

Staff observation aggression scale, SOAS: Presentation and evaluation

T. Palmstierna and B. Wistedt

Department of Psychiatry (Head: Olle Östman, M.D.)
Västerås Central Hospital and Department of
Psychiatry (Head: Börje Wistedt, M.D.), Karolinska
Institute, Danderyd Hospital, Sweden

ABSTRACT – A new psychiatric report and rating scale assessing severity and frequency of aggressive behaviour is presented and evaluated. It is based on the staff's standardized reports of aggressive incidents. By using a special aggression report form, comprehensive and standardized information is obtained, thereby permitting scoring and further analysis of different aspects of aggressive incidents. The reliability of scoring is tested and found to be good as is the scale's capacity to discriminate between different patterns of aggressive behaviour in different groups of patients. As a result of this and because of the simplicity of the scale, it is thought to be a potentially useful tool in scientific research on aggressive behaviour from psychiatric inpatients.

Received April 4, 1987; accepted for publication May 24, 1987

Aggressive and assaultative behaviour from patients in psychiatric wards is an important problem, but when trying to assess the efficiency of different remedies to prevent or treat assaultive behaviour, methodological difficulties arise. Already existing registers and reporting routines, primarily designed for purposes other than scientific research have been used (1-3), and it is also possible to assess aggression and hostility in terms of separate items on conventional psychiatric rating scales such as the Nurses' Observation Scale for Inpatient Evaluation, NOSIE (4), the Comprehensive Psychiatric Rating Scale, CPRS (5) and the Brief Psychiatric Rating Scale, BRPS (6). All these different methods of obtaining information on aggressive behaviour have disadvantages. Reporting routines to registers primarily designed for purposes other than scientific, probably vary from time to time, and interview-based rating scales, on the other hand, have the disadvantage of either relying on the patient's own subjective reports of violent and aggressive acts, or the objective behaviour of the patient during the short period of the interview. Despite the recently

pointed out importance of defining operationally and rating the seriousness of violent acts (7), very few attempts have been made to construct and evaluate rating scales, designed to assess aggressive inpatient behaviour.

The aim of this paper is three-fold. The Staff Observation Aggression Scale, SOAS is described; a reliability study is performed; and a brief clinical application is presented principally aimed at demonstrating the capacity of the scale to discriminate between patterns of aggressive behaviour among patients.

Description of the scale

The Staff Observation Aggression Scale, SOAS, is constructed to assess the degree and frequency of violent and assaultive acts from psychiatric and psychogeriatric inpatients. The scale is based on systematized reporting of aggressive events, combined with operationally defined criteria for assessing the degree of each event. The advantage of systematizing observations made by ordinary nursing staff is

obvious, they (and seldom the researcher) have the best opportunity of fully observing the patient throughout the day.

In order to make the scale easier to handle, and thereby interfere less with ordinary tasks for the nursing staff, the procedure for observing an aggressive event is separated from the procedure for scoring it. In order to standardize the reporting procedures, a specially designed report form has been constructed. The report form consists of pa-

tient identification, a short instruction to the observer and a report section (Fig.1). In such a section, it is possible to report an aggressive event and in the instruction the observer is asked to report each aggressive event by indicating at least one of the given alternatives (several are possible) in each column of a separate report section. The five separate columns of which a report section consists, explore (from left to right) five separate and consecutive aspects of aggressive events. Of these five

STAFF OBSERVATION AGGRESSION SCALE Report form When the patient shows any behaviour regarded as aggressive by observing staff, time is noted together with at least one mark in each of the five columns. All incidents of aggressive behaviour should be reported.			PATIENT IDENTIFICATION	
PROVOCATION	MEANS USED BY PATIENT	AIM OF AGGRESSION	CONSEQUENCE(S) FOR VICTIM(S)	MEASURE TO STOP AGGRESSION
<input type="checkbox"/> NO UNDERSTAND-ABLE PROVOCATION	<input type="checkbox"/> Verbal, no threat	<input type="checkbox"/> Nothing/no-body	<input type="checkbox"/> None	<input type="checkbox"/> None
PROVOKED BY:	<input type="checkbox"/> Verbal, physical threat	<input type="checkbox"/> Other object	OBJECTS:	<input type="checkbox"/> Talk with patient
<input type="checkbox"/> Other patient(s)	<input type="checkbox"/> Hand	<input type="checkbox"/> Staff member	<input type="checkbox"/> Damaged, not replaced	<input type="checkbox"/> Calmly brought away
<input type="checkbox"/> Help with ADL	<input type="checkbox"/> Foot	<input type="checkbox"/> Other patient	<input type="checkbox"/> Damaged, replaced	<input type="checkbox"/> Peroral medication
<input type="checkbox"/> Staff demanding pat to take medication	<input type="checkbox"/> Chair	<input type="checkbox"/> Other person	PERSONS:	<input type="checkbox"/> Parenteral medication
<input type="checkbox"/> Patient denied something	<input type="checkbox"/> Glass		<input type="checkbox"/> Felt threatened	<input type="checkbox"/> Held with force
<input type="checkbox"/> Other:.....	<input type="checkbox"/> Teeth		<input type="checkbox"/> Pain < 10 min	
.....	<input type="checkbox"/> Tries to strangle		<input type="checkbox"/> Pain > 10 min	
	<input type="checkbox"/> Knife		<input type="checkbox"/> Visible injury	
	<input type="checkbox"/> Other:		<input type="checkbox"/> Need for treatment	
			<input type="checkbox"/> Need for treatment by physician	
Date:	Time:	Signature:		

Fig. 1. SOAS report form consisting of patient identification, a short instruction to the observer and a report section.

aspects, the first and the last describe the immediate cause and the measures taken to stop the aggression. The three "central" aspects, i.e. "Means", "Aim" and "Result", describe the aggressive event without regard to the more or less relevant surrounding stimuli and precautions.

These "core" items are possible to score with operationally defined criteria (Appendix 1). Each of the items range from 0 to 4 points in an ordinal scale, where 0 points is defined as no means, no aim and no resulting injury/damage and 4 points is the highest possible score on each item. As is readily seen, if no means are used or there is no aim of the aggression, an observer is not able to recognize it. However, aggressive acts could of course be observed even if no injury/damage occurs. This means that the items "Means" and "Aim" will only have a range between 1 and 4 points, while the item "Result" will have a range from 0 to 4 points. Of course, one could choose scoring criteria giving the same range for all three items, but we have preferred to have scoring criteria conformed rather than have conformity in the range of points for the items.

With SOAS, numerous information concerning different aggressive events is obtained, which could be dealt with in many separate ways. In this study, global severity of an aggressive event is defined by the sum of the three items "Means", "Aim" and "Result". (The range of this sum is theoretically 0-12 points, but in reality the range will be 2-12 points). Aggressive events are demarcated into three categories, mild, moderate and severe. Where mild is defined as 2-5 points, moderate as 6-8 points and severe as at least 9 points.

Reliability of SOAS

A widely accepted numerical measure of reliability is the intraclass correlation (ICC)(8). ICC describes how well the inter-item variation (item-profile) within a rating is reproduced by different raters and also how well scoring "levels", expressed as the sum of a group of items, are reproduced by different raters. However, ICC as a measure of interrater reliability has certain limitations. Since the ICC on the inter-item profile is based on a quotient containing a measure of the variation between separate items

within a rating, it becomes more accurate, the more items the scale contains and the greater the variation is between scored items. These limitations are worth pointing out, since SOAS does not fulfill the optimal conditions for applying ICC measures of reliability. It is quite rare that two or more observers are present at the same aggressive event at the same time and it is even rarer that a greater number of events could be observed by the same set of observers. Moreover, the scoreable number of items in the scale is only three. Therefore, we have chosen to perform two different kinds of reliability studies.

Reliability study 1

To overcome the difficulties of getting a greater number of identical observers observing several aggressive events, permitting a proper use of the ICC measures, four different aggressive events were briefly and identically *described* to 12 staff members. Without any previous training, they indicated which of the alternatives given in the report form (Fig.1) they found to be most appropriate. Their "reports" were then scored and ICC measures were calculated for both item-profile and total score between ratings.

The ICC of total score between raters is very high (0.96). For three of the events, ICC of item-profile is high (0.72, 0.73 and 0.98). One event, No.2, has a surprisingly low ICC-value, 0.05. This different ICC value is not, however, explained by a great variability in the ratings of total score, but by a small variability in the total score itself (see Table 1). In event No. 2, 10 of 12 reports gave the score 2 in the item "Means", and only two reports gave the score 1 point. On item "Aim", all 12 reports gave the score 2 points, and on item "Results", nine reports gave the score 2, and only three reports gave the score 0 points. Since there is practically no variation at all between either the 12 raters, or be-

Table 1
Variation of total score between ratings in reliability study (number of reporters = 12)

Event no.	Minimum value	Maximum value	Mean
	3	5	
2	4	6	
3	8	9	
4	9	10	

Table 2
Variation of total score between ratings in reliability study 2

Event no.	Reporter no.	Scores			Total
		Means	Aim	Result	
			2		
2	2		2		
	3		3		
	4		3		
	5		2		
	6		2		
4	7		2		
	8		2	0	
	9		2		
	10		2		
	11		2	0	
	10		2	2	
	11		2	2	
6	12		2	2	7
	13		2	2	7
	11		2	2	7
	14		2	3	8
	10		2		8
8	15		3	0	5
	16		3	0	5
	15		3	0	6
9	16		3	0	6
	17		2		5
10	18		2	0	4
	4		2		5
	19		2	2	7
	11		2	2	6

tween the three items within each rating, the surprisingly low ICC in this event's item-profile is more likely to be due to the limitations of the ICC measure (according to the discussion above) than to an extremely low interreporter reliability.

Reliability study 2

During a 6-week period all aggressive events from patients on a ward for acute patients involuntarily admitted were reported. During this period, 10 aggressive events were reported by two or more staff members simultaneously and independently (Table 2). As can be seen, only five of the events were observed by at least three observers, the minimum requirement for calculating any correlations. Furthermore, in none of these events were the same set of observers present. Calculation of numerical values to describe the interreporter reliability in this part of the study would not provide any further information than could be obtained from the table itself. The variation between the ratings is very small. Re-

garding the individual items, it can be noted that in item "Aim" there is no variation at all between reports, and in items "Means" and "Results", only two and four of the individual reports differ. In no case is there more than 1 point difference between reports in an event, and in no case are there differences in more than one item of a reported event. Also in the total score, there are therefore very small differences between reports. In no event was there a difference of more than one point between reports.

Clinical application of SOAS

Material

Twenty-eight demented and not permanently bedridden patients were randomly selected from two psychogeriatric long-stay wards. From a ward for acute patients, mostly involuntarily admitted, all individuals who during the observation period were treated as inpatients were observed (Table 3). The patients were classified into three major diagnostic groups; demented patients, acute admitted schizophrenic patients and acute admitted non-schizophrenic patients. (In the tables, these groups are referred to as "Dem", "Sz" and "Nsz" respectively). Since the purpose of this part of the study was mainly to demonstrate the discriminating capacity of SOAS, the diagnoses of dementia and schizophrenia were in accordance with ICD-8 which was the clinical diagnostic practice current in Sweden at the time of the study. The study was performed in two consecutive periods of 2 and 4 weeks respectively. During these periods, staff members reported all aggressive events from the patients on the report form. During the second period, the possibility of reporting the kind of provocations was introduced.

Table 3
Characteristics of patients included in clinical application of SOAS

Category	Number of patients			Mean age
	Men	Women		
Dem				
Sz				
Nsz				

Dem = demented patients; Sz = acutely admitted schizophrenic patients; Nsz = acutely admitted non-schizophrenic patients.

Table 4
Distribution of scores in separate scoreable items in acute admitted patients (Number of scored events = 46)

Item	0 points	1 point	2 points	3 points	4 points
Means	0	14	25	7	0
Aim	0	9	31	5	0
Result	16	19	10		0

In "Means" and "Aim" it is not possible to receive 0 points.

Statistics

Severity of aggressive events is classified as described above and, as with the kind of provocation and measures taken by staff to prevent further aggression from the patient, the significance levels of the differences is calculated with chi-square tests.

To analyze the differences in frequency of aggressive behaviour, the "individual aggression frequency" is determined for each patient by calculating a quotient where the total number of aggressive events for each patient is divided by the number of observation days. Since this measure is, as will be seen below, very skewly distributed, a nonparametrical test based on ranking of all 95 patients according to their individual aggression frequency, the Kruskal-Wallis H-test, was performed (9). (This test is an extension of the Wilcoxon rank sum test, applicable when the means of three or more samples are to be tested based on ranking of the samples). The test function, Ht, with correction for ties, is approximately chi-square distributed with $df=k-1$, k =number of samples compared.

Results

During the observation periods, 150 events were reported. Of these, 124 were classified as mild, 26 as moderate, and no event was classified as severe. The distribution of scores for separate items is shown in Tables 4 and 5. As can be expected from

Table 5
Distribution of scores in separate scoreable items in demented long-stay patients (number of scored events = 104)

Item	0 points	1 point	2 points	3 points	4 points
Means	0	7	97	0	0
Aim	0	3	98	3	0
Result	56	38	10	0	0

In "Means" and "Aim" it is not possible to receive 0 points.

the scoring criteria which define separate scores from the least possible to the almost worst possible aggressive events, most events receive scores in the lower part of the range instead of being evenly distributed, since severe assaults are quite rare even within psychiatric facilities (1).

The range of individual aggression frequencies varies greatly (Table 6). Table 6 also shows that the mean values of individual aggression frequencies lies very far from the middle of the range, indicating a highly skewed distribution. The three mean values of "individual aggression frequencies" differ significantly, $P<0.05$, with the highest value for demented patients and the lowest for non-schizophrenic patients (Kruskal-Wallis H-test with correction for ties, $H_t=6,13$, $df=2$)

Differences in severity of aggressive events are also demonstrated. Not surprisingly, the demented patients have the highest proportion of mild events, and the schizophrenic patients have the highest proportion of more severe events (38% of the events from schizophrenics were classified as moderate, compared with 20% in the non-schizophrenic group and 10% in the demented group). These differences in proportions of mild and moderate aggressive events between the groups differ significantly, chi square=13.1, $df=2$, $P<0.01$.

Differences in the kind of provocation between the groups were analyzed by grouping the different kinds of provocation into three categories; events with no identifiable provocation, events preceded by staff helping the patients with activities of daily life (ADL) and events preceded by other identifiable provocation from staff or fellow patients. Since the possibility of reporting the kind of provocation was not introduced until the last 4 weeks of the study, this analysis only includes 101 reported ag-

Table 6
Frequency of aggressive events in different categories

Category	Total number of reported events	Individual aggression frequency (Events/day)	
		Range	Mean
Dem	104		
Sz	26		
Nsz	20		

The three means differ significantly from each other, $P<0.05$, when an analysis of variance based on ranks is performed (ref. 9).

Table 7
Provocation preceding an aggressive event in different groups

Category	Number of events			Total
	Without provocation	ADL-provoked	Other	
Dem	5	64		
Sz	9	1		
Nsz	2	0		

gressive events (Table 7). Even though it is not possible to perform a chi-square test because of an insufficient number of expected events in several cells of this table, it is readily seen that demented patients are mostly provoked by help with ADL (86% of the events). In the acutely admitted group of patients, it is interesting to note that schizophrenics have more unprovoked aggressive events than non-schizophrenic patients. (chi square=6.31; $P < 0.05$).

There are also differences in whether there were any measures taken by the staff to prevent further aggression. In the demented group, no measures was taken in 99 of the 104 events reported (95%), while this was the case in only 11 of the 46 events reported from the acute admitted patients (24%). (chi square=79.1; $P < 0.001$.)

In this brief application of SOAS, demented patients were mostly provoked by ADL-help from the staff and manifested highly frequent but mild aggressive behaviour which the staff did not take any immediate measure to stop. On the other hand, schizophrenic patients were mostly not understandably provoked and they manifested less frequent but more serious aggressive behaviour than the demented patients. Even though these results are in accordance with those of several other authors investigating violent behaviour from institutionalized patients (3,10,11) they should, however, be interpreted with caution, since this is a fairly small investigation and the aim of this clinical application was primarily to demonstrate the discriminating capacity of this report method.

Conclusion

The Staff Observation Aggression Scale, SOAS, seems to have both good inter-reporter reliability as well as a capacity to discriminate between different aggressive patterns with special regard to frequency, kind of provocation and severity of aggressive

events. It is worth noting that the inter-reporter reliability is shown to be good without any previous training or education of the staff on how to use the report form. Therefore, it should be easy to introduce the method in psychiatric wards without interfering with the observing nursing staff's ordinary tasks. Even though further validation of the scale and investigation of its sensitivity to change still remains, it is reasonable to assume that it could be a useful tool in scientific research on aggressive behaviour from psychiatric inpatients.

Acknowledgements

The authors wish to thank Bernard Huitfeldt, Astra ALAB, for his indispensable help in providing appropriate methods for statistical analysis. This study was in part supported by grants from the Swedish Work Environment Fund.

References

1. Ekblom B. Acts of violence by patients in mental hospitals. Stockholm: Scandinavian University Press, 1970.
2. Tardiff K, Sweillam A. Assault, suicide and mental illness. *Arch Gen Psychiatry* 1980;37:164-169.
3. Rofman ES, Askinazi C, Fant E. The prediction of dangerous behavior in emergency civil commitment. *Am J Psychiatry* 1980;137:1061-1064.
4. Honigfeld G, Gillis RD, Klett CJ. NOSIE-30: A treatment sensitive ward behavior scale. *Psychol Rep* 1966;19:180-182.
5. Åsberg M, Montgomery SA, Schalling D, Sedvall G. Comprehensive psychopathological rating scale. *Acta psychiatr scand* 1978; Suppl. 271.
6. Overall J, Gorham DR. The brief psychiatric rating scale. *Psychol Rep* 1962;10:799-812.
7. Edwards JG, Reid WH. Violence in psychiatric facilities in Europe and the United States. In: Lion JR, Reid WH, eds. Assaults within psychiatric facilities. New York: Grune and Stratton, Inc. 1983:131-142.
8. Winer BJ. Statistical principles in experimental design. 2nd edition New York: McGraw-Hill, 1971.
9. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. *J Am Stat Ass* 1962;47:583-621.
10. Tardiff K. A survey of assault by chronic patients in a state hospital system. In: Lion JR, Reid WH, eds. Assaults within psychiatric facilities. New York: Grune and Stratton, Inc. 1983:3-20.
11. Albee GW. Patterns of aggression in psychopathology. *J Consult Psychology* 1950;14:465-468.

Address

Tom Palmstierna, M.D.
Psychiatric Clinic I
Västerås Central Hospital
S 721 89 Västerås
Sweden

Appendix 1

Scoring criteria for SOAS

For each reported incident, it is possible to score the items specified below. It should be noted that it is possible to mark more than one alternative in each column of the report form. If this is the case, the item in question should be scored according to the mark giving the highest score. In the case of a missing mark in one item of a report, this is regarded as a mark for the mildest possible alternative in this column with regard to the marks of the other columns and is scored accordingly. (This means not necessarily zero, since for example if there is a mark in "Result" for a bruise, it cannot have occurred without the patient having made use of at least a hand, i.e. 2 points for "Means".) If several incidents occur from the same patient during a 60-min period, these are regarded as one incident. The scores for this kind of "multiple incident" are defined for each item by the marks giving the highest score during this 60 min-period.

Means:

- 0 points Nothing was used.
- 1 point Verbal aggression.
- 2 points The patient makes use of parts of his own body. (Teeth and attempts to strangle excluded here, see below).
- 3 points The patient makes use of ordinary and easily available objects, bites or tries to strangle someone.
- 4 points The patient makes use of objects generally regarded as dangerous in an aggressive situation, such as knives, scissors, tools, etc.

Aim:

- 0 points Nothing was aimed at. (Observe the difference to "no result"; aggression could have an aim even though no damage/injury occurred).
- 1 point Object (other than the object eventually made use of, which is scored in item "Means").
- 2 points Staff member(s).
- 3 points Other patient(s).
- 4 points Other person than above mentioned categories.

Result:

- 0 points No damage/injury.
- 1 point *Objects:* Visible damage but still possible to make use of. *Individuals:* Felt threatened or brief pain (<10 min), no visible injury.
- 2 points *Objects:* Damaged and need to be replaced. *Individuals:* Physical pain >10 min or visible injury not requiring any treatment.
- 3 points *Individuals:* Injury requiring some kind of treatment but not necessarily by a physician.
- 4 points *Individuals:* Injury requiring some kind of treatment or supervision prescribed or performed by a physician.