Static-99's differences in scoring and predictive accuracy as described in the previous section. Recently, researchers have also found that the Static-99 was less accurate in predicting sexual recidivism with other offender subgroups such as those with different offense (i.e., pedophile vs. rapist) and demographic (i.e., age) characteristics (Anderson & Hanson, 2010; Brouilette-Alarie, Proulx, Helmus, & Hanson, 2011; Smallbone, 2011; Varela et al., 2011). Historically, evaluating sexual offense recidivism has relied on the comparison of

individual information to group data, but due to the variability in Static-99 scoring with offender sub-groups, researchers (Brouilette et al., 2011) are starting to evaluate alternative approaches. While many evaluators in other fields adequately use group comparison to interpret individual scores (i.e., IQ), the consequences of inaccurate interpretations are more dire for sexual offenders. The Static-99 is far more accurate than clinical judgment and continues to provide the field with useful information, but researchers are always striving to improve the accuracy of actuarial measures to further protect communities from dangerous offenders. This section reflects on current movements in the field to increase the predictive accuracy of sexual offender risk assessments by highlighting areas needing improvement and upcoming areas of research.

The Static-99 is by far the most researched actuarial measure for sexual recidivism (Parent, Guay, & Knight, 2011), but also has more variability in its predictive accuracy across studies than would be expected by chance (Anderson & Hanson, 2010). The AUC is usually .70 (Anderson & Hanson, 2010), but scores have ranged from .92 (Thornton, 2002) to .50 (Langstrom, 2004). The reason for these disparities is unknown, but researchers have hypothesized it could be attributed to variability in research methods or that there truly is a difference in the predictive accuracy across groups of offenders (Anderson & Hanson, 2010). Sreenivasan, Weinberger, Frances, and Cusworth-Walker (2010) also surmised that variations in predictive accuracy across studies are attributable to differences between the normed sample and the individual tested. Another possible explanation is that the Static-99 is not measuring a unitary construct. The Static-99's internal consistency measured by Cronbach's alpha varied from .80 in the original sample (Hanson & Thornton, 2000) to .469 in the current study. It also remained weak when computed for Black, White, and Latino groups. Poor internal consistency may be a reflection of the Static-99 measuring multiple constructs, which may be differentially relevant to

sexual recidivism.

In their initial meta-analysis, Hanson and Bussiere (1998) found antisocial orientation and sexual deviance to be predictors of sexual recidivism and this was again supported in the next meta-analysis (Hanson & Morton-Bourgon, 2004). Other investigators, though, have proposed more specific risks attendant to these two distinct categories of offense history, such that antisocial orientation only predicted nonsexual recidivism, while sexual deviance only predicted sexual recidivism (Barbaree, Langton & Peacock, 2006; Brouilette-Alarie et al., 2011; Roberts, Doren, & Thornton, 2002). Brouilette-Alarie, et al. also determined through factor analysis that only two Static-99 items contributed to the antisocial orientation: prior nonsexual violence and prior sentencing dates. It would be interesting to see if these two categories produced the same results with non-White offenders. In the current study, Blacks scored higher on these two items. Blacks' higher scoring pattern on items related to antisociality is concerning since they are arrested at disproportionate rates for violent offenses in the United States (Blumstein, 2009), which could overestimate their risk to sexually reoffend and may be more predictive of violent or general offense risk.

Despite difficulties in predictive accuracy with offender subgroups, the Static-99 continues to be used because it remains more accurate than clinical judgment alone (Bengtson & Langstrom, 2007). Actuarial measures are used with other offender subgroups such as those with developmental disabilities, severe mental illness, and women despite the rather homogenous nature of the standardization sample. Researchers, evaluators, and clinicians continuously strive to provide the best services possible with the resources available, while at the same time, they are working to improve actuarial accuracy. The remainder of this section summarizes two areas that researchers are exploring which could enhance the accuracy of actuarial measures with offender

subgroups, such as those who are non-White.

First, researchers should continue their efforts to understand the dimensions of risk for various offender groups. It will be important to conduct validation studies with various ethnic groups, offenses, and demographic information to increase predictive accuracy and develop norms for these groups. The Personality Assessment Inventory (PAI; Morey, 2007) conducted its standardization sample with a US census matched sample, which included data on age, gender, and race. When interpreting the PAI, users can compare individuals to the score distribution from their diverse sample (Morey, 2007). The availability of similar information when using the Static-99 would allow evaluators to provide a more accurate and individualized risk assessment.

Second, since the Static-99 provides a narrow view of risk, it should be used in conjunction with other empirically supported measures to cast a wider net of risk factors and portray a potentially more accurate view of risk potential (Anderson & Hanson, 2010). Because it is unclear how risk factors vary across subgroups, having a larger group of factors to choose from will portray a broader picture of risk for the individual offender. Too many factors increase the risk of error, but too few also create potential error. The addition of dynamic (changeable) risk factors to static risk factors could provide a more flexible and individualized approach to risk, though it has yet to demonstrate superior risk prediction. Anderson and Hanson examined the predictive validity of a risk assessment combining the Stable-2007, a measure of empirically supported dynamic factors, with the Static-99. The resulting AUC was .81, as compared with an AUC of .77 using the Static-99 alone. While dynamic factors remain an area in need of research, it would be useful to conduct studies with non-White offenders evaluating if the dynamic variables (i.e., Static variables) are able to predict risk to sexually recidivate.

Despite the Static-99's limitations, it remains the most accurate tool available to measure

risk when compared to clinical judgment. In the past two years, researchers have begun to

expand their thinking about risk and acknowledge there is still a significant amount of

information to be acquired about what variables best predict risk of sexual recidivism. The

autonomous nature of each offense makes it difficult to generalize variables for a large group of

offenders. Utilizing dynamic factors and exploring the construct of risk for offender subgroups

are promising directions that could increase the predictive accuracy of sexual offender risk

assessments.

Limitations of the Study

Our reliance on archival data left us with some missing and marginally reliable

demographic data. Sample sizes for the Black and Latino groups were small compared to the

White group, albeit adequate for statistical power. Of greatest consequence, we did not have

recidivism data to explore for ethnic differences in predictive validity of Static-99 items.

Closing Remarks

In the past two years and since the current study was proposed, there has been exciting

and considerable research that considered a more individualized approach to potentially address

not only the variations in the Static-99's development and performance, but also provided

another perspective of risk as a construct. Future researchers should continue to replicate the

above studies and continue to explore the dimensions of risk, but on a more individualized level.

While the ability to predict sexual recidivism has come a long way since using purely clinical

judgment, there remains a longer journey to increase the accuracy of actuarial assessments to

predict sexual recidivism. The current study sought to contribute to a small, but growing amount

of research that is designed to support cross-culture actuarial assessments.

References

American Psychological Association (2003). *Ethical principles of psychologists and code of*

conduct. Washington, DC: American Psychological Association.

Anastasi, A. & Urbina, S. (1997). *Psychological testing.* Upper Saddle River, New Jersey:

Prentice Hall.

Anderson, D. & Hanson, R.K. (2010). Static-99: An actuarial tool to assess risk of sexual and

violent recidivism among sex offenders. In R. Otto & K. Douglas (Eds.), *Handbook of*

Violence Risk Assessment (pp. 251-268). New York, NY: Taylor & Francis.

Association for the Treatment of Sexual Abusers. (2005) Practice standards and guidelines for

members of the association for the treatment of sexual abusers.

Barbaree, H.E., Langton, C.M. & Peacock, E.J. (2006). The factor structure of static actuarial

items: Its relation to prediction. *Sexual Abuse: A Journal of Research and Treatment, 18,*

207-226.

Bengtson, S. & Langstrom, N. (2007). Unguided clinical and actuarial assessment of re-

offending risk: A direct comparison with sex offenders in Denmark. *Sexual Abuse: A*

Journal of Research and Treatment, 19, 135-153. doi: 10.1177/107906320701900205.

Blumstein, A. (2009). Race and the criminal justice system. *Race and Social Problems, 1,* 183-

186. doi: 10.1007/s12552-009-9022-2.

Brouilette-Alarie, Proulx, J. Helmus, L., & Hanson, R.K. (2011). *Predictive validity of the Static-*

99R dimensions. Paper presented at the Association for Treatment of Sexual Abusers

Annual Conference, Toronto, Canada. Abstract retrieved from

http://www.atsa.com/pdfs/Conf2011/F-30.pdf.

Chapman, J.F., Desai, R.A., Falzer, P.R., Borum, R. (2006). Violence risk and race in a sample

    of youth in juvenile detention. *Youth Violence and Juvenile Justice, 4,* 170-184. doi:

    10.1177/1541204006286316

Claussen-Schultz, A.M., Pearce, M.W., Schopp, R.F. (2004). Dangerous, risk, and capital

    sentencing. *Psychology, Public Policy, and the Law, 10,* 471-491. doi: 10.1037/1076-

    8971.10.4.471

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155-159. doi: 0033-2909/92

Connelly, C. and Williamson, S. (2000). A Review of the Research Literature on Serious

    Violent and Sexual Offenders, Edinburgh: The Scottish Executive Central Research Unit.

Craig, L.A. & Beech, AR. (2010). Towards a guide to best practice in conducting actuarial risk

    assessments with sex offenders., *Aggression and Violent Behavior, 15,* 278-293 doi:

    10.1016/j.avb.2010.01.007.

Criminal History Registry Board (2004). Sex Offender Registry Board: Registration,

    Classification, and Dissemination. (Mass Regulations 803 CMR 1:00).

De Apodaca, R.F., Schultz, J.M., Anderson, A.N., & McLennan, M.D. (2005). Young,

    unassimilated Hispanic offenders: Absolutist vs. relativist cultural assumptions. *Sexuality*

    & Culture: An Interdisciplinary Quarterly, 9, 3-23 doi: 10.1007/s12119-005-1012-y.

Demuth, S. & Steffensmeier, D. (2004). Ethnicity effects on sentence outcomes in large urban

    courts: Comparisons among, white, black, and Latino defendants. *Social Science*

    Quarterly, 85, 994-10011 doi: 10.1111/j.0038-4941.2004.00255.x.

Doren, D.M. (2004) Stability of the interpretative risk percentages for the RRASOR and Static-

    99. *Sexual Abuse: A Journal of Research and Treatment, 16,* 25-36 doi: 1079-

    0632/04/0100-0025/0.

Federal Bureau of Investigation (2003). Age-specific arrest rates and race-specific arrest rates for

      selected offenses, 1993-2001. *Uniform Crime Reports, November 2003*.

Forbes, S.M. (2007). Race differences in scores of actuarial measures of sex offender risk

      assessment. *Proquest Information and Learning Company*. (UMI No. 3293542).

Frazier, P.A. (1993). A comparative study of male and female rape victims seen at a hospital-

      based rape crisis program, *Journal of Interpersonal Violence, 8*, 64-76 doi:

      10.1177/088626093008001005.

Groth-Marnat, G. (2003). *Handbook of psychological assessment.* Hoboken, New Jersey: John

      Wiley & Sons.

Grubin, D. (1998). Sex Offending against children: Understanding the Risk. Police Research

      Series Paper 99. London Home Office.

Hanson, R.K. (1997). *The development of a brief actuarial risk scale for sex offender recidivism.*

      User Report No. 1997-04. Ottawa: Department of the Solicitor General of Canada.

Hanson, R.K. (1998). What do we know about sex offender risk assessment? *Psychology, Public

      Policy, and the Law, 4,* 50-72 doi: 10.1037/1076-8971.4.1-2.50.

Hanson, R. K. (2001). *Age and sexual recidivism: A comparison of rapists and child molesters.*

      User Report 2001-01. Ottawa: Department of the Solicitor General of Canada.

      Department of the Solicitor General of Canada website, www.sgc.gc.ca

Hanson, R. K. (2002). Recidivism and age: Follow-up data on 4,673 sexual offenders. *Journal of

      Interpersonal Violence, 17*, 1046-1062 doi: 10.1177/088626002236659.

Hanson, R.K. (2009). New developments in actuarial scoring. Presentation on June, 18, 2009 in

      Nashua, NH.

Hanson, R.K. & Bussiere, M.T. (1996). *Predictors of sexual offender recidivism: A meta-*

analysis. User Report No. 1996-04. Ottawa: Department of the Solicitor General of

Canada.

Hanson, R.K. & Bussiere, M.T. (1998). Predicting relapse: A meta-analysis of sexual offender

recidivism studies. *Journal of Consulting and Clinical Psychology, 66,* 348-362 doi:

10.1037/0022-006X.66.2.348.

Hanson, R.K. & Harris, A. (1998). *Dynamic predictors of sexual recidivism.* User Report No.

1998-01. Ottawa: Department of the Solicitor General of Canada.

Hanson, R.K., Harris, A.J., Scott, T. & Helmus, L. (2007). Assessing the risk of sexual offenders

on community supervision: The Dynamic Supervision Project. User Report, Corrections

research. Ottawa, ON, Canada: Public Safety Canada.

Hanson, R.K., Helmus, L., & Thornton, D. (2010). Predicting recidivism amongst sexual

offenders: A multi-site study of Static-2002. *Law and Human Behavior, 34,* 198-211 doi:

10.1007/s10979-009-9180-1.

Hanson, R.K. & Thornton, D. (1999). *Static-99: Improving actuarial risk assessments for sex

offenders.* User Report No. 1999-02. Ottawa: Department of the Solicitor General of

Canada.

Hanson, R.K. & Thornton, D. (2000) Improving risk assessments for sex offenders: A

comparison of three actuarial scales. *Law and Human Behavior, 24,* 119-136 doi:

10.1023/A:1005482921333.

Hanson, R. K., & Thornton, D. (2003). *Notes on the development of Static-2002* (User Report

2003-01). Ottawa, Ontario, Canada: Department of the Solicitor General of Canada.

Harris, A., Phenix, A., Hanson, R.K., & Thornton, D. (2003). Static-99 Coding Rules Revised-

2003. Solicitor General of Canada.

Harrison, K. (2007). The high-risk sex offender strategy in England and Wales: Is chemical

     castration an option? *The Howard Journal of Criminal Justice, 46,* 16-31 doi:

     10.1111/j.1468-2311.2007.00451.x.

Heilbrun, A.B., Emory, U., Cross, J.M. (1979). An analysis of rape patterns in White and Black

     rapists. *The Journal of Social Psychology, 108,* 83-87.

Helmus, L., Thornton, D., Hanson, R.K. & Babchishin, K.M. (2011). Assessing the risk of older

     sex offenders: Developing the Static-99R and Static-2002R. (Corrections Research User

     Report 2011-01). Ottawa, ON, Canada: Public Safety.

Hoffer, J., Chasiotis, A., Friedlemeier, W., Busch, H, & Campos, D. (2005). The

     measurement of implicit motives in three cultures: Power and affiliation in Cameroon,

     Costa Rica, and Germany, *Journal of Cross-Cultural Psychology, 36,* 689-716 doi:

     10.1177/0022022105280510.

Jackson, R.L. & Hess, D.T. (2007). Evaluation for civil commitment of sex offenders: A survey

     of experts. *Sexual Abuse: A Journal of Research and Treatment, 19,* 425-448 doi:

     10.1007/s11194-007-9062-3.

Kazdin, A. E. (2003). *Research Design in Clinical Psychology*. Boston: Allyn and Bacon.

Langstrom, N. (2004). Accuracy of actuarial procedures for assessment of sexual offender

     recidivism risk may vary across ethnicity. *Sexual Abuse: A Journal of Research and*

     Treatment, 16, 107-120 doi: 10.1177/107906320401600202.

Leguizamo, A., Carrasco, N., & Peltzman, B. (2008). *Offense, Victim, and Offender*

     Characteristics of Incarcerated Sex Offenders: Comparison of Three Ethnic Groups.

     Unpublished Data

Levenson, J.S. & Prescott, D.S. (2009). Treatment experiences of civilly committed sex

offenders: A consumer satisfaction survey. *Sexual Abuse: A Journal of Research and Treatment, 21,* 6-20 doi: 10.1177/1079063208325205.

Macdonald, P. L., & Gardner R. C. (2000). Type I error rate comparisons of post hoc procedures for chi-square tables. *Educational and Psychological Measurement, 60,* 735-754 doi: 10.1177/0013164400605008.

Maxwell, C.D., Robinson, A.L. & Post, L.A. (2003). The impact of race on the adjudication of sexual assault and other violent crimes. *Journal of Criminal Justice, 31,* 523-538 doi: 10.1016/j.jcrimjus.2003.08.005.

McGrath, R.J., Cumming, G.F., & Burchard, B.L. (2003). *Current practices and trends in sexual abuser management: The Safer Society 2002 Nationwide Survey.* Brandon, VT: Safer Society.

Morey, L. (2007). *Personality Assessment Inventory.* Lutz, FL: Psychological Assessment Resources.

Mossman, D. (2006). Another look at interpreting risk categories. *Sexual Abuse: A Journal of Research and Treatment, 18,* 41-63 doi: 1079-0632/06/0100-0041/1.

Muir, G. & MacLeod, M.D. (2003). The demographic and spatial patterns of recorded rape in a large UK metropolitan area. *Psychology, Crime, & Law, 9,* 345-355 doi: 10.1080/1068316031000093423.

Padilla, A.M. & Borsato, G.N. (2008). Issues in culturally appropriate psychoeducational assessment. In L.A. Suzuki, P.J. Meller, & J.G. Ponterotto (Eds.), *Handbook of multicultural assessment* (3rd pp. 5-21). San Francisco: Jossey-Bass.

Padilla, A.M. & Medina, A. (1996). Cross-cultural sensitivity in assessment using tests in culturally appropriate ways. In L.A. Suzuki, P.J. Meller, & J.G. Ponterotto (Eds.),

Handbook of multicultural assessment (pp. 3-28) San Francisco: Jossey-Bass.

Parent, G., Guay, J., & Knight, R.A. (2011). An assessment of long-term risk of recidivism by

adult sex offenders: One size doesn't fit all. *Criminal Justice and Behavior, 38,* 188-209

doi: 10.1177/0093854810388238.

Roberts, C. F., Doren, D. M., & Thornton, D. (2002). Dimensions associated with assessments of

sex offender recidivism risk. *Criminal Justice and Behavior, 29*, 569-589 doi:

10.1177/009385402236733.

Schwalbe, C.S., Fraser, M.W., Day, S.H., & Arnold, E.M. (2004). North Carolina assessment of

risk (NCAR): Reliability and predictive validity with juvenile offenders. *Journal of*

Offender Rehabilitation, 40, 1-22 doi: 10.1300/J076v40n01-01.

Schwalbe, C.S., Fraser, M.W., Day, S.H., & Cooley, V. (2006). Classifying juvenile offenders

according to risk of recidivism. *Criminal Justice and Behavior, 33,* 305-324 doi:

10.1177/0093854806286451.

Schwalbe, C.S., Fraser, M.W., & Day, S.H. (2007). Predictive validity of the joint risk

assessment matrix with juvenile offenders. *Criminal Justice and Behavior, 34,* 348-361

doi: 10.1177/0093854806292244.

Sjostedt, G. & Langstrom, N. (2001). Actuarial assessment of sex offender recidivism risk: A

cross-validation of the RRASOR and the Static-99 in Sweden. *Law and Human Behavior,*

25, 629-645 doi: 10.1023/A:1012758307983.

Snyder, H.N. (2003). Juvenile arrests 2001. *Juvenile Justice Bulletin,* Office of Juvenile Justice

and Delinquency Prevention, US Department of Justice.

Spaulding, L. (1998). Florida's 1997 chemical castration law: A return to the dark ages. *Florida*

State University Law Review. Retrieved November 5, 2008 from

http://www.law.fsu.edu/journals/lawreview/frames/252/spalfram.html.

Sreenivasan, S., Weinberger, L.E., Frances, A., & Cuswroth-Walker, S. (2010). Alice in actuarial-land: Through the looking glass of changing Static-99 norms. *The Journal of the American Academy of Psychiatry and the Law, 38,* 400-406.

Sudo, J., Sato, M., Obata, S., & Yamagami, Y. (2006). Exploring the possibility of risk assessment of Japanese sexual offenders using Static-99. *Criminal Behavior and Mental Health,* 16, 146-154 doi: 10.10002/cbm.622.

Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research and Treatment, 14*, 139-153 doi: 1079-0632/02/0400-0139/0.

U.S. Department of Justice (2007). *Federal prosecution of child sex exploitation offenders, 2006.* Bureau of Justice Statistics Report NCJ-219412.

U.S. Department of Justice (2003). *Recidivism of sex offenders released from prison in 1994*. Bureau of Justice Statistics Report NCJ-198281.

U.S. Department of Justice (1997). *Criminal victimization in the United States,* 1994. Bureau of Justice Statistics Report NCJ-162126.

U.S. Department of Justice (1997). *Sex offenses and offenders.* Bureau of Justice Statistics Report NCJ-163392.

Van de Vijver, F. J. R. & Poortinga, Y.H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment, 13*, 29-37 doi: 10.1027/1015-5759.13.1.29.

Van de Vijver, F.(2000). The nature of bias. In. R.H. Dana (Ed.) *Handbook of Cross-Cultural and Multicultural Personality Assessment* (pp. 87-106). Lawrence Erlbaum Associates, Publishers: Mahwah, New Jersey.

Varela, J.G., Boccaccini, M.T., Gonzalez, E., Murrie, D.C., & Caperton, J. (2011) Do the Static-

99 and Static-99R perform similarly for White, Black, and Latino offenders? Presentation

at Annual Psychology and Law Conference.

Walsh, Z., Swogger, M.T., & Kosson, D.S. (2004). Psychopathy, IQ and violence in European

American and African American county jail inmates. *Journal of Consulting and Clinical

Psychology, 72,* 1165-1169 doi: 10.1037/0022-006X.72.6.1165.

West, J., & Templer, D.I. (1994). Child molestation, rape, and ethnicity. *Psychological Reports,

75,* 1326.

Wheeler, J. & George, W.H. (2005). Race and sexual offending. In Race, culture, psychology,

and law. Barrett, K.H. (Ed.); George, W.H. (Ed.) pp. 391-402. Thousand Oaks, CA: Sage

Publications Inc.

Wolak, J., Mitchell, K., & Finkelhor, D. (2003). *Internet sex crimes against minors: The

response of law enforcement*. National Center for Missing and Exploited Children.

Wood, J. (2006). Profiling high-risk offenders: A review of 136 cases. *The Howard Journal of

Criminal Justice, 45,* 307-320 doi: 10.1111/j.1468-2311.2006.00423.x.

Zott, L.M. (2008). *Sex Offenders and Public Policy.* Detroit: Greenhaven Press.

Appendix
Static-99 Coding Form

Question Number	Risk Factor	Codes	Score
1	Young	Aged 25 or older Aged 18 – 24.99	0 1
2	Ever Lived With	Ever lived with lover for at least two years? Yes No	0 1
3	Index nonsexual violence Any Convictions	No Yes	0 1
4	Prior nonsexual violence Any Convictions	No Yes	0 1
5	Prior Sex Offences	Charges Convictions None None 6+ 4+	0 1 2 3
6	Prior sentencing dates (excluding index)	3 or less 4 or more	0 1
7	Any convictions for non-contact sex offences	No Yes	0 1
8	Any Unrelated Victims	No Yes	0 1
9	Any Stranger Victims	No Yes	0 1
10	Any Male Victims	No Yes	0 1
	Total Score	**Add up scores from individual risk factors**	

Score   Label for Risk Category
  0,1   Low
  2,3   Moderate-Low
  4,5   Moderate-High
  6 plus   High