

Predictive Validity of the Structured Assessment of Protective Factors (SAPROF): A Meta-Analytic Review

Amanda Flake, B.A., & Adam P. Natoli, Ph.D.
Sam Houston State University

Introduction

Risk assessment research has primarily focused on enumerating factors underlying risk for future misconduct (e.g., violence, sexual, general misconduct); however, relatively less is known about protective factors and how they can inform risk assessment. This is a notable gap because protective factors are associated with a lower likelihood of negative outcomes when compared to risk factors and can reduce their negative impact.

The Structured Assessment of Protective Factors for violence risk (SAPROF) has become one of the most widely used instruments for assessing protective factors in adult risk evaluations. Nonetheless, most adult-based risk assessments continue to partially or entirely disregard protective factors. Given the importance of considering both risk and protective factors, and the inherent benefits of doing so, this study used meta-analysis to summarize predictive validity research on the SAPROF.

Although the SAPROF was originally developed to exclusively assess protective factors for violence risk, researchers have also examined its validity in the context of other types of misconduct. Thus, this meta-analysis summarizes the SAPROF's predictive validity regarding three outcomes: abstinence from any misconduct, from violence, and from sexual misconduct.

Method

Inclusion Criteria and Literature Search

For inclusion in the meta-analysis, a study had to (a) include the SAPROF, (b) analyze the SAPROF's ability to explain variability in abstinence from future misconduct and (c) provide enough statistical information that an effect size statistic representing the difference in each SAPROF score between those who did and did not engage in future misconduct could be derived. To gather published studies for potential inclusion, an extensive database search was performed in late-2020 using ProQuest and EBSCOhost (including Medline). After removing duplicate articles 246 articles were screened for eligibility and the full text of 62 articles were assessed. Of these 62 articles, 43 articles were eliminated for various reasons. This procedure yielded a total of 19 publications including 20 separate studies (154 effect sizes, total N = 3,182) for inclusion in the meta-analysis.

Calculation of Effect Sizes and Statistical Procedure

Meta-analysis was conducted using the software JASP. In addition to a combined effect size d and computation of 95% confidence intervals around these effect sizes, a combined Z was calculated for each effect size to extend the information obtained from d by considering sample size, which varied widely across studies. The current meta-analysis examined the results of two heterogeneity variance estimators known to perform well given our objectives, characteristics of the included studies, and the properties of study data. For both d and Z , a positive sign was assigned when individuals who abstained from future misconduct obtained greater SAPROF scores than individuals who engaged in future misconduct, and a negative sign was assigned when this relationship was reversed.

Assessment of Publication Bias

Given the potential implications of the current meta-analyses, three strategies were employed to address issues of publication bias. A fail-safe N (FSN; Rosenthal, 1979, 1991), funnel plot analyses and Egger's regression tests of funnel plot asymmetry (Egger et al., 1997), and Duval and Tweedie's (2000a, 2000b) trim-and-fill procedure.

Risk assessments maintain an influential role in the justice system, but the failure to examine both risk and protective factors can create implicitly biased evaluations. The SAPROF has become one of the most widely used instruments for assessing protective factors in adult risk evaluations. Meta-analytic results offer evidence supporting the SAPROF's predictive validity and admissibility regarding both the Daubert and Frye standards.

Results

Predictive validity for general misconduct

As Table 2 summarizes, individuals who abstained from any form of misconduct following their risk evaluation obtained significantly higher SAPROF scores, with results indicating the strongest predictor being the FPJ ($d = 0.803$, standard error [SE] = 0.128, FSN = 24535).

Score	Number of effect sizes	n	Combined effect size d (SE)	95% CI	Combined Z	p	Fail-safe N	Egger's Test z	Corrected for publication bias
TPS	20	3182	0.721 (.081)	0.563, 0.879	8.944	<.001	93156	0.748 ($p = .454$)	NA
FPJ	7	630	0.803 (.128)	0.551, 1.055	6.256	<.001	24535	-0.811 ($p = .417$)	NA
Internal	14	1982	0.524 (.108)	0.313, 0.735	4.857	<.001	23091	0.904 ($p = .366$)	NA
Motivational	14	1982	0.618 (.075)	0.472, 0.764	8.289	<.001	35012	0.308 ($p = .758$)	NA
External	14	1980	0.172 (.066)	0.044, 0.301	2.627	.009	4697	-1.064 ($p = .288$)	NA

Note. TPS = Total protective score; FPJ = Final protection judgment rating.

Predictive validity for violent and sexual misconduct

Trends in the SAPROF's predictive validity similar to the first meta-analysis emerged when abstinence from violence was examined individually ($k = 16$) whereas meta-analysis produced diverging results pertaining to the SAPROF's ability to predict an individual's abstinence from engaging in future sexual misconduct

Score	Number of effect sizes	n	Combined effect size d (SE)	95% CI	Combined Z	p	Fail-safe N	Egger's Test z	Corrected for publication bias
Violence									
TPS	16	2836	0.744 (.093)	0.562, 0.927	7.989	<.001	49993	0.796 ($p = .426$)	NA
FPJ	5	486	0.841 (.127)	0.593, 1.090	6.634	<.001	7278	-0.067 ($p = .947$)	0.98 (0.72, 1.23)
Internal	11	1741	0.533 (.124)	0.291, 0.776	4.312	<.001	14087	0.133 ($p = .894$)	0.43 (0.18, 0.68)
Motivational	11	1741	0.650 (.096)	0.462, 0.838	6.770	<.001	17594	1.281 ($p = .200$)	NA
External	11	1739	0.215 (.092)	0.035, 0.395	2.340	.019	4562	-1.342 ($p = .180$)	NA
Sexual Misconduct									
TPS	6	1361	0.277 (.103)	0.075, 0.479	2.682	.007	2004	-0.660 ($p = .510$)	NA
FPJ	2	146	0.516 (.012)	0.493, 0.539	43.658	<.001	831	NA	NA
Internal	5	1132	0.310 (.137)	0.042, 0.578	2.267	.023	1612	-0.602 ($p = .547$)	NA
Motivational	5	1132	0.277 (.057)	0.166, 0.388	4.885	<.001	1044	-0.467 ($p = .641$)	NA
External	5	1130	0.181 (.069)	0.046, 0.316	2.621	.009	753	-1.568 ($p = .117$)	NA

Note. TPS = Total protective score; FPJ = Final protection judgment rating.

Regarding sexual misconduct, the SAPROF Internal domain scale ($d = 0.310$, SE = 0.137, FSN = 1612) emerged as the strongest predictor, performing better than the SAPROF TPS ($d = 0.277$, SE = 0.103, FSN = 2004), while the External domain scale ($d = 0.181$, SE = 0.069, FSN = 753) was again found to be the poorest estimate of abstinence (see Table 4).

Mean effects across outcomes

Comparison of mean effects across outcomes indicated several significant differences (see Table 6). Estimated effects with respect to abstinence from sexual misconduct were significantly smaller than the estimated effects obtained when SAPROF ratings were used to differentiate individuals who engaged in any form of misconduct from those who abstained. Comparisons of mean effects between the outcomes of abstinence from any misconduct and abstinence from violence only revealed a single significant difference. Specifically, significantly larger estimated effects were found when the FPJ ($Z_{Diff}^* = -38.000$, $p < .001$) was used to predict abstinence from violence. Finally, except for the SAPROF External domain scale ($Z_{Diff}^* = 1.478$, $p = .139$), all estimated effects were found to be significantly larger when SAPROF ratings were used to predict abstinence from violence than when used to predict individuals' abstinence from sexual misconduct. Results of mean effects comparisons across the five SAPROF ratings within each outcome produced multiple significant differences and are reported in Table 7.

Score	Violent versus Sexual		Any versus Violent		Any versus Sexual	
	Z_{Diff}^*	2-tailed p -value for Z	Z_{Diff}^*	2-tailed p -value for Z	Z_{Diff}^*	2-tailed p -value for Z
TPS	-46.700	<.001	1.917	.055	-20.182	<.001
FPJ	2.826	.005	-38.000	<.001	2.474	.013
Internal	-17.154	<.001	0.563	.574	-7.379	<.001
Motivational	9.564	<.001	1.524	.128	18.944	<.001
External	1.478	.139	1.654	.098	3.000	.003

Note. TPS = Total protective score; FPJ = Final protection judgment rating.

Score	Outcome	FPJ	Internal	Motivational	External
		Z_{Diff}^* (p)	Z_{Diff}^* (p)	Z_{Diff}^* (p)	Z_{Diff}^* (p)
TPS	Any	1.745 (.081)	-7.296 (<.001)	17.167 (<.001)	36.600 (<.001)
	Violence	2.853 (.004)	-6.806 (<.001)	-31.333 (<.001)	529.000 (<.001)
	Sexual	-2.626 (.009)	0.971 (.332)	0.000 (1.00)	2.824 (.005)
FPJ	Any	13.950 (<.001)	3.491 (<.001)	10.177 (<.001)	17.886 (<.001)
	Violence	102.667 (<.001)	6.161 (<.001)	17.886 (<.001)	17.886 (<.001)
	Sexual	-1.648 (.099)	-5.311 (<.001)	-5.877 (<.001)	-5.877 (<.001)
Internal	Any		-2.848 (.004)	8.381 (<.001)	8.381 (<.001)
	Violence		-4.179 (<.001)	9.938 (<.001)	9.938 (<.001)
	Sexual		0.413 (.680)	1.897 (.059)	1.897 (.059)
Motivational	Any			49.556 (<.001)	49.556 (<.001)
	Violence			9.938 (<.001)	9.938 (<.001)
	Sexual			-8.000 (<.001)	-8.000 (<.001)

Note. TPS = Total protective score; FPJ = Final protection judgment rating.

Discussion

Every individual possesses a variety of biological and psychological characteristics that have the ability to make them either vulnerable to or protected against potential behavioral health issues. Risk assessment protocols have been globally implemented in criminal justice settings to prioritize risk reduction strategies and to help prevent the continuation of serious crime before it occurs. Although knowledge of risk factors has increased greatly over the years, far less is known about factors that may buffer against the negative effects posed by the presence of such risk factors.

Whereas it remains unclear whether risk assessment tools help to reduce future offending behavior, research has provided evidence suggesting protective factors can help prevent future misconduct. The SAPROF is one of the most widely used instruments for evaluating these protective factors but, as summarized above, evaluations of this instrument's psychometric properties, particularly its predictive validity, have produced mixed findings.

Existing psychometric investigations of the SAPROF have provided varying degrees of support for this instrument's predictive validity; meta-analytic synthesis of results has the potential to summarize this evidence and calculate a more precise estimation of the SAPROF's predictive validity by providing a systematic compilation of the magnitude of differences in SAPROF scores between individuals who had or had not engaged in misconduct following their evaluation. Meta-analysis also allows for the influence of potential moderating variables on the SAPROF's predictive validity (e.g., type of misconduct) to be directly examined. Accordingly, this study used a meta-analytic approach to evaluate the SAPROF's predictive validity, synthesizing results from all published research examining this issue.

Comprehensive and efficacious risk assessment should involve the evaluation of both risk and protective factors. The current meta-analyses demonstrated the adequate predictive validity of the SAPROF in assessing protective factors associated with abstinence from engaging in any form of misconduct, as well as abstinence from future violence and sexual misconduct, independently. Combined with the SAPROF's standardized procedures for application, accepted reliability and known potential error rate, and growing acceptance within the field, these meta-analytic results, which describe the empirical summarization of the currently available peer-reviewed research, offer evidence supporting the instrument's admissibility regarding both the Daubert and Frye standards. Continued research is needed before these claims can be generalized to women and gender diverse adults, but current results are promising. In conclusion, the SAPROF is opined to be a valid measure of protective factors with respect to violence, sexual misconduct, and general future misconduct.

References available by request