Reliability and validity of the interactive observation scale for psychiatric inpatients applied by nursing aides in daily ward practice

Abstract

The main characteristic of the nursing Interactive Observation Scale for Psychiatric Inpatients (IOSPI) is the necessity of interaction between raters and patients during assessment. The aim of this study was to evaluate the reliability and validity of the scale in the “real” world of daily ward practice and to determine whether the IOSPI can increase the interaction time between raters and patients and influence the raters’ opinion about mental illness. All inpatients of a general university hospital psychiatric ward were assessed daily over a period of two months by 9 nursing aides during the morning and afternoon shifts, with 273 pairs of daily observations. Once a week the patients were interviewed by a psychiatrist who filled in the Brief Psychiatric Rating Scale (BPRS). The IOSPI total score was found to show significant test-retest reliability (interclass correlation coefficient = 0.83) and significant correlation with the BPRS total score (r = 0.69), meeting the criteria of concurrent validity. The instrument can also discriminate between patients in need of further inpatient treatment from those about to be discharged (negative predictive value for discharge = 0.91). Using this scale, the interaction time between nursing aides and patients increased significantly (t = 2.93, P<0.05) and their opinion about the mental illness changed. The “social restrictiveness” factor of the opinion scale about mental illness showed a significant reduction (t = 4.27, P<0.01) and the “interpersonal etiology” factor tended to increase (t = 1.98, P = 0.08). The IOSPI was confirmed as a reliable and valid scale and as an efficient tool to stimulate the therapeutic attitudes of the nursing staff.

Introduction

The Interactive Observation Scale for Psychiatric Inpatients (IOSPI; 1) differs from previous nursing scales by presenting as its main characteristic the need of interaction between raters and patients during assessment. Previous nursing assessment scales vary from the direct observation of the specific behaviors (2-4) to the subjective integration of observations which result in a clinical judgment (5,6). Within these limits
there are scales which join direct observation with interviews or clinical judgments (7-9). The scale “Routine Assessment of Patient Progress” (10) uses interaction in some of its items, but only the IOSPI uses a standardized interaction in all of its items.

The IOSPI requires that the rater not only observe but also use interaction to obtain the relevant information for rating each of the items. In the IOSPI items related to desirable behavior expression (10 items), there are three response options: spontaneous occurrence (grade 0), occurrence after stimulation by the nursing personnel (grade 1), and non-occurrence even when stimulated (grade 2). In the items related to inadequate behaviors, assessment stimulates confrontation of these behaviors or the observation of environmental variables which favor their occurrence. In these items, the response options are absence of behavior (grade 0), presence of behavior in particular conditions or behavior susceptible to confrontation (grade 1), and constant presence of behaviors unbending to confrontation (grade 2). This instrument was designed to provide an assessment tool as well as to promote therapeutic attitudes.

It has been previously found that the IOSPI applied by nursing aides with research training has significant inter-rater reliability (interclass correlation coefficient (ICC) of the total score = 0.66). The total scores for the IOSPI showed a significant correlation with the total scores for the Brief Psychiatric Rating Scale (BPRS) filled in by psychiatrists (Pearson correlation coefficient = 0.66), meeting the criteria for concurrent validity. The criteria for predictive validity were also met, since the scores can accurately reflect changes in functioning over time and can also discriminate the patients in need of further inpatient treatment from those about to be discharged (1).

The factorial structure of the IOSPI was closely similar to that of the Nurse’s Observation Scale for Inpatient Evaluation (4), except for depressive mood (1).

The aim of the present study was to evaluate whether the IOSPI reliability and validity could be confirmed when applied in daily practice on a general hospital psychiatric ward. An additional aim was to determine whether the scale application can increase the time of interaction between the nursing aides and the patients, and influence their opinion about the mental illness.

Subjects and Methods

Subjects

Seventeen inpatients admitted to the psychiatric unit of a general university hospital over a period of two months were included in this study. There were 5 males and 12 females aged on average 33 (SD = 12.4) years. They were interviewed by a psychiatrist (J.E.C.H.) trained in the Structured Clinical Interview for DSM-III-R (11), translated and adapted to Portuguese (12), and the following diagnoses were obtained: schizophrenia (5 patients), bipolar affective disorder (8 patients), major depression (2 patients), and organic psychoses (2 patients). The patients agreed to participate in the study and the protocol was approved by the Hospital Ethics Committee.

Evaluation scales

The IOSPI consists of 16 items to be rated on a clearly defined 3-point scale indicating increasing order of inadequacy (Appendix 1). The items which could not be evaluated during the shifts were included in the option “non-evaluated” (maximum of 2 items). The total score was calculated by summing the scores of each item multiplied by the 16/number of evaluated items ratio.

The BPRS (13) has been widely used by psychiatrists to study changes in psychopathology. The instrument was the Bech et al. (14) version of the BPRS, which adds anchor points and reduces severity to five lev-
Nursing scale increased interaction between nursing aides and patients

The translation and adaptation to Portuguese of this BPRS version present significant levels of reliability and validity (15).

The OMI - Opinions about Mental Illness - (16) was filled in by the nursing aides. This scale consists of 51 items (statements) with six response points ranging from “totally agree” to “totally disagree” and has been translated and validated for Portuguese (17). The items were clustered according to seven factors which received the following denominations: authoritarianism, benevolence, mental hygiene ideology, social restrictive-ness, interpersonal etiology, mental effort etiology, and minority view.

**IOSPI raters**

The IOSPI was applied by 9 nursing aides aged on average 39.2 (SD = 6.4) years. All raters had at least a six-year experience with work in the unit, and voluntarily agreed to participate in the study. The nursing aides were trained in the IOSPI by a researcher with experience in scale application (L.J.P.). The training was completed within approximately 20 h (usually distributed over 4 shifts) and included the following: a) reading and discussing the register protocol and the scale manual with the trainer; b) joint observation of 1, 2 and 3 patients in three successive shifts by the raters and trainer and independent filling in of the scale, followed by discussion of the results. In the last joint observation, the percentage of agreement between raters and trainer was higher than 70%.

**Validity**

Once a week, all the patients were assessed by a psychiatrist (J.E.C.H.) trained in the BPRS and blinded to the results of the IOSPI scores. The evaluations were compared with the IOSPI average score of pairs of observation effected on the same days, providing 70 comparisons (concurrent validity). Reliability in filling in the BPRS was evaluated by comparing the results of the 10 interviews during the study with those obtained by a second psychiatrist (A.W.Z.). The ICC of these evaluations was 0.67. The IOSPI average score of pairs of daily observation was used to determine the cutoff score for detecting the patients in a condition to be medically discharged (predictive validity). The clinicians deciding on discharge were blind to the results of the IOSPI scores.

**Reliability**

When the training was successfully completed, 9 patients of the psychiatric unit were evaluated daily from Monday to Friday by 3 nursing aides during the morning shift (7:00 to 12:00 h) and another 3 nursing aides during the afternoon shift (12:00 to 17:00 h), in a rotation system. Each nursing aide observed up to 3 patients per shift, filling in the IOSPI by the end of the shift. Immediately after being filled in, the scales were placed in individual envelopes and evaluated at the end of the study. Whenever a patient was discharged, the one admitted to the same bed was included in the study. Over a period of two months, 17 patients were included in the study, producing a total of 273 pairs of observations. This allowed assessment of test-retest reliability. Since retest was performed by a different rater, and the interval between test and retest was short (morning and afternoon of the same day), this procedure may be viewed as an inter-rater reliability measure.

**Interaction time**

Prior to training in the IOSPI and before any information was given about the study, the interaction time between raters and patients was assessed by videotaping four sessions of a routine activity of the unit named “free activity”, for a total of 240 min. This is a nonstructured activity carried out several times a week in the morning as well as in the
afternoon, always in the same room, with occupational therapy material and recreational games. Thirty minutes before the activity, a video camera was placed always in the same position in order to obtain the widest angle shot. The participants were informed that the activity would be recorded on video tape for research purpose. During the second month of the routine application of the IOSPI, four sessions of “free activities” were again videotaped, for a total of 207 min under the same conditions as for assessment prior to any contact with the IOSPI scale. After completion of data collection for the study, the previous videotaping and those effected during the application of the scales were independently analyzed by two observers who recorded the time the nursing aides remained in the activity and the time they interacted with the patients. The operative definition for interaction was: “when the nursing aide stands by the patient’s side: verbally addressing the patient, and/or listening to a verbal communication of the patient, and/or helping the patient to perform an activity, and/or participating in a common activity with the patient”.

Before training in the IOSPI and by the end of the study, the nursing aides filled in the OMI.

**Statistical analysis**

Reliability of the IOSPI was evaluated on the basis of the ICC of the total scores independently obtained by the morning and afternoon nursing aides (18,19). In order to determine whether the reliability indices varied accordingly to the rater, the same procedure was used to compare each nursing aide’s scores with the scores of all the others who had evaluated the same patient during other periods of the day.

The concurrent validity of the IOSPI was evaluated by comparing the correlation between the weekly scores of the BPRS and the average total scores of the IOSPI on the same day using the Pearson correlation coefficient (18).

To evaluate the predictive validity of the instrument, the negative and positive predictive values (NPV and PPV) were calculated when patients in a condition to be medically discharged were detected. The patients medically discharged during a maximum period of 10 days after reaching the cutoff score were considered to be ready for discharge. To determine the cutoff score, a true positive (sensitivity) x false positive (1-specificity) curve was constructed to estimate the best point by receiver-operating-characteristic (ROC) analysis (20).

The study of the interaction time was effected using data from 6 nursing aides, since the study had been carried out under the psychiatric unit’s usual conditions of operation and, consequently, only they were present in the shifts in which the taped activities were performed before as well as during the use of the scale. The time proportion in which the aides-patients interaction during the videotaped activity occurred in relation to the time the aides were available to participate in this activity was calculated. The paired Student t-test was used to evaluate the statistical significance of the differences between proportions before and during the use of the IOSPI.

The OMI raw scores were transformed to “Sten” scores, providing a standard comparison among the seven factors of the scale in question, ranging from 1 to 10 (20). The transformed scores obtained before and during the use of the IOSPI were compared by the paired Student t-test.

**Results**

**Reliability**

The ICC value for the IOSPI total scores was 0.831 (P≤0.0001). The ICC values, obtained by comparing each nursing aide’s assessment with those of all others who as-
Nursing scale increased interaction between nursing aides and patients

sessed the same patients, ranged from 0.66 to 0.91, all of them being statistically significant (P<0.05).

Validity

A significant correlation was found between the average total score of pairs of observations of IOSPI and BPRS (r = 0.69, P<0.001). In the predictive validity study, the best cutoff point by ROC analysis was score 0, which presented sensitivity = 0.6 and specificity = 0.9. With this cutoff point, PPV = 0.57 and NPV = 0.91 were obtained.

Interaction

The interaction time recorded by the two observers presented an ICC = 0.99 (P<0.001). Figure 1 shows the interaction time proportions of nursing aides and patients in relation to the available time for each nursing aide and for the average of all assessments before and during the IOSPI application. These proportions increased significantly during the use of the scale (t = 2.93, d.f. = 5, P = 0.033, paired Student t-test).

Opinions about mental illness

The mean scores of the OMI factors transformed to “Sten” scores, obtained before and during the IOSPI application, are depicted in Table 1. During the IOSPI application, the score for the “social restrictiveness” factor was significantly reduced (t = 4.27, d.f. = 8, P = 0.003) and the score for the “interpersonal etiology” factor was increased, approaching statistical significance (t = 1.98, d.f. = 8, P = 0.08).

Discussion

In this study, the IOSPI reliability previously observed in an experimental situation (1) was maintained even during routine application by the ordinary nursing team of a general hospital psychiatric unit. Under these conditions, the IOSPI also maintained concurrent validity criteria, since its scores were significantly correlated with those obtained after psychiatric assessment using the BPRS.

The IOSPI can also discriminate those in need of further inpatient treatment since it showed a high NPV for discharge on the subsequent 10 days. The PPV was lower and this can be attributed to the fact that discharge does not depend only on remission of the symptoms, but also on the social and family conditions needed to shelter the patient after discharge. These observations allow us to consider the IOSPI as a tool that meets the predictive validity criteria, at least for the identification of patients in no condition to be discharged.

The present results also suggest that the

![Figure 1. Proportion of interaction time between inpatients and nursing aides during the videotaped activities. Lines show the results for each nursing aide and squares the average before (open squares) and during (filled squares) the IOSPI application. *P<0.05 vs before IOSPI application (Student t-test).]

<table>
<thead>
<tr>
<th>Factors</th>
<th>Mean ± SD</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Before</td>
<td>During</td>
</tr>
<tr>
<td>Authoritarianism</td>
<td>7.00 ± 1.93</td>
<td>6.67 ± 2.56</td>
</tr>
<tr>
<td>Benevolence</td>
<td>5.56 ± 1.81</td>
<td>5.33 ± 0.71</td>
</tr>
<tr>
<td>Mental hygiene</td>
<td>3.33 ± 1.41</td>
<td>3.67 ± 1.80</td>
</tr>
<tr>
<td>Social restrictiveness</td>
<td>6.89 ± 1.17</td>
<td>5.44 ± 1.42</td>
</tr>
<tr>
<td>Interpersonal etiology</td>
<td>4.78 ± 1.86</td>
<td>6.22 ± 1.09</td>
</tr>
<tr>
<td>Mental effort etiology</td>
<td>6.00 ± 1.00</td>
<td>6.33 ± 1.87</td>
</tr>
<tr>
<td>Minority view</td>
<td>6.22 ± 1.39</td>
<td>5.78 ± 1.20</td>
</tr>
</tbody>
</table>

1OMI factor raw scores were transformed to “Sten” scores ranging from 1 to 10, providing a standard comparison among the factors. *P<0.01 and +P<0.1, before vs during (paired Student t-test).
routine application of the IOSPI favors a greater interaction between nursing aides and patients, possibly because filling in the scale requires interaction. Apart from the Routine Assessment of Patient Progress (10) which uses interaction in some of its items, the IOSPI is the only nursing scale known to us that uses a standardized interaction in all of its items.

Increased interaction observed during the IOSPI application is particularly important, and can contribute to increased treatment compliance (21) creating opportunities for therapeutic interactions on the part of the nursing aides. Other proceedings to stimulate interaction, such as increased number of team participants, did not show satisfactory results (22).

The opinion of the nursing aides about mental illness was also an objective of this study. The results obtained with the OMI suggest that after the application of the IOSPI the nursing aides were less restrictive with the mentally ill and also tended to believe that mental illness arises from interpersonal experience. It is surprising that the simple application of an assessment scale can change the opinion about mental illness of those who apply it. Among the possible explanations for this observation we may suggest increased interaction with the patients and the knowledge provided by the training in the use of the scale and by the systematized observations of the patients.

In conclusion, the IOSPI proved to be a reliable and valid assessment tool, even when applied to the “real” world of daily ward practice. Besides being an assessment tool, this scale can increase the interaction time between nursing aides and patients, contributing to the stimulation of therapeutic attitudes among nursing aides.

References

19. McNeil BJ, Keeler E & Adelstein Sj
Appendix 1. Interactive observation scale for psychiatric inpatients.

Instructions: For each item, choose the grade (0, 1 or 2) which best describes what you can observe during this period. If this classification is not possible and/or new situations appear and/or the patient presents physical limitations that interfere with the execution of items 1, 2, 3 and 4, choose NE (non-evaluated) and write the reason down overleaf.

1. SELF-CARE
   (0) **Spontaneously takes care of him/herself.** Although not individually solicited, washes his/her face, bathes, keeps him/herself clean.
   (1) **Takes care of him/herself only when solicited.** Only when individually stimulated, does he/she wash his/her face, bathe, and keep him/herself clean.
   (2) **Refuses to take care of him/herself.** Even if the patient is individually stimulated, he/she does not wash the face, does not bathe, and does not keep him/herself clean.
   (NE) Non-evaluated.

2. LEAVES THE BED
   (0) **Spontaneously.** Although not individually solicited, gets up at the scheduled time.
   (1) **Only when stimulated.** Gets up at the scheduled time only after being stimulated, and may return to bed afterwards.
   (2) **Does not get up even when stimulated.** Remains in bed all the time, even after being stimulated to get up.
   (NE) Non-evaluated.

3. PARTICIPATION IN ROUTINES
   (0) **Spontaneously.** Although not individually solicited, follows the ward routines concerning individual and group activities, i.e., individual interviews, use of the dining room, operative group, and recreational activities.
   (1) **Only when stimulated.** Only when solicited and individually stimulated does he/she follow the ward routines concerning individual and group activities, i.e., individual interviews, use of the dining room, operative group, and recreational activities.
   (2) **Does not participate even when stimulated.** Even when solicited and individually stimulated, does not follow the ward routines concerning individual and group activities, i.e., refuses to participate in at least one of these activities: individual interviews, use of the dining room, operative group, and recreational activities.
   (NE) Non-evaluated.

4. PERFORMS OCCUPATIONAL ACTIVITIES
   (0) **Spontaneously.** Though not individually solicited, starts and continues to perform occupational activities.
(1) *Only when stimulated.* Needs to be individually solicited and stimulated to start and continue to perform occupational activities.

(2) *Does not perform even when stimulated.* Even when individually stimulated, does not succeed in performing the proposed activities.

(NE) Non-evaluated.

5. SHOWS INTEREST

(0) *Spontaneously.* The patient shows interest, through questions, comments, suggestions or criticism, in at least some of these activities: eating, walking, riding, TV watching, self-care, discharge from hospital.

(1) *Only when stimulated.* Only when stimulated does the patient show interest, through questions, comments, suggestions or criticism, in at least some of these activities: eating, walking, riding, TV watching, self-care, discharge from hospital.

(2) *Does not show interest even when stimulated.* Even when stimulated, does not succeed in showing interest through questions, comments, suggestions or criticism, in at least some of these activities: eating, walking, riding, TV watching, self-care, discharge from hospital.

(NE) Non-evaluated.

6. VERBAL COMMUNICATION

(0) *Spontaneously.* Although not directly solicited, initiates verbal communication with patients or technicians.

(1) *Gives verbal answers only when stimulated.* Only when directly solicited and stimulated, answers to verbal communication, but does not continue or initiate new communication.

(2) *Does not answer verbally even when stimulated.* Even when directly solicited, does not respond to communication.

(NE) Non-evaluated.

7. SOCIAL INTERACTION

(0) *Spontaneously.* Although not directly solicited, even if the routines do not concern groups, stays next to other patients and/or technicians, in the same place, and interacts with them.

(1) *Only in group routines.* When the routines concern groups, stays next to other patients and/or technicians, in the same place, and interacts with them.

(2) *Does not interact even during group routines.* When the routines concern groups, refuses to participate or stay next to other patients and/or technicians, in the same place, and does not interact with them.

(NE) Non-evaluated.

8. REFERENCE TO UNREAL FACTS

(0) *Absent.* The verbal communications are related to apparently real facts.

(1) *Suggestive communications; however, accepts confrontation.* The verbal communications seem to be related to apparently unreal facts; however, facing arguments that suggest the unreality of these facts, shows doubt or reformulates his/her communication.

(2) *Suggestive communications, does not accept confrontation.* The verbal communica-
tions seem to be related to apparently unreal facts; the patient rejects any arguments that suggest the unreality of these facts and maintains the previous communication.

(NE) Non-evaluated.

9. HALLUCINATIONS

(0) **Absent.** No clues that he/she is going to talk, smile, or gesticulate with him/herself; no mention of voices and/or visions of things that do not actually exist.

(1) **Behavior suggesting hallucinations.** Speaks, smiles or gesticulates with him/herself, apparently responding to unreal stimulation; however, denies that this occurs or that it is related to voices and/or visions of things that do not actually exist.

(2) **Comments or behaviors suggesting hallucinations.** Refers to voices and/or visions of things that do not actually exist as if they were real, presenting or not behaviors suggesting false perceptions as speaking, smiling or gesticulating to him/herself, apparently responding to unreal stimulation.

(NE) Non-evaluated.

10. ORIENTATION

(0) **Remains oriented.** Oriented as far as being in a hospital, recognizes persons and identifies them in relation to locality, moves inside this locality, identifying his/her functions.

(1) **Oscillates concerning orientation.** Sometimes recognizes the persons and the locality where he/she is, other times makes confusions about places and/or persons.

(2) **Disoriented.** Disoriented concerning persons and localities.

(NE) Non-evaluated.

11. INCREASED PSYCHOMOTRICITY

(0) **Absent.** The movements occur in normal or lessened rhythm.

(1) **Occasionally unquiet.** Stays in predetermined places; however, sometimes executes movements in continuous rhythm (for instance: rubs the hands, beats the feet, etc.) suggesting psychomotor uneasiness.

(2) **Psychomotor agitation.** Does not stay in predetermined places, showing movements in accelerated rhythm as unfinished, abrupt actions, suggesting psychomotor agitation.

(NE) Non-evaluated.

12. ACCELERATED SPEECH

(0) **Absent.** Expresses him/herself verbally in normal or diminished rhythm.

(1) **Present; however, listens to the interlocutor.** Expresses him/herself in accelerated rhythm, succeeds in stopping to listen to the interlocutor.

(2) **Present, does not listen to the interlocutor.** Expresses him/herself verbally in accelerated rhythm, does not succeed in stopping to listen to the interlocutor.

(NE) Non-evaluated.

13. IRRITABILITY

(0) **Absent or proportionate.** No signs of irritation in the absence of situations that justify it or when something displeases him/her; reacts not showing irritation and/or disproportionate hostility, being possible to detect what actually has displeased him/her.
(1) **Disproportionate with a motive.** When something displeases him/her, reacts showing great irritation and/or hostility disproportionate to the situation; however, it is possible to detect what actually has displeased him/her.

(2) **Present without a motive.** Shows irritation and/or hostility, not being possible to detect what actually has displeased him/her, independently of the intensity of the reaction. (NE) Non-evaluated.

14. **ACCEPTANCE OF LIMITS**

(0) **Present.** Easily recognizes and accepts the limits established for his/her behavior.

(1) **Accepts with difficulty.** After confrontation only, recognizes and accepts the limits established for his/her behavior.

(2) **Does not accept.** Even after confrontation, does not recognizes and does not accept the limits established for his/her behavior. (NE) Non-evaluated.

15. **SELF-ESTEEM EXPRESSION**

(0) **Positive.** The comments about him/herself and his/her life manifest a positive self-esteem.

(1) **Negative.** The comments about him/herself and his/her life refer, predominantly, to being worthless.

(2) **Negative with suicidal ideas.** The comments which are made in relation to him/herself and his/her life refer, predominantly, to being worthless, absence of plans for the future and/or manifestation of intention to end his/her own life. (NE) Non-evaluated.

16. **UNUSUAL BEHAVIOR**

(0) **Absent.** Does not show behaviors rarely observed in most of the people such as: repetition of gestures or actions, grimaces, gestures or actions without apparent reason, imitation of other people’s actions, permanence in the same position during long periods of time.

(1) **Present; however, interrupted by interaction.** Shows behaviors rarely seen in most of the people such as: repetition of gestures or actions, grimaces, gestures or actions without apparent reason, imitation of other people’s actions, permanence in the same position during long periods of time; however, the behavior is interrupted when the other person succeeds in establishing an interaction.

(2) **Present and noninterrupted by interaction.** Shows behaviors rarely observed in most of the people such as: repetition of gestures or actions, grimaces, gestures or actions without apparent reason, imitation of other people’s actions, permanence in the same position during long periods of time; however, the behavior is not interrupted when the other person succeeds in establishing an interaction. (NE) Non-evaluated.

**Observations:**