

# Differences between SAPROF 2<sup>nd</sup> Edition and SAPROF version 1

May 2012

M. de Vries Robbé & V. de Vogel

***The paragraphs on Research with the SAPROF-RV (V1 p20) and From SAPROF-RV to SAPROF Version 1 (V1 p20-21), have been replaced by the paragraph The SAPROF (2nd Ed p18-20).***

## *The SAPROF*

In 2006, we conducted a pilot study with the SAPROF-RV in two Dutch forensic psychiatric settings (Van der Hoeven Kliniek and Pompekliniek) and one Dutch forensic outpatient setting (De Waag). The major aim of the study was to evaluate the practical applicability of the SAPROF-RV. In this study, 20 raters coded the SAPROF-RV for 40 patients and critically reviewed the item descriptions and coding procedures. Raters were asked if they thought the items were relevant for their daily work and if they considered the guidelines and item descriptions to be clear. All raters were positive about the usefulness of the SAPROF, especially with respect to the positive approach towards the patient and the suggestions offered for treatment. Some example remarks were: “The SAPROF provides mental health professionals and patients with a more positive starting-point”; “The SAPROF makes us more aware of what we still can and need to invest in”. They also had several suggestions for improvements of the item descriptions and the SAPROF in general. Based on these comments and a literature review update, we adapted the SAPROF-RV, which resulted in the present SAPROF.

The adaptations consisted of eliminating items that were difficult to code (7. *Religion / philosophy of life*<sup>1</sup>; 16. *Satisfying sexual relationship with adult partner*), adding items that were missing (5. *Self-control*; 11. *Life goals*) and splitting the item 4 (*Positive attitude towards intervention and authority*) into 9 (*Motivation for treatment*) and 10 (*Attitudes towards authority*). Furthermore, the item names and descriptions were formulated more neutrally and unambiguously, the coding instructions became more detailed and several examples were added. Also, an additional literature search was conducted for the new items 5 (*Self-control*) and 11 (*Life goals*). For the other items, the literature was updated (literature published between 2002-2007) and the rationale for the items was modified when needed. Finally, the categorization of the items and the names of the sub-scales were changed. Initially, the sub-scales of the HCR-20 (Historical scale, Clinical scale and Risk management scale) had been used as a framework for the structure of the SAPROF-RV. The reason for this was the aim of

---

<sup>1</sup> Religion has become part of the broader item Life goals in the SAPROF.

maximum correspondence between the SAPROF and the HCR-20. However, many users considered this categorization somewhat artificial and illogical. In the HCR-20, a clear distinction is made between past, present and future. However, in the SAPROF, almost all items are dynamic. In the process of evaluating the SAPROF, we were able to distinguish between items that refer to aspects of the individual (internal) and items that are external. This distinction was also considered valuable for future risk management. Furthermore, several of the items could be related to motivation for treatment, and - more general - to positive participation in society. The present SAPROF is divided into three new sub-scales: Internal items, Motivational items and External items (see Table 1). Internal items are personal characteristics that can be protective. Motivational items comprise protective factors that reflect an individual's motivation to participate in society in a positive manner. External items deal with protection from outside factors.

***The paragraph Applications (V1 p25) has been replaced by the two paragraphs Applications and Gender differences (2nd Ed p21-22).***

## **Applications**

The SAPROF can be used by mental health experts trained in risk assessments<sup>2</sup> who work in the prison system, (forensic) psychiatric institutions (in- and outpatient) and the probation service. The SAPROF can be applied in cases where risk assessment is needed, for example, in forensic mental health evaluations for the court, when a decision on extension or termination of forensic psychiatric treatment needs to be made, before leave privileges are granted, or at the beginning of a resocialisation phase. Risk assessment can also be useful at the start of inpatient therapy, in order to assess the risk of violent behavior within the institution and to draw up a treatment plan. Repeated assessments are recommended if there are changes in an individual's context; after all, an individual's risk level and protective factors can fluctuate over time. Repeated assessments of risk and protective factors are especially important because those responsible for supervising individuals are encouraged to reflect on a regular basis on the necessity of changes in supervision and intervention policies.

## **Gender differences**

Most research on risk assessment has been carried out among male populations; this is true

---

<sup>2</sup> Specific training in the SAPROF is recommended. For more information about workshops you may contact the authors or Forum Educatief (see inside cover page of this manual).

for research on the risk factors included in risk assessment tools, as well as for research on the validity of these tools (see Garcia-Mansilla, Rosenfeld, & Nicholls, 2009; Odgers & Moretti, 2002; Odgers et al., 2007). However, research has shown that the predictive validity of both the HCR-20 and the *Psychopathy Checklist-Revised* (PCL-R; Hare, 1991, 2003) is considerably lower for women compared to men (De Vogel & De Ruiter, 2005) and that for women different risk factors may well be involved (De Vogel & De Vries Robbé, in press; Funk, 1999; Odgers & Moretti, 2002; Odgers et al., 2007). Therefore, a new risk tool was developed recently as an addition to the HCR-20 for violence risk assessment in women; the *Female Additional Manual* (FAM; de Vogel, de Vries Robbé, van Kalmthout & Place, 2011). With the addition of the FAM the authors hope to create a more gender sensitive assessment of risk factors for violent behavior.

Like most risk focused SPJ tools, the SAPROF was also initially developed to assess adult males with a history of violence who suffer from a mental or personality disorder. However, it has been suggested that women may respond differently to protective factors (Rumgay, 2004). Nevertheless, since there is little known about protective factors specifically for women and results to date with the SAPROF in female populations show promising results (see *Prospective clinical studies*, p. 34), for now the SAPROF is considered to be a useful tool for women as well as for men. The assessor is advised to consider the limited body of evidence for the value of protective factors for women when drawing conclusions regarding risk judgment in women. Future empirical research will have to prove the value of the SAPROF for female populations, we therefore recommend that research results of the SAPROF are described separately for men and women.

***A paragraph on Time frame for coding was added (2nd Ed p25).***

**Time frame for coding**

The items should be coded dynamically, predominantly based on information from the past 6 months and the current plans regarding the near future. Coding is done with the context in mind for which the assessment is being carried out. The items should thus be coded with a similar time frame as used for the Risk management items in the HCR-20. For three items the time frame is different: *Intelligence* (item 1) is coded based on intelligence testing carried out in the past 6 years; *Secure attachment in childhood* (item 2) concerns the entire childhood; and *Self-control* (item 5) incorporates information from the past 12 months in order to be able to establish persevering self-control. The overall SAPROF assessment is valid for 12 months, providing that the context stays the same.

***In the paragraph Marking the critical items a remark was added on the maximum number of critical items (2nd Ed p27).***

## ***2. Marking the critical items***

The assessor is advised to use the option of coding critical items sparingly. It is advised to mark a maximum of 4 *Key* and 4 *Goal* items as otherwise the essence of the critical items will disappear; it applies to items that are clearly more important for the individual than the other items.

***In the paragraph Coding the Final Protection Judgment a section was added on the possibility of clinically using a 5-point judgment scale (2nd Ed p27).***

## ***3. Coding the Final Protection Judgment***

For clinical use it might be useful to apply a five-point rating scale for the final judgment: 1) Low; 2) Low to Moderate; 3) Moderate; 4) Moderate to High; and 5) High protection. Applying a five-point scale instead of a three-point scale makes it easier to pinpoint nuances; in a forensic population where the treatment progress is usually slow, it can be useful and motivating to be able to show small changes. In addition, research at the Van der Hoeven Kliniek showed higher predictive validities for final judgments made on five-point scales than on three-point scales (see Research with the SAPROF, p. 30).

***In the paragraph Coding the Integrative Final Risk Judgment remarks were added on the time frame for coding and a reference to risk formulation and scenarios (2nd Ed p27).***

## ***4. Coding the Integrative Final Risk Judgment***

The Final Risk Judgment should be made for the coming year. In addition to the final judgments it is advised to formulate the most likely risk scenarios that describe the context, frequency, duration and possible time frame in which violence could take place, as well as identify potential victims (see Risk formulation, p. 29).

***The paragraph Using the SAPROF in practice (V1 p30-31) has been replaced by the paragraphs Using the SAPROF in practice and Risk formulation (2nd Ed p28-29).***

### **Using the SAPROF in practice**

At the start of an inpatient treatment, mental health professionals may find that mainly the external items 15 (*Professional care*), 16 (*Living circumstances*) and 17 (*External control*) are providing protection from relapse into violent behavior. Ideally the other dynamic protective factors will be addressed and strengthened during the course of treatment so eventually a shift in item scores takes place where the protection from the external treatment factors is replaced by protection from motivational and internal factors (see also Research with the SAPROF, p. 30).

Mental health professionals in the Van der Hoeven Kliniek who frequently use the SAPROF for clinical purposes stated that the tool can be helpful in formulating treatment goals, justifying transfer to another stage of treatment, phasing treatment and facilitating risk communication among those involved in the treatment of the patient (Van den Broek & De Vries Robbé, 2008). It may be useful for mental health professionals to go over the SAPROF factors together with their patients, and discuss which factors need attention during treatment. This positive, collaborative approach – looking for options and possibilities together – can be motivating for both mental health professionals and patients/clients.

### **Risk formulation**

The SAPROF findings should be integrated into a report on the overall risk assessment. When reporting on the findings from risk assessment tools and communicating about risk and protective factors, the assessor is advised not to mention scores but instead provide a clear description of the relevant factors and the final judgment on the likelihood of hypothesized future violence. In this risk formulation process, the assessor should attempt to formulate a narrative on the most likely scenario for a patient's route to violence. Issues to consider in the formulation process are the nature, severity, imminence and possible victims of the expected violent behavior. The applied risk assessment tools should give the assessor insight into the question how likely it is the putative violent behavior will take place, which factors could be risk-enhancing (risk factors) and which factors provide risk-reduction (protective factors). Regarding protective factors it would be valuable to describe which factors are most important in preventing the patient from becoming violent and, in addition, which factors offer the most potential for improvement in the level of protection and should thus become first priority for treatment intervention.

***The paragraph Research with the SAPROF Version 1 (V1 p21-23) has been replaced by the updated research chapter Research with the SAPROF (2nd Ed p30-36).***

## **Research with the SAPROF**

### *Retrospective file studies*

Research with the SAPROF is being conducted in various settings around the world. Findings from validation studies in The Netherlands concern different samples of forensic psychiatric patients. Two retrospective file studies were carried out with discharged male violent (N=105) and sexual (N=83) offenders, who had completed mandatory treatment in a forensic psychiatric hospital of an average length of 5.5 years (De Vries Robbé & De Vogel, in press; De Vries Robbé, De Vogel, & De Spa, 2011; De Vries Robbé, De Vogel, Koster, & Bogaerts, in preparation). Patients in both samples were primarily diagnosed with personality disorders and to a much lesser extent with major mental disorders. In the sexual offender sample all patients had been convicted for hands-on sexual offenses, about a fifth of these included child victims. Based on file information, the SAPROF and the HCR-20 were coded at the end of treatment. For 70 cases the files were coded by two raters. Results showed excellent interrater reliability for the SAPROF total score as well as the Final Protection Judgment for both the violent (ICC = .88; ICC = .85,  $p < .01$ , single measure)<sup>3</sup> and sexual offender (ICC = .85; ICC = .73,  $p < .01$ , single measure) samples.

Outcome data for all offenders were retrieved from official criminal records with a follow-up period in the community of at least 3 years after discharge. Predictive validity was examined by using Receiver Operating Characteristics analyses (see Area Under the Curve, p. 31). Table 3 shows the predictive validity for violent recidivism of SAPROF post-treatment ratings for both groups of offenders as well as the combined sample. As Table 3 indicates, the post-treatment SAPROF total score ratings for both the violent and the sexual offender samples showed good predictive validity for violent reconvictions at short-term (1 year) as well as medium-term (3 year) and long-term (average 11 year) follow-up. Although these results were equally good for violent and sexual offenders, analyses revealed different SAPROF factors that were most predictive of no future (sexual) violence for the two offender groups: *Self-control*, *Work* and *Finances* for the violent offenders; and *Coping*, *Self-control*, *Motivation* and *Attitudes* for the sexual offenders.

---

<sup>3</sup> Interrater reliability was studied using the Intraclass Correlation Coefficient (ICC; McGraw & Wong, 1996). The critical values used for ICCs (single measure) are: ICC  $\geq .75$  = excellent;  $.60 \leq \text{ICC} < .75$  = good;  $.40 \leq \text{ICC} < .60$  = moderate; ICC  $< .40$  = poor (Fleiss, 1986).

By subtracting the SAPROF total scores from the HCR-20 total scores a new measure was created in which violence risk is counterbalanced by the available protection. Table 3 shows that overall the predictive accuracy of this combined measure was higher than that of the HCR-20 alone, indicating that the addition of the SAPROF adds to the predictive power of risk-only tools. At long-term follow-up, the difference in predictive accuracy between the SAPROF and the HCR-20 and between the combined measure and the HCR-20 was significant ( $X^2(1, 188) = 9.1, p < .01$ ;  $X^2(1, 188) = 13.4, p < .01$ ). The Final Protection Judgment and the Integrative Final Risk Judgment predicted violent recidivism almost equally well. In general, final judgments made on a five-point rating scale predicted violence more accurately than those made on a three-point rating scale (see 3. *Coding the Final Protection Judgment*, p. 27). Nevertheless, AUC-values for the total scores on both tools were slightly better than those for the final judgments.

#### **Area Under the Curve**

A widely used statistical method to assess predictive validity of tools is receiver operating characteristics (ROC) analysis (Douglas et al., 2007; Rice & Harris, 1995). The major advantage of this statistical method is its insensitivity to base rates. The ROC analyses result in a plot of the true positive rate (sensitivity) against the false positive rate (1 minus specificity) for every possible cut-off score of the tool. The area under the curve (AUC) can be interpreted as the probability that a randomly selected recidivist would score higher on the risk tool or lower on the protection tool than a randomly selected non-recidivist. An AUC of .50 represents chance prediction, and an AUC of 1.0 perfect prediction. In general, AUC values of .70 and above are considered moderate to large, and above .75 large (Douglas et al., 2007). Since protective factors are meant to predict non-recidivism, the AUC values for the SAPROF represent the accuracy of predictions for non-recidivism.

**Table 3.** Predictive validity (AUC-values) for reconvictions for a violent offense of the SAPROF post-treatment ratings for discharged offenders (N=188 ♂)

Follow-up	1 year			3 year			8/15 year (M)		
	Violent N=105	Sexual N=83	Total N=188	Violent N=105	Sexual N=83	Total N=188	Violent N=105	Sexual N=83	Total N=188
Total score SAPROF	.85**	.83**	<b>.85**</b>	.74**	.77**	<b>.75**</b>	.71**	.74**	<b>.73**</b>
Total score HCR-20	.81**	.91**	<b>.84**</b>	.68*	.80**	<b>.73**</b>	.67**	.67**	<b>.64**</b>
Total HCR-20 – Total SAPROF	.85**	.89**	<b>.87**</b>	.72**	.80**	<b>.76**</b>	.71**	.72**	<b>.70**</b>
Final Protection Judgment 3-pt	.82**	.79*	<b>.80**</b>	.71**	.73**	<b>.72**</b>	.67**	.66*	<b>.66**</b>
Final Protection Judgment 5-pt	.83**	.81*	<b>.83**</b>	.72**	.71*	<b>.71**</b>	.69**	.65*	<b>.67**</b>
Integrative Final Risk Judgment 3-pt	.80**	.79*	<b>.79**</b>	.65*	.70*	<b>.67**</b>	.66*	.67**	<b>.66**</b>
Integrative Final Risk Judgment 5-pt	.85**	.83**	<b>.84**</b>	.71**	.73**	<b>.72**</b>	.68**	.68**	<b>.68**</b>

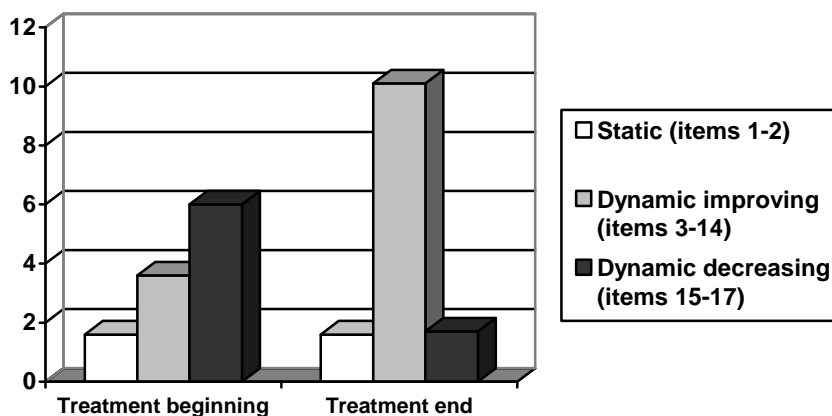
Note. \* =  $p < .05$ , \*\* =  $p < .01$  (two-tailed).



## Measuring changes during treatment with the SAPROF

For 120 of the 188 cases, the SAPROF was also coded at the beginning of treatment, based on the case history and file information from the first treatment phase. Figure 2 shows the changes in average SAPROF scores between the ratings at the beginning of treatment and the ratings at the end of treatment. Items are divided into Static (item 1-2), Dynamic improving (items 3-14) and Dynamic decreasing (items 15-17), according to their expected direction of change during treatment (see also Using the SAPROF in practice, p. 28). Comparison of the pre- and post-treatment scores showed a significant increase in the total scores on the Dynamic improving SAPROF items during treatment. As expected, the external Dynamic decreasing items showed a significant decrease towards the end of treatment. This clear shift in protection from External to Motivational and Internal is what is aimed for in treatment. In fact, the changes in SAPROF Dynamic improving item scores during treatment proved predictive of no new convictions for violent offenses long after treatment had ended (AUC = .75 at 10-year follow-up,  $p < .01$ ). Thus, patients who showed more treatment progress as measured by the SAPROF, recidivated less after treatment had ended.

**Figure 2.** Pre- and post treatment ratings SAPROF (N=120)



Note. Scores on items 1-2: 1.6 for both ratings; scores on items 3-14 changed from 3.6 to 10.1,  $p < .001$ ; scores on items 15-17 changed from 6.0 to 1.7,  $p < .001$ .

### Prospective clinical studies

Preliminary results from a prospective study including the assessments of 315 (male and female) inpatient forensic psychiatric patients was carried out to examine the clinical interrater reliability and predictive validity of the SAPROF for violent incidents during treatment (De Vries Robbé & De Vogel, 2011; De Vries Robbé, De Vogel, Wever, & Douglas, in preparation). Assessments were carried out by three independent raters from different disciplines (nursing staff member, treatment supervisor and researcher) and final scores on each item were agreed upon in a consensus meeting. Good interrater reliability was found for the SAPROF total score between the three raters of 250 multidisciplinary assessments (ICC = .70,  $p < .01$ , single measure). Consensus scores were used to calculate the predictive validity. Table 4 shows the predictive validity of SAPROF and HCR-20 scores for violent incidents towards others during the year following the assessment. Results from this clinical prospective study for the male population were very similar to those of the retrospective file studies. Again, the SAPROF predicted as well for male violent offenders as for male sexual offenders and the total scores predicted slightly better than the final judgments. Predictive values for female offenders were fairly good, although not significant in this small sample and lower than for the male patients. Again different factors were best predictors for the different groups: *Self-control, Attitudes, Work, Motivation* and *Medication* for the male violent offenders; *Coping, Leisure activities, Attitudes* and *Network* for the male sexual offenders; and *Intelligence, Coping, Work* and *Finances* for the female offenders. Overall, SAPROF scores had good predictive validity and combined HCR-SAPROF scores outperformed predictions by either instrument alone.

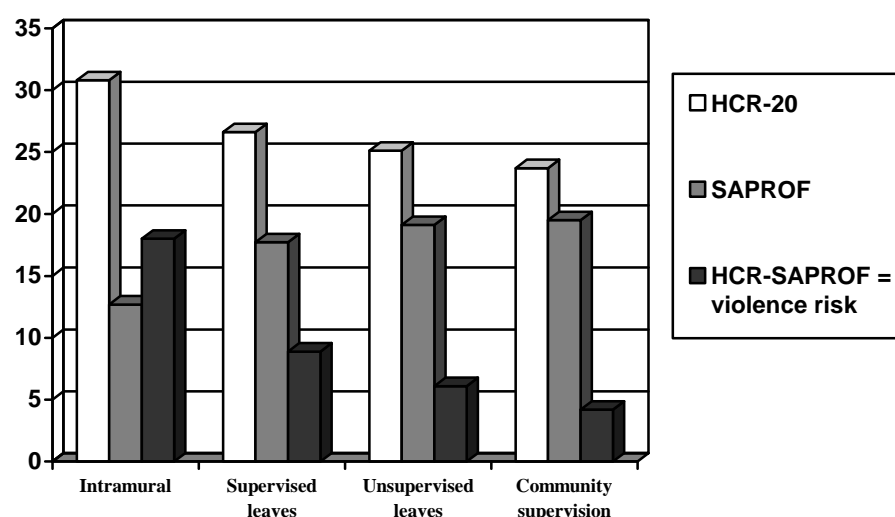
**Table 4.** Predictive validity for violent incidents during treatment for the SAPROF clinical ratings (N=315, 1 year follow-up)

Follow-up	Violent N=148 ♂	Sexual N=97 ♂	Total ♂ N=245 ♂	Total ♀ N=70 ♀	Total N=315
Total score SAPROF	.77**	.81**	.78**	.70	.77**
Total score HCR-20	.74**	.85**	.79**	.78*	.79**
Total HCR-20 – SAPROF	.81**	.84**	.82**	.76*	.81**
Final Protection Judgment	.69*	.73**	.70**	.69	.70**
Integrative Final Risk Judgment	.75**	.81**	.77**	.72*	.76**

Note. \* =  $p < .05$ , \*\* =  $p < .01$  (two-tailed). Final judgments are made on a five-point scale.

As in the retrospective file studies, the SAPROF proved dynamic and patients' protective factors showed to be increasing as treatment progressed (see Figure 3). At the same time the ratings on the dynamic Clinical and Risk management factors of the HCR-20 showed to decrease over time. Together the improved protective factors and the diminished risk factors resulted in an overall reduction in violence risk. It was found that there were far fewer incidents of violence at the later stages in treatment and that predictions of no violent incidents by the SAPROF factors were particularly good at these later stages in treatment, when patients had had more chance to build up their protective factors during treatment (De Vries Robbé & De Vogel, 2011; De Vries Robbé et al., in preparation). These results show the applicability of the SAPROF to (forensic) clinical practice and the usefulness for measuring treatment progress. Further prospective studies will have to corroborate these findings.

**Figure 3.** Risk- and protective factor total scores at different stages of treatment (N=315)



Note. Violent incident rates at the different stages of treatment: Intramural 29%; Supervised leaves 15%; Unsupervised leaves 7%; Community supervision 3%.

Summarizing, the SAPROF can be reliably coded and has demonstrated to be a strong predictor of future violence both during and after clinical treatment. These results are consistent for different types of forensic psychiatric offenders at the Van der Hoeven Kliniek. Further research in different countries and settings will have to consolidate these findings and in addition will have to focus on the usability of the SAPROF for different groups of offenders such as general offenders and offenders in outpatient treatment settings.

**The content of the coding instruction for the items has not changed, except for some minor clarifications:**

**Item 1 Intelligence:**

***‘or higher’ and ‘or lower’ were deleted from the coding box.***

***In the coding instruction more guidelines are given concerning the IQ cut-off scores that should be used for coding the item.***

**1. Intelligence**

*Coding*

<b>2</b>	Intelligence above average.
<b>1</b>	Average intelligence.
<b>0</b>	Intelligence below average.

Generally, an IQ cut-off of 1 standard deviation above or under the average has been adopted. As an IQ of 100 is generally being considered average in most intelligence tests, a score of 2 would be given when a person has an IQ of 115 or higher (above average) and a score of 0 would be given when a person has an IQ of 85 or lower (below average).

**Item 14 Intimate relationship:**

***The coding instructions now specify that ‘an extended period of time’ = at least 12 months.***

**14. Intimate relationship**

If an individual currently has an intimate relationship which has lasted for an extended period of time (at least 12 months), whether it is a marriage or not, a minimum score of 1 should be given.