Running head: ASSESSMENT OF SEX AND AGGRESSION

Validation and Revision of the Multidimensional Assessment of Sex and Aggression

Raymond A. Knight & David D. Cerce

Brandeis University

FINAL REPORT Archive

Approved By: A. Howes

Date: 10/17/01

Abstract

The Multidimensional Assessment of Sex and Aggression (the MASA) was initially created to supplement the often poorly represented information in the archival records of sex offenders and to provide sufficient data to classify adult sex offenders. It has now been revised four times, expanding the breadth of its assessment, simplifying its language to make it appropriate for juveniles, and computerizing its administration. This article summarizes some of the recent reliability and validity analyses that have been calculated on a wide variety of samples including college students, community non-criminals, non-sex offending criminals, and adult and juvenile sex offenders. Continued reliability and cross-sample stability of factor structures and the intercorrelations across its scales suggest that the inventory shows promise as a useful assessment instrument for sex offenders.

Validation and Revision of the Multidimensional Assessment of Sex and Aggression

The high prevalence of sexual aggression (Koss, Woodruff, & Koss, 1991; Wyatt, 1992) and the seriousness of the consequences of such aggression to its victims (Burgess & Holmstrom, 1974; Hanson, 1990) are well documented. The widespread concern of society about sexual aggression is reflected in the numerous legislative initiatives that have been directed at reducing its incidence. Recent legislation has created sexual predator laws, required community notification about high risk offenders, and mandated the treatment of offenders (Grubin & Prentky, 1993; Prentky, 1996). Adequate implementation of such directives requires a solid foundation in the assessment and prediction of sexual aggression. The goals of enhancing our ability to identify potential sex offenders and to assess the risk of re-offending of known offenders, of making adequate dispositional decisions about convicted offenders, of identifying and treating the specific deficits of sex offenders, of evaluating the efficacy of intervention techniques, and of predicting recidivism all depend on the adequacy of our theoretical models and our ability to assess the critical domains of sexual aggression reliably and validly.

If we are going to progress toward achieving these goals, we need to identify or develop a standardized, assessment instrument for sex offenders that can guide adjudication, serve the function of pre- and post-treatment evaluation, and provide the data for valid risk assessments and predictions of recidivism. It is critical that that this assessment inventory or battery not only be reliable and valid and have adequate standardized norms, but also that it be comprehensive and easily administered and

processed, so that it will be widely used. The generation of a model of sexual aggression that will be useful for guiding decisions requires the analysis of multiple domains, measured on sufficiently large samples. The extensive use of an efficient, reliable, valid, standardized multivariate assessment tool would contribute substantially to the establishment of data bases required to generate such a model.

The option of adapting existing, well standardized inventories to serve this evaluation need has not proven viable. Although the major self-report inventories that are currently available (e.g., MMPI-2, MCMI, MSI) are appropriate for and helpful with the assessments for which they were created, they are suboptimal for the assessment of sex offenders, because they do not assess the most critical sexual aggression domains (see Knight, Rosenberg, & Schneider, 1985; Prentky & Knight, 1991). Each of these major assessment instruments has its own distinct advantages, but is also burdened with significant disadvantages that compromise its utility for this purpose.

The MMPI, which is the most frequently studied psychometric instrument for sex offenders and other criminal populations (see Gearing, 1979; Knight et al., 1985), provides a rich source of empirically validated data for comparative purposes and contains the most extensively researched scales for faking good and bad. Unfortunately, large numbers of sex offenders produce profiles that match non-offender groups (Marshall & Hall, 1995), and within clearly defined subgroups of sex offenders MMPI profiles have been characterized more by their heterogeneity than by their similarities (Erickson, Luxenberg, Walbek, & Seely, 1987; Hall, Maiuro, Vitaliano, & Proctor, 1986; Marshall & Hall, 1995). Moreover, cluster analytic studies of sex offenders have yielded inconsistent results across studies (Anderson, Kunce, & Rich,

1979; Kalichman, Szymanowski, McKee, Taylor, & Craig, 1989; Schlank, 1995; Shealy, Kalichman, Henderson, Szymanowski, & McKee, 1991), and when types have emerged, they have often pooled subjects with quite different offense histories (Marshall & Hall, 1995).

The MCMI has fared somewhat better than the MMPI with sex offenders. It is sensitive to the antisocial and narcissistic features that are so prevalent in correctional settings (McNeil & Meyer, 1990), and it has scales that show the same stability for sex offenders as other deviant populations (Langevin et al., 1988). It has yielded some meaningful cluster groupings of sexual offenders (Bard & Knight, 1986), and the factor structure of the responses of an inmate sample approximated those found in clinical populations (Langevin et al., 1988). Moreover, some differences among both adult and juvenile sex offender groups on the MCMI parallel the differences found in previous diagnostic literature (Carpenter, Peed, & Eastman, 1995; Chantry & Craig, 1994). Regretfully, only a handful of studies using the MCMI with sexual offenders have been carried out, and it shares with the MMPI and the MSI the general difficulties that we discuss below.

The MSI has the distinct advantage of providing information about sexual behavior and pathology not tapped by traditional psychological tests. The scales of the MSI have shown fair internal consistencies in independent assessments of the instrument (Kalichman, Henderson, Shealy, & Dwyer, 1992), and its test-retest reliabilities range from .64 to .92 over a 21-day period (Nichols & Molinder, 1984). Importantly, some of its scales have been related to improvement in sex offender treatment programs (Minor, Marques, Day, & Nelson, 1990; Simkins, Ward, Bowman, &

Rinck, 1989). Unfortunately, despite its widespread use, relatively little empirical work has been done to assess its validity. Moreover, although there is some evidence of correlations with greater pathology on MMPI scales, these correlations have not been found to be consistent across samples (Kalichman et al., 1992). The transparency of its items and potential response bias contamination are also potential drawbacks. The Rape, Sexual Obsessions, and Cognitive Distortions and Immaturity scales all have negative relations with the Marlowe-Crowne (Kalichman et al., 1992). In a sample of child molesters many of the MSI scales correlated substantially with the MMPI F and K scales (Rape, Exhibitionism, Sexual Obsessions, Premature Ejaculation, Cognitive Distortions and Immaturity, and Justifications). In a cluster analytic studies of sex offenders (Kalichman et al., 1989; Schlank, 1995) only the Paraphilias Scale has shown cross-study discriminatory power. In addition, none of the MSI sexual dysfunction subscales were related to sex drive or sexual fantasy on the Derogatis Sexual Functioning Inventory (Kalichman et al. 1992), even though these have been found to be important components of sexual aggression (Knight, 1995).

The most telling problems with these three inventories, however, are three shared difficulties that make them all questionable instruments for evaluating sex offenders. First, no data exist on the usefulness of any of their scales for making dispositional decisions at any level for sex offenders, and no data on the most appropriate scale cutoffs or on the hit rates for various dispositional decisions are evident. Second, none of the inventories sufficiently sample all the domains that have been found critical in assessing sexual aggression (see Knight et al., 1985; Prentky & Knight, 1991). Third, most studies on these inventories have assessed offenders who

admit their guilt. Substantial evidence indicates deniers differ significantly from admitters, report little psychopathology, and present themselves in a favorable light (Langevin, 1988; Lanyon & Lutz, 1984). An adequate assessment tool for sex offenders must provide better solutions to the duplicity problem.

These problems explain the disenchantment with such instruments that has arisen among practitioners and researchers (e.g., Marshall & Hall, 1995). In our own research program, it also became clear to us that if we were going to integrate adequately the role of sexual behavior, cognitions, and fantasy and offense planning into the taxonomic systems we were developing for sex offenders, we needed to develop a self-report inventory that met these assessment needs. Consequently, we created the Multidimensional Assessment of Sex and Aggression, the MASA, a self-report inventory that assesses all domains necessary for classification in our taxonomic systems (see Knight, Prentky, & Cerce, 1994). It is the purpose of this article to summarize some of the recent data on the development and validation of this instrument.

Method

History and Design of the MASA

We have previously described in detail our methodology for constructing the original MASA (Knight et al., 1994). In brief, it involved the specification of multiple domains that our research had shown important in the assessment of sexual aggression, the creation of an extensive item pool covering all these domains, the rating by experienced clinicians of the appropriateness of items for each domain, the selection of the most suitable items for each domain, the rewriting of the chosen items to

maximize their relevance to the domains, the assessment of domain coverage, the creation of supplemental items for areas that were not adequately represented, and finally the preliminary testing of the original version of the MASA on 127 Massachusetts Treatment Center (MTC) sex offenders and the re-administration of the MASA to 35 of these offenders to assess reliability.

This first version of the MASA, which focused more exclusively on adult rapists, assessed social competence, juvenile and adult antisocial behavior, anger and anger management, expressive aggression, sadism, sexual deviance and paraphilias, sexual preoccupation and compulsivity, offense planning, hostility toward women, and pornography use. These are the domains most critical for classification in our rapist typology, MTC Rapist Typology, Version 3 (MTC:R3). Since testing the original version, we have revised the MASA four times and retested it on generalization samples.

In the first revision of the MASA, which was completed seven years ago, we incorporated the assessment of additional domains that our research program had identified as important for evaluating sex offenders. For instance, items were added that assess those developmental antecedents that we have found (a) to discriminate sexually coercive non-criminals from non-sexually coercive males (Knight, 1993), (b) to predict criminal recidivism (Knight, 1999), (c) to discriminate early-onset from late-onset sex offenders (Knight & Prentky, 1993), and (d) to be correlated with the amount of injury done to victims during sexual assaults (Prentky, Knight, Sims-Knight, Straus, Rokous, & Cerce, l989). Moreover, in this first revision an attempt to evaluate components of Hare's Psychopathy Checklist (Hare, 1980; Hare, Harpur, Hakstian, Forth, Hart, & Newman, 1990) was introduced, as well as improved lie and fake good

scales.

In the second revision of the MASA we simplified the language and made it suitable for juveniles by incorporating alternative age-appropriate questions both on social competence and on sexual attitudes, behavior, cognitions, and fantasies. In revising the core of the inventory we once again item analyzed scales to assure the highest internal consistency. Our analyses from both the first and second revisions of the MASA indicated that the scales that we had introduced to assess response bias required more work. Response biases, or responding to a range of questionnaire items on some basis other than the specific item content, plague all of psychometric assessment (Paulhus, 1991), but especially the assessment of sex offenders, who present problems not regularly associated with other patient or criminal populations (Marshall & Hall, 1995). Some of these problems stem from the fact that these offenders must talk openly about sexual behavior, a topic that engenders much anxiety and discomfort in our society. More importantly, the offender is asked to admit to behaviors that are not only socially unacceptable, but are, in fact, illegal. To date we have applied the control technique of demand reduction, by guaranteeing subjects anonymity. From the success of our reliability, internal consistency, factor analytic, and cross-group consistency analyses reported below, this has proved to be a successful strategy. If the MASA is, however, to have practical utility, it must incorporate assessments of various response biases, so that their presence can be evaluated and taken into account in situations in which anonymity cannot be guaranteed. This was a major focus of our third and fourth revisions, which we will describe briefly in the discussion section.

Participants

Participants in the original MTC sample. The participants in the validation study of Version 1 of the MASA were 127 incarcerated sex offenders, who had been civilly committed at the time of data collection (1990-1991) to the MTC in Bridgewater. This sample included approximately 60% of the committed residents of the MTC during the period of data collection and constituted a good representation of the entire population. It comprised repetitive offenders against adult women and children. For the purposes of this report, the term rapist (n = 59) refers to an adult male whose sexual offenses were committed exclusively against adult women (i.e., 16 years of age or older). A child molester was defined as someone who had at least one sexual offense against a victim under the age of 16 (n = 68).

Participants in the generalization samples validating Versions 2 and 3. Two waves of generalization samples have been tested on the MASA. In the first wave, using a paper-and pencil version of the first revision of the MASA, we tested various groups of subjects from prison facilities in Pennsylvania, Vermont, and Ontario, Canada, and we tested non-criminal controls in the USA and Canada. These groups included 127 college students, 60 unemployed, non-criminal community controls from Ontario and 33 non-criminal prison employees from multiple sites, 162 non-sexual offending criminals, 95 incarcerated rapists, and 45 sex offenders with extra-familial child victims. In the second wave, using both computer and paper-and-pencil versions of revision 2 of the MASA, we tested 578 criminals and non-criminals in New Jersey, Minnesota, Virginia, and Massachusetts. In this article we will report the results of 406 of these subjects, 131 juvenile sex offenders from Virginia and Minnesota, and 275

adult sex offenders from New Jersey and Minnesota prisons. All juvenile offenders had been charged with and were being treated for at least one sexual crime involving sexual contact with a victim. All adult sex offenders had been convicted of at least one sexual crime involving sexual contact with a victim, and were currently incarcerated for such a crime.

Procedure

Subject selection for all testing involved a simple two-step process. Potential volunteers were identified and approached by on-site personnel. In some institutions this involved advertising and in others possible participants were contacted through program personnel. During the second wave of testing, parental or legal guardian permission had to be obtained for juveniles before the testing team came on site. When the testing team arrived in the institution to administer the test, interested participants were convened in groups of 7 to 12 subjects. They were informed in more detail about the nature of the study, about the kind of material they would be asked to answer, about the protection of confidentiality they were guaranteed and about the Writ of Confidentiality we had been awarded from NIMH, and about the fee they would be paid for their participation ($18.00). A strong plea was made for honesty, and the potential future benefits of adequate assessment for offenders like themselves was stressed. After informed consent statements had been explained by a visiting research team member and signed both by the offender and the research team representative, either the paper-and-pencil or the computer versions of the MASA were distributed and a standard set of instructions was given. If offenders had difficulty reading the inventory, arrangements were made for one of our team to read the inventory to him.

Results

Overview

The intent of the present article is to present a representative sampling of the analyses that we have done on the first three versions of the MASA and to give an overview of the MASA's reliability and validity. Toward that end we will first summarize the already reported reliability and validity analyses of the first version of the MASA calculated on the original MTC sample, give a sampling of our factor analyses of that version, and summarize the relations among the factor domains. Second, using three samples from the first generalization study of the MASA (Version 2), we will demonstrate the consistency of the individual factor scales across these new samples and using a slightly different breakdown of the samples in the first generalization study, we will summarize the congruence of the pattern of correlations among factor domains across both the original sample and these new samples. Third, using samples of both juvenile and adult sex offenders from our second generalization study of the MASA (Version 3), we will provide evidence for the utility of the MASA for juveniles sex offenders by showing their comparable internal consistencies and test-retest reliabilities on the factor scales to those of adult sex offenders. We will also map the pattern of the relations among the factor domains for the juveniles onto the patterns of relations we had reported for the adults.

Analyses of the Original Sample

Reliability and validity analyses. For the original sample we calculated the internal consistencies and test-retest reliabilities for a set of rational scales that had been designed to measure the critical domains for classification in MTC:R3. The high coefficient alphas for all these scales (94% greater than .70) and high test-retest

reliabilities (only two scales--Vandalism in Adulthood and Impulsivity in the Offense yielding reliabilities < .70) indicated that reasonable reliability had been achieved (Knight et al., 1994).

For these rational scales we also reported concurrent validity coefficients, derived by correlating each scale with a parallel, independent assessment of the same domain, which was created by rating the information provided in the participants' archival records. These analyses indicated that only the domains of sexualization, sexual aggression, and sexual offense planning failed to show adequate concurrent validity coefficients. A comparison between offenders' answers to the MASA scales for these domains and the information garnered from their archival files indicated that far more sexual preoccupation, deviance, compulsiveness, inadequacy, and sadistic fantasies and behaviors were reported on the MASA than were evident in the archival files, suggesting that the MASA provided greater validity and coverage of the relevant information than the criminal and clinical files.

Factor analyses of the MASA: The example of offense planning. Although the general rational scales had reasonably high internal consistencies, suggesting that the items in each scale assessed the same general construct, we wanted to explore the factor structure within each domain. Consequently, for each of ten separate domains, which are listed in the left hand column of Table 2, we calculated principal components analyses (Hartman, 1967), extracted all factors with eigenvalues greater than 1, and rotated these factors to VARIMAX criteria. We examined these preliminary factors to determine the number of core factors that were cohesive and theoretically meaningful. We then recalculated the principle components analyses, limiting the extraction to the

number of core factors. We will present the analysis of the Offense Planning domain to illustrate the outcome of this procedure for one domain.

In the original offense planning principal components analysis ten factors were extracted. An examination of this solution revealed that the last four factors were either single item or weak, splinter factors. Consequently, a six factor solution was specified, and this solution, which accounted for 65% of the variance is presented in Table 1. In the clinical and criminal literature offense planning is often discussed as a univocal construct (Rosenberg & Knight, 1988). The factor analysis of the offense planning items challenges this notion and suggests that offense planning is a multidimensional construct, comprising the six relatively independent factors of this analysis.

A brief consideration of these factors reveals that each represents a theoretically meaningful and separable component of offense planning and pre-offense fantasy. The first factor closely approximates a construct that has often been described in the clinical literature. It involves fantasies that Cohen, Garofalo, Boucher, and Seghorn (1971) attributed to their compensatory rapist type and Groth, Burgess, and Holmstrom (l977) saw as characteristic of their similarly defined power-reassurance rapist. Hazelwood (1987) has referred to these as pseudo-unselfish fantasies, and Marshall (1989) discussed them in the context of seeking intimacy, which is the descriptor we have chosen. In these fantasies the rapist ignores the agonistic nature of the sexual assault and fantasies that his sexual overtures will elicit a positive response in the victim. The second factor, Aggressive/Violent Fantasies, taps the offender's fantasies about physically harming, frightening, and even killing the victim. The third factor, Planning the Offense: Victim Type and Crime Location, captures the offender's forethought in

seeking a particular victim and fantasies about a particular location for an assault. The fourth, Sexual Fantasies, includes the offender's fantasies about what sexual acts he would perform, or would have the victim do to or for him. The fifth, Eluding Apprehension, taps his plans to elude apprehension after the crime. The sixth and final factor, Planning the Offense: Weapons and Paraphernalia, focuses on the weapons and paraphernalia (his "rape kit") that he planned to take with him for his crimes.

The relatively low intercorrelation among these offense planning characteristics suggests that they are ripe for cluster analyses that might be informative either from a criminal investigative analysis perspective (Knight, Warren, Reboussin, & Soley, 1998) or from a psychotherapeutic vantage (Pithers, 1990). If distinct clusters of planning could be identified and replicated, these could also have important implications for crime scene analyses and for structuring relapse prevention interventions. The greater differentiation of offense planning into distinct components of planning also provides potential resolutions to the problems we encountered with the clearly inaccurate global representation of this construct in MTC:R3 (Knight, 1999).

Correlations among the factor domains in the original sample. We generated factor scales for each of the factors in the ten domains by standardizing each item and averaging over all items that loaded >.40 on each factor. To analyze the relations between the various components in the MASA, we correlated the factor scores of each domain with the factor scores of the other domains for the 59 rapists who had taken the original version of the MASA. Table 2 presents a summary of the intercorrelations among the factor domains. The number of factors in each domain is presented in parentheses below the factor domain name in the left column of the table. Above the

diagonal is the average correlation among the factors for the two domains. Below the diagonal is the percent of the correlations between the factors in the two domains that reached .01 significance. Thus, the average correlation gives some notion of the overall general level of relation between the two domains, and the percent gives an indication of the pervasiveness of the relationship across the factors in the two domains.

There are several important relationships in this table that should be noted.

(1) Social competence was completely independent of all the other domains. None of the correlations of its two factor scales, Relationships and Independence, reached .01 significance with any other factor scale.

(2) As expected, juvenile and adult antisocial behavior were highly related.

(3) Antisocial behavior was moderately related to expressive aggression, but relatively independent of other factors.

(4) As expected, pervasive anger, expressive aggression, and sadism were all interrelated. Both sadism and pervasive anger were strongly related to expressive aggression, and expressive aggression was strongly related to pervasive anger, but sadism and pervasive anger were only weakly related.

(5) Sexual drive was strongly related to pervasive anger, expressive aggression, sadism, and offense planning.

(6) Pornography use was strongly related to both sexual drive and expressive aggression.

(7) The paraphilias were moderately related to expressive aggression, sexual drive, and offense planning.

First Generalization Study Analyses

Consistencies of the factor scales across samples. The first revision of the MASA was given to a variety of samples including sex offenders in general prison settings (n = 140), non-sex offending criminals (n = 162), and non-criminals (n = 220). We repeated the principal components analyses described above on each of these groups and were able to replicate most of the factor structures of the original sample. Table 3 presents the average Cronbach alphas for these same factor scales on the three new replication samples. For instance, for the sample in the first column, who are the sex offenders in the first replication sample, and consequently the sample that most closely approximated the 127 offenders in the original MTC sample, the .87 represents the average of the alphas for this group for the Independence (α = .81) and Relationships (α = .92) factor scales in the social competence domain. An examination of the α's in this table reveals that the scales developed on the original sample also cohere in the replication samples. The rare exceptions were predominantly in the normal sample (e.g., expressive aggression, sadism, paraphilias, and offense planning), where the particular behaviors were infrequent. Indeed, very few non-criminals answered any of the items in the offense planning section.

Consistencies of the relations between factor domains across adult samples. Figure 1 presents a summary of the consistencies in the correlations among the factor domains for five groups of adult subjects selected from in the original MTC sample and from the samples in the first generalization study: the original MTC rapists (n = 59), college students (n = 127), community normals (n = 93), non-sexually offending criminals (n = 162), and rapists in prison (n = 95). In these consistency analyses we focused only on rapists and not child molesters, because the MASA was originally

developed for rapists, and only incorporated more extensive assessment of child molesters in Versions 3, 4, and 5. The numbers in the figure refer to the number of different groups (up to 5) who reached our criterion for a significant relation between each pair of factor domains. Thus, both blanks (zero groups reached criterion) and 5's indicate perfect agreement across the five groups. Because the offense planning questions were only given consistently to the two sex offender groups, a dot was used to indicate perfect cross-group agreement of the two groups. Note in the legend of this figure that different levels of shading from dark (5 group agreement) to no shading (0, 1, or 2 group agreement) also correspond to the different levels of group agreement. Perfect two group consistency on offense planning is indicated by light diagonal lines. These same shading designations will be used in Figure 2 to provide a comparative backdrop of adult group agreement for the pattern of correlations of the juveniles.

The diagonal divides the matrix into the same two different assessments of relation that we employed in Table 2. Above the diagonal the criterion is based on the average correlation between factor scales in the two domains, and below the diagonal the criterion refers to the percent of correlations across the two domains that reached significance ($p < .01$). To meet the average correlation criterion above the diagonal, the average correlation among the factor scales across a factor domain pair must have reached at least .01 significance (e.g., as indicated in Figure 1 for the juvenile antisocial domain [6 factor scales] and the adult antisocial domain [4 factor scales], the average of the 24 correlations between all factor scales was significantly different from 0 at $p < .01$ for all five groups, and is consequently marked with dark shading and "5") . The numbers in the figure below the diagonal refer to the number of groups for whom the

percent of correlations (e.g., out of 24 in the above example) that reached .01 significance across the factor scales in the two domains was greater than 40%. The average correlation gives some indication of the overall level of relation between the two domains, and the percent of significant correlations reflects the pervasiveness of the relation across the factor scales in the two domains. It is noteworthy that the pattern of relations among antisocial behavior, sexual drive, violence, and offense planning that were found in the original sample of rapists was replicated across all groups tested here, including the non-criminals.

Second Generalization Study Analyses: The Utility of the MASA for Juveniles

The second revision of the MASA (Version 3) focused on two goals--creating a computerized form of the MASA and making the language and content appropriate for juveniles. Version 3 of the MASA was administered to 131 juvenile sex offenders from five different inpatient facilities in Minnesota and Virginia and to 275 adult sex offenders from two prisons and one treatment center in Minnesota and from one treatment center in New Jersey. Of the 131 juveniles, 121 were administered the computer form, 81 took the paper-and-pencil form, and 71 took both forms. Of the 275 adult sex offenders, 50 adults took both forms of the MASA. In analyzing the internal consistencies for the juveniles, five juveniles were dropped because their protocols indicated random responding in part of the inventory.

Reliability of the MASA for juvenile sex offenders. In addressing the question of whether Revision 3 of the MASA was an appropriate assessment tool for juveniles, we examined the internal consistencies and test-retest reliabilities of the factor scales on juvenile sex offenders we tested in our second generalization study. Table 4 presents

the average internal consistencies and test-retest reliabilities for the ten factor domains for both the juvenile and adult samples in the second generalization study. The internal consistencies averages are based on 38 of the 41 original factor scales. Two of the factors comprised a single variable, precluding calculation of internal consistencies. In addition, we dropped the fifth paraphilia factor, zoophilia/necrophilia, because it did not cohere in the prior analyses of the first generalization study. The test-retest reliabilities are based on all 41 factor scales.

As can be seen in Table 4, the internal consistencies on these factors were consistently high. For the juvenile sex offenders approximately 90% of the factor scale α's were greater than .70, and 67% were greater than .80. For the adults 92% exceeded .70, and 84% were greater than .80. Although the internal consistencies of the juveniles were slightly lower than those of the adults, they were still high and clearly support the use of these factor scales for juveniles. Both the juveniles and the adults showed poor consistency on the items in the Sexual Behavior factor in the sexual drive domain, $\underline{r}$ = .21 and .20, respectively, and both showed suboptimal consistency on the Voyeurism factor in the paraphilia domain, $\underline{r}$ = .59 and .59, respectively. Although Sexual Behavior (i.e., the frequency of sexual activity) factor did not cohere, both the Sexual Preoccupation and Sexual Compulsion factors, the other factor scales in the sexual drive domain, did evidence high internal consistency for both juveniles and adults. The average of the low Sexual Behavior factors with the high Preoccupation and Compulsion factors yielded the averages presented in Table 4 for the adults ($\underline{r}$ = .66) and juveniles ($\underline{r}$ = .62) for the sexual drive domain. In addition to these common problems in consistency, the juveniles showed inferior consistency on factor scales that

were clearly less appropriate for this young sample: Independence in the social competence domain and Drugs and Vandalism in the adult antisocial domain. In the computer form, which used the subject's age as a criterion for which questions were asked, a large proportion of the juvenile sample were not asked the adult antisocial items because they were too young (i.e., less than 17 years old). Moreover, those who had reached their seventeenth birthday and were asked the adult antisocial items, often had their responses to these items affected by their incarceration during their brief adult life.

The average correlations for each domain between the factor scales for the written and computer administrations of the MASA are also presented in Table 4. The correlations are sufficiently high that they can be considered reasonable assessments of test-retest reliability. For the juveniles 93% of the reliabilities were greater than .60 and approximately 80% exceeded .70. For the adults 98% were greater than .60, and approximately 90% exceeded .70. The test-retest reliabilities of the juveniles, although slightly lower than those of the adults, were nonetheless high and support the use of these same factors for juveniles. It is noteworthy that the factors with low test-retest reliabilities for the juveniles were exclusively in those domains that could be considered less appropriate for a juvenile sample—social competence and adult antisocial. Both juveniles and adults showed high test-retest correlations on both the Voyeurism and Sexual Behavior factors, despite the low internal consistencies found for these factor scales.

Comparison of the relation among factor domains for juveniles and adults. Figure 2 depicts the comparison of the pattern of correlations across factor domains that the

juveniles produced with the patterns of the previous adult samples. To illustrate the relation of the juveniles to the previous adult samples, we deleted the numbers in the body of Figure 1, which represented the number of adult groups that were in agreement for a particular cell, and we retained only the shading to indicate this agreement. We then mapped the agreement between factor domains for the juveniles onto the relations produced by the adult groups in the first generalization study, using the # and ** symbols. Because the adult antisocial factor scales were inappropriate for and not answered by a large number of juveniles, we deleted that factor domain, but maintained the same domain numbering as in Figure 1. Graduated shading indicates 5 group, 4 group, or 3 group criterion achievement, as we described earlier for Figure 1. The offense planning domain is shaded somewhat differently, because of the smaller number of groups (two) previously given this section (light diagonal lines indicating two group agreement for the specific factor domain pair).

Figure 2 presents a mapping of the pattern of correlations of the juveniles onto the shaded patterns of the previous groups. Above the diagonal two asterisks (**) indicate that the juveniles (n=131) reached the criterion of agreement for the average correlation, and below the diagonal the pound sign (#) indicates that juveniles reached the criterion for the percent of correlations that reached $\underline{p} < .01$ significance. The figure clearly indicates that there was considerable agreement between the juveniles and the previous adult samples in their patterns of correlations between domains. Of the 72 cross-domain cells in the 9 X 9 matrix, 38 had had perfect agreement (either 0 or 5 groups; see Figure 1) among the five previous adult groups. In only two of these cells were the juveniles completely discrepant with the other groups in their pattern of

responding. In all other relations at least two or three adults groups showed a relation between the domains (e.g., the pervasive anger and juvenile antisocial behavior correlations showed 2 adult group agreement for the average correlation and 3 adult group agreement for the percent of correlations greater than 40%). It is interesting that the only two completely discrepant cells involved the relation between juvenile antisocial behavior and social competence, where there was a relation for the juveniles, but not for any of the adult groups. A finding in our retrospective study sheds some light on this disagreement (Knight & Prentky, 1993). Juvenile antisocial behavior and social competence were found to be the major distinguishing characteristics between offenders charged with sexual offenses as juveniles and those who had no sexually coercive behavior until adulthood. This suggests that these two characteristics might be related to being apprehended for sexual coercion as a juvenile, and thus might produce the correlation in apprehended juvenile samples that we see here. In general, the pattern of correlations strongly supports the comparability of juvenile sex offenders' patterns of responding to other groups' patterns across these factor domains.

Table 5 presents two other ways to summarize these results. The first, the left numbers without parentheses in each box, used the values of the average correlation between two domains for each group as data points and correlated each group with every other group on these average correlations. The second, the right numbers in parentheses in each box, used the percentage of correlations that were significant between two domains for each group as data points and again calculated correlations between pairs of groups on these percentages. For both ways of evaluating cross-group compatibility of domain relations, the juvenile showed both high correlations with

all adult groups that paralleled the relations among the adult groups. This consistency with adult groups suggests again that the juveniles' pattern of responding on the MASA was comparable to that of the other groups and indicates that the MASA can serve as an appropriate assessment tool for juveniles.

Discussion

As is evident in the selective data presented in this article, the MASA shows great promise for fulfilling the much needed role of a comprehensive assessment tool for both juvenile and adult offenders, and indeed for both criminal and non-criminal samples. In the original sample the rational scales, which had been created to measure specific theoretical domains found important in the classification of rapists, demonstrated high internal consistency and reasonable cross-temporal stability. Moreover, these scales either correlated highly with companion scales that had been rated using archival records, or evidenced considerably more frequent admission of sexual deviance, violence, and sadism than was recorded in the archival files. This suggests that the scales were likely to have been tapping true variance and to have captured information not recorded in the archival sources.

The factor analyses calculated on the original MTC sample have not only been informative about potentially important cohesive constructs in the select MTC sample, but they also provided structures that we found to cohere across multiple, radically different samples. The factor analysis of the offense planning items, presented here as a representative example of the results of these analyses, illustrates the potential insights these factor analyses have yielded for assessment, taxonomic structuring, criminal investigative analyses, and therapeutic intervention.

The factor structures generated on the MTC sample were, with some exceptions substantially recoverable across generalization samples, and the factor scales derived from these analyses had high internal consistencies and test-retest reliabilities in both adult and juvenile samples. Moreover, the pattern of correlations among the factor domains was found to be consistent across six samples that included sex offenders and non-sexual offenders, criminals and non-criminals, and adults and juveniles. The consistency of the pattern of correlations between domains across samples has provided substantial support for the possibility of creating a unified theory of sexual aggression that encompasses diverse populations. Across all samples antisocial behavior emerged as essentially independent of sex drive, expressive aggression, and offense planning. In contrast, sexual drive was invariably related to expressive aggression, and sadism, and it appeared to be a major driving force of offense planning. On the basis of the consistency of this pattern and also the similarly congruent patterns in data on developmental antecedents of the components of sexual aggression in these same samples, Knight (1997) has proposed a tripartite developmental model of sexual aggression that combines: (a) sex drive/promiscuity, which is frequently predicted by early sexual abuse and sometimes indirectly predicted by physical abuse; (b) predatory personality/emotional detachment, which is preceded by physical abuse and appears to play an equivalent role across criminal and non-criminal samples, and (c) impulsivity or antisocial behavior, which covaries strongly with alcohol and drug abuse, and is the major discriminator between criminal and non-criminal samples. The model integrates these data with the work of Malamuth (in press) and with recent developments in the study of the components of psychopathy (Patrick &

Zempolich, in press) and personality disorder (Berenbaum, 1995) in an attempt to interweave biological/genetic, developmental/experiential, and societal/attitudinal etiological factors.

Although the MASA has already proven a rich source of numerous insights into various components of sexual aggression, it will never be able to serve the role of a viable assessment tool for sexual aggression without solving the critical problem of duplicity. As we indicated in the introduction, the issues of denial and lying are especially problematic for sex offenders. To date we have addressed this response bias problem by applying the control technique of demand reduction by promising and guaranteeing anonymity to all participants in our studies. From the success of our reliability, internal consistency, factor analytic, and cross-group consistency analyses, this has proven to be a successful strategy.

If the MASA is, however, to have practical utility, it must incorporate assessments of various response biases, so that their presence can be evaluated and taken into account in situations in which anonymity cannot be guaranteed. In the early versions of the MASA our use of an abbreviated Marlowe-Crowne scale (Saunders, 1991) and a newly created Sexual Behavior Lie scale, which asked about common sexual behaviors (e.g., masturbation, viewing sexual materials) that non-defensive respondents should admit, but defensive respondents might deny, proved suboptimal. In our most recent revision to the MASA we have decided to make a concerted effort to address this problem, implementing three independent response bias assessment strategies and testing their validity under varying conditions that should increase or decrease duplicity (giving feedback to clinical staff about test results versus

guaranteeing the anonymity of responses, using instructional sets to fake good or fake bad versus a set to reply honestly, testing inmates at various stages in their treatment history, and comparing offenders whose responses closely match key information from their criminal files to those for whom there is a significant discrepancy).

The first strategy we are implementing attacks the problem from the traditional content perspective, which although not successful as a single strategy for sex offenders, might contribute to the detection of bias, when combined with the other strategies. Moreover, there have been some recent developments in the content approach that might improve its sensitivity. Paulhus (1991) has convincingly demonstrated that socially desirable responding (SDR) is not a univocal construct, but rather comprises two kinds responding, what he calls "self-deceptive positivity" (in which the responder gives an overly positive representation of himself) and "impression management" (in which the responder deliberately tailors his responses to his audience). Paulhus (1991) has developed a questionnaire, the Balanced Inventory of Desirable Responding. We have adapted its response format to the five point true-false response format of the MASA. In addition, we constructed analogues of the MMPI-2 True Response Inconsistency Scale (TRIN) and Variable Response Inconsistency Scale (VRIN) for the MASA to assess acquiescence and random responding (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989). Finally, we are attempting to improve the partially successful Sexual Behavior Lie Scale, by changing its response format and adding new items.

Our second strategy involves the implementation one of the more promising recent techniques for assessing response biases. This approach focuses on the

process of questionnaire responding rather than simply on response content (e.g., Holden, 1995). The Authorware software we have used to program the MASA has the option of measuring response latencies to questions. In Versions 4 and 5 of the MASA the average differential latencies to selective, reading-length-matched questions that ask about neutral and sensitive areas of various response domains are recorded. Holden (1995) has demonstrated that respondents who are dissimulating by denying deviant or negative behaviors or cognitions take longer than honest respondents to endorse items that describe negative characteristics about themselves. Not only can the response latency measures be used to detect fakers, but Holden and Hibbs (1995) have provided evidence that the use of item response latencies can account for dissimulation variance not accounted for by standard validity indexes, and thus can add incremental validity to detecting duplicity.

The third strategy uses item response theory (IRT) to generate an appropriateness measurement aimed at identifying dishonest respondents. Zickar and Drasgow (1996) have demonstrated that when attempting to lower false positive misclassifications this approach more successfully identified faking respondents than the social desirability content approach

We have recently begun an analogue study in which we administered the MASA to male college students under three instructional sets—honest responding, faking good (encouraging them to lie to make a positive impression), and faking bad (encouraging them to lie to make a negative impression). Preliminary results indicate that Paulhus' self-deceptive positivity scale has only moderate success in differentiating fake bad subjects from honest subjects and no success in differentiating fake goods.

Consistent with the nature of the manipulation, his impression management scale was very successful in differentiating fake bad subjects, but just missed significance in discriminating fake goods from honest respondents. Similarly, our revised sex lie scale had only moderate success in differentiating fake bads and no ability to distinguish fake good subjects. In contrast, Holden's reaction time measure discriminated both types of lying from honest responding.

We have also completed a study of 95 non-civilly committed sexual offenders at the MTC program, who were quite concerned about the issue of civil commitment and thus were primed to be very defensive. We tested them under conditions of anonymity and feedback to the clinical staff. Although we have not had the opportunity to analyze these data, it is already clear from the testing and data entry that the groups responded differently in the two conditions. Our hope is that by simultaneously applying to these data and to the other studies in progress all three of duplicity assessment approaches--content, latency, and IRT appropriateness measurements--we will be able to provide a sufficiently high level of duplicity detection to allow the MASA to be used as a clinical assessment instrument. Preliminary analyses are very promising, but suggest that the confluence of these three strategies will be necessary to achieve adequate detection hit rates. If we are successful in this research, it will be possible to apply the rich information that the MASA has yielded as a research instrument to clinical and dispositional decisions.

Although the MASA was originally conceived and constructed as an assessment instrument for rapists, the logistics of testing in institutions has resulted in a large number of child molesters taking the inventory. The determination of an offender's

status as a rapist, child molester, mixed-age offender has most frequently been made after testing by ratings from the offender's criminal history. The original factor analyses of the ten domains and the replication factor analyses in the sex offender sample both included child molesters. Consequently, the factor scales thus far generated are equally appropriate for both child molesters and rapists. Many of the child molesters tested were dismayed by their perception that critical components of their offense and sexual histories were not adequately assessed by the questions in the test. They made many excellent suggestions for changes and additions. Using both their recommendations and also incorporating domains that we had found important in our child molester typology research (Knight, 1992), we have in Versions 4 and 5 created complete new subroutines assessing fixation on children, sexual preference, identification of the range of victim ages and the sex of victims, amount of contact with children, and offense modus operandi specific to child molesters. We are currently validating these new additions on sex offender samples in Maine, Massachusetts, and Minnesota.

The current version the MASA (Version 5), which exists only in a computerized format, takes full advantage of contingency based questioning and has added modules for a detailed developmental history, greater assessment of adolescent social competence, as well as the more extensive evaluation of domains unique to child molesters, just described. It now has the potential to fill the role of a standardized, comprehensive self-report instrument for sex offenders. If, as now seems possible, the duplicity problem can be adequately addressed, the next step in its development will be the creation of computer algorithms to provide user friendly feedback to aid in clinical and dispositional decision making.

References

Anderson, W. P., Kunce, J. T., & Rich, B. (1979). Sex offenders: Three personality types. Journal of Clinical Psychology, 35, 671-676.

Bard, L. A., & Knight, R. A. (1987). Sex offender subtypes and the MCMI. In C. Green (Ed.), Conference on the Millon inventories (MCMI, MBHI, MAPI) (pp. 133-137). Minneapolis: National Computer Systems.

Berenbaum, H. (1995). Peculiarity. In C. G. Costello (Ed.), Personality characteristics of the personality disordered (pp. 206-241). New York: Wiley.

Burgess, A. W., & Holmstrom, L. L. (1974). Rape trauma syndrome. American Journal of Psychiatry, 131, 981-986.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). Manual for the restandardized Minnesota Multiphasic Personality Inventory: MMPI-2. An administrative and interpretive guide. Minneapolis, MN: University of Minnesota Press.

Carpenter, D. R., Peed, S. F., & Eastman, B. (1995). Personality characteristics of adolescent sexual offenders: A pilot study. Sexual Abuse: A Journal of Research and Treatment, 7, 195-203.

Chantry, K., & Craig, R. J. (1994). Psychological screening of sexually violent offenders with the MCMI. Journal of Clinical Psychology, 50, 430-435.

Cohen, M. L., Garofalo, R. F., Boucher, R., & Seghorn, T. (1971). The psychology of rapists. Seminars in Psychiatry, 3, 307-327.

Erickson, W. D., Luxenberg, M. G., Walbek, N. H., & Seely, R. K. (1987). Frequency of MMPI two-point code types among sex offenders. Journal of Consulting

and Clinical Psychology, 55, 566-570.

Gearing, M. L. (1979). MMPI as a primary differentiator and predictor of behavior in prison: A methodological critique and review of the recent literature. Psychological Bulletin, 86, 929-963.

Groth, A. N., Burgess, A. W., & Holmstrom, L. L. (1977). Rape: Power, anger, and sexuality. American Journal of Psychiatry, 134, 1239-1243.

Grubin, D. & Prentky, R. (1993). Sexual psychopathy laws. Criminal Behaviour and Mental Health, 3, 381-392.

Hall, G. C. N., Maiuro, R. D., Vitaliano, P. P., & Proctor, W. C. (1986). The utility of the MMPI with men who have sexually assaulted children. Journal of Consulting and Clinical Psychology, 54, 493-496.

Hanson, R. K. (1990). The psychological impact of sexual assault on women and children: A review. Annals of Sex Research, 3, 187-232.

Hare, R. D. (1980). A research scale for the assessment of psychopathy in criminal populations. Personality and Individual Differences, 1, 111-119.

Hare, R. D., Harpur, T. J., Hakstian, A. R., Forth, A. E., Hart, S. D., & Newman, J. P. (1990). The revised Psychopathy Checklist: Reliability and factor structure. Psychological Assessment, 2, 338-341.

Hartman, H. H. (1967). Modern factor analysis. Chicago: University of Chicago Press.

Hazelwood, R. R. (1987). Analyzing the rape and profiling the offender. In R. R. Hazelwood & A. W. Burgess (Eds.), Practical aspects of rape investigation: A multidisciplinary approach (pp. 169-199). New York: Elsevier.

Holden, R. R. (1995). Response latency detection of fakers on personnel tests. Canadian Journal of Behavioural Science, 27, 343-355.

Holden, R. R., & Hibbs, N. (1995). Incremental validity of response latencies for detecting fakers on a personality test. Journal of Research in Personality, 29, 362-372.

Kalichman, S. C., Szymanowski, D., McKee, G., Taylor, J., & Craig, M. E. (1989). Cluster analytically derived MMPI profile subgroups of incarcerated adult rapists. Journal of Clinical Psychology, 45, 149-155.

Kalichman, S. C., Henderson, M. C., Shealy, L. S., & Dwyer, M. (1992). Psychometric properties of the Multiphasic Sex Inventory in assessing sex offenders. Criminal Justice and Behavior, 19, 384-395.

Knight, R. A. (1992). The generation and corroboration of a taxonomic model for child molesters. In W. O'Donohue & J. H. Geer (Eds.), The sexual abuse of children: Theory, research, and therapy, Vol. 2 (pp. 24-70). Hillsdale, NJ: Lawrence Erlbaum.

Knight, R. A. (1993, November). Developmental antecedents of sexual aggression. Paper presented at the 12th Annual Conference of The Association for the Treatment of Sexual Abusers, Boston, MA.

Knight, R. A. (1995, October). A unified developmental theory of sexual aggression: Models in the making. Paper presented at the 14th Annual Meeting of the Association for the Treatment of Sexual Abusers, New Orleans, Louisiana.

Knight, R. A. (1997, October). A unified model of sexual aggression: Consistencies and differences across non-criminal and criminal samples. Paper presented at the 16th Annual Meeting of the Association for the Treatment of Sexual Abusers, Arlington, Virginia.

Knight, R. A. (1999). Validation of a typology for rapists. Journal of Interpersonal Violence, 14,

Knight, R. A., & Prentky, R. A. (l990). Classifying sexual offenders: The development and corroboration of taxonomic models. In W. L. Marshall, D. R. Laws, & H. E. Barbaree (Eds.), The handbook of sexual assault: Issues, theories, and treatment of the offender (pp. 27-52). New York: Plenum.

Knight, R. A., & Prentky, R. A. (1993). Exploring the characteristics for classifying juvenile sexual offenders. In H. E. Barbaree, W. L. Marshall, & S. M. Hudson (Eds.), The juvenile sex offender (pp. 45-83). New York: Guilford Press.

Knight, R. A., Prentky, R. A., & Cerce, D. (1994). The development, reliability, and validity of an inventory for the Multidimensional Assessment of Sex and Aggression. Criminal Justice and Behavior, 21, 72-94.

Knight, R. A., Rosenberg, R., & Schneider, B. (1985). Classification of sexual offenders: Perspectives, methods and validation. In A. Burgess (Ed.), Rape and sexual assault: A research handbook (pp. 222-293). New York: Garland.

Knight, R. A., Warren, J. I., Reboussin, R., & Soley, B. J. (1998). Predicting rapist type from crime-scene variables. Criminal Justice and Behavior, 25, 46-80.

Koss, M. P., Woodruff, W. J., & Koss, P. (1991). Criminal victimization among primary medical patients: Prevalence, incidence, and physician usage. Behavioral Sciences and the Law, 9, 85-96.

Langevin, R., Lang, R., Reynolds, R., Wright, P, Garrels, D., Marchese, V. Handy, L., Pugh, G. & Frenzel, R. (1988). Personality and sexual anomalies: An examination of the Millon Clinical Multiaxial Inventory. Annals of Sex Research, 1, 13-