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

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

Criminal Justice and Behavior, 36(6), 567-590. doi: 10.1177/0093854809333610

While assessment of violence risk should play a crucial role in sentencing, release, case management, and rehabilitation methods, there are very few meta-analyses comparing different risk assessment instruments and their predictive validity (Andrews & Bonta, 2006; Heilbrun, 1997; Campbell, French, & Gendreau, 2009). More specifically, meta-analyses look at the predictive validity for predicting risk, identifying risk reduction targets, and monitoring risk level changes (Campbell et al., 2009). To properly inform professionals in the field, the present authors conduct a meta-analysis on a wide range of instruments used to assess risk in adult offenders and forensic patients.

First generation risk assessments during the mid-20th century were unstructured clinical judgments where error and bias were common (Grove et al., 2000; Monahan & Steadman, 1994; Rice, 1997). Second generation assessments of risk (e.g., Violence Risk Assessment Guide [VRAG])became standardized tools based on items statistically predictive of recidivism, but were criticized for lacking theoretical foundation and having only static, unchanging items such as criminal history (Campbell et al., 2009). Third generation tools (e.g., Level of Supervision Inventory-Revised [LSI-R]; Historical, Clinical, and Risk Management Violence Risk Assessment Scheme [HCR-20]) aim to predict risk and identify criminogenic needs that could help reduce risk (Andrews et al., 2006; Bonta, 2002). Instruments from this generation select risk factors based on theories related to criminality and violence, and are empirically supported (Andrews & Bonta, 2006; Gendreau, Goggin, French, & Smith, 2006). Furthermore, third

generation tools include dynamic risk factors that can vary with time and be influenced by other variables (e.g., social, contextual, biological) (Douglas & Skeem, 2005). The most current, fourth generation instruments (e.g., Level of Service/Case Management Inventory [LS/CMI]), are integrated into risk management, intervention and treatment selection, and assessment of rehabilitation progress in order for evaluators to document changes in criminogenic needs, and to identify areas of success within a case management plan and intervention strategies that may require revision to increase success of risk reduction (Campbell et al., 2009). Given the variety of tools available to the professionals tasked with conducting risk assessments, they are confronted with selecting the appropriate assessment tools for the specific population and setting, and the type of construct (general violence vs. sexual violence) being assessed. Aside from these concerns, they also need to consider the proper assessment administration method to ensure predictive accuracy (Campbell et al., 2009).

The current authors indicated that there are four meta-analyses that examined risk prediction tools for adults. Gendreau et al. (1996) compared five different instruments for general recidivism and found that the LSI-R had the strongest effect size and the remaining tools were moderately predictive. Gendreau, Goggin, and Law (1997) evaluated the LSI-R and the Minnesota Multiphasic Personality Inventory (MMPI) with other risk measures, and non-MMPI antisocial personality instruments. Again, the LSI-R had the best predictive value when compared to other measures. Gendreau, Goggin, and Smith (2002) examined violent recidivism and found that the LSI-R just barely out predicted the Psychopathy Checklist-Revised (PCL-R). Walters (2006) compared a group of structured/actuarial instruments to self-report measures and found that self-report tools were predictive if was constructed by items empirically tied to risk (e.g., antisocial attitudes).

The purpose of the study was to identify which risk assessment tools can best predict future non-sexual violence, and compare these tools in five areas: type of risk measure, type of risk assessment generation, type of risk factors (static vs. dynamic vs. static/dynamic), method of administration, and correction relevant content (Campell et al., 2009). The risk measures included the HCR-20, LSI/LSI-R, PCL/PCL-R, PCL:SV, SIR scale, and VRAG. Eighty-eight studies were examined and 185 effect sizes for violent recidivism and 76 effect sizes for institutional violence were produced within the study. Results of the meta-analysis are described below.

Prediction of Violent Recidivism

Comparisons of Static and Dynamic-Based Instruments

Campbell et al. (2009) found that instruments mostly comprised of dynamic risk factors created the strongest effect size for violent recidivism, which suggests that they may have an advantage over measures using static factors to predict risk.

Comparisons of Risk Assessment Generation

Third generation instruments resulted in a better estimate than second generation (Campbell et al., 2009). Both of which are replications of previous studies (Gendreau et al., 1996; Schwalbe, 2007). Of the instruments examined, they appeared to be similar in their ability to predict violent reoffending. The fourth generation tools produced the strongest predictive estimate of the different generation. Additional research on this newer generation of risk instruments is needed (Campbell et al., 2009).

Comparisons of Risk Instruments

Results showed that each risk measure predicted violent recidivism with at least a moderate degree of success, with VRAG having the strongest effect size. The Statistical

3

Information of Recidivism (SIR) scale and the PCL-R were comparable to LSI-R in predicting

violent recidivism. Parallel to previous findings, current results indicated that the risk measures

are moderately to highly intercorrelated, indicating a significant overlap between common risk

measures (Campbell et al., 2009).

Comparisons of Content Relevance and Administration Method

Current results revealed that the content relevant to criminal behavior and risk was more

accurate in predicting violent reoffending than those with less relevant content. This brings

caution to using the MMPI to predict future violence, although it is commonly used today.

Campbell et al. (2009) findings suggest that self-report measures should be included in the

violence risk assessment, but not as the sole measure.

The file review only and file review plus interview methods produced the largest

predictive validity for predicting violent recidivism. However, the results may be affected by not

separating relevant and less-relevant content self-report measures (Campbell et al., 2009).

Prediction of Institutional Violence

Comparisons of Static and Dynamic-Based Instruments

Results showed that risk instruments with static factors generated the largest effect size

than those with dynamic or combined (static/dynamic) factors.

Comparisons of Risk Assessment Generation

In contrast to violence reoffending, second generation instruments, those based on

criminal history and other static variables predicted more accurately than third generation tools.

It may be that the static factors were more valuable when evaluating institutional violence due to

the short-term (i.e., one-year) follow-up assessments. Dynamic factors may be more useful for

long-term predictions of institutional violence because they may require more time to mature (Campbell et al., 2009).

Comparisons of Risk Instruments

Unlike predicting violent recidivism, there was much more variability within each risk instrument when predicting institutional violence. Although criminal history had the largest effect size, different types of crimes were included in this category that made it difficult to interpret. The HCR-20 had the largest effect size for institutional violence, but the data was derived from the risk score, not the clinical prediction judgment. In addition, the data was extracted from the forensic psychiatric samples, which are not generalizable to non-psychiatric correctional facilities. The PCL-SV and the LSI-R had moderate ability to predict violence. The PCL-R and VRAG generated small effect sizes when predicting institutional violence (Campbell et al., 2009).

Comparisons of Content Relevance and Administration Method

The file review only and file review plus interview methods produced the largest predictive validity for predicting institutional violence. However, the results may be affected by not separating relevant and less-relevant content self-report measures (Campbell et al., 2009).

Limitations of the present study should be considered. First, coding of important variables of potential moderators was not always possible due to insufficient information on variables (e.g., violence history) studied in previous literature. Second, studies that examined institutional violence did not provide details about previous levels of institutional violence. Also, information about violent vs. non-violent index offenses was lacking for 56% of the effect sizes. This prohibited the authors from examining the moderating effects of index offense severity on predictive validity. In addition, the samples were predominantly male that the results cannot be

applied to females. Lastly, 88% of the effect sizes were generated from low or moderate risk of

reoffending; therefore, additional research is warranted to generalize the results to high risk

offenders (Campbell et al., 2009).

In summary, this study found little difference between actuarial and structured tools in

predicting violent reoffending, which does not suggest they are equally informative for case

planning, particularly when the goal is to reduce risk. Future research should identify predictive

factors that are related to the nature/context of an offender's violent behavior, examine the

incremental validity of risk instruments, adjust the composition of the comparison group (i.e., not

include recidivism of other types of crime in the non-recidivism group) for violent recidivism,

and identify a standard method to interpret the confidence intervals (Campbell et. al, 2009).