ABSTRACT – **Background and Objectives:** According to Spanish legislation, the psychological harm suffered by the victim of a criminal act is determined by assessing its impact on the victim’s mental state. Usually, the victim’s pain and suffering is estimated by administering clinical scales. The aim of the present study was to explore the effectiveness of psychopathological assessment using commonly used scales in clinical practice and whose results are presented as legal evidence in a forensic context in order to detect malingered psychological sequelae (anxiety, depression and low self-esteem) in victims of intimate partner violence in forensic contexts.

**Methods:** In the present study three scales based in a clinical setting and regularly used in a forensic context were administered (BDI, STAI and Rosenberg) to assess malingering of symptoms. The sample comprised 66 women: 36 students, and 30 real victims. The non-clinical sample was evaluated twice: the first time they gave sincere responses, and the second time they were instructed to answer as if they were victims. The real victims underwent testing in a forensic context.

**Results and Conclusions:** The results of our research show that, even without previous knowledge of the scales, people can distort the test results by malingering symptoms that are normally accepted as sequelae of intimate partner violence, especially depression and low self-esteem; however, the results for anxiety, were less homogeneous. Although these tests are used extensively in clinical psychology, our study confirms that, just by themselves, they are not a reliable source of information in a forensic context.
Introduction

According to the Spanish legislation regarding intimate partner violence that came into effect in 2004 (Ley Orgánica 1/2004), psychological violence is defined as a situation in which the victim is seen to be suffering psychological distress as a result of the crime committed. Therefore, from a legal perspective, pain and suffering (regardless of any economic compensation they may entail as evidence of civil responsibility within the criminal proceedings) need to be demonstrated. This is especially relevant nowadays, when considering that the number of criminal prosecutions for intimate partner violence (IPV) in Spain are increasing in a steady manner – up to 72.1%, from 47,262 cases in 2002 to 81,301 cases in 2007.

These cases rely on external reports, usually from public services or private organizations, to determine the existence of psychological sequelae that will be used as prosecution evidence in court. In order for this to happen, the psychological assessment must establish that the alleged offence is the cause of such sequelae.

Although the psychological harm suffered by victims of a criminal act is identified by assessing its impact on their mental state, clinical reports usually assess the presence of psychological sequelae using clinical scales that mainly measure depression, anxiety, sexual dysfunction, and similar. However, the scientific literature establishes that the presence of these disorders and/or symptoms is legally, not a sufficient proof of the existence of psychological damage, especially given its possible interaction with previous or associated disorders. Other studies on gender-based violence show as well, how the psychological sequelae caused by battering are the same as those seen in victims of other types of violent crime.

Although other conditions associated with intimate partner violence – such as depression and anxiety – can be included in the clinical diagnosis, psychological signs can only be established on a forensic level when post-traumatic stress disorder (PTSD) is present.

The diagnosis of a specific disorder is sufficient in a clinical assessment, but not in a legal-forensic context, in which the cause and effect relationship between the criminal act and the psychological sequelae must be established beyond any reasonable doubt. The possibility that the alleged symptomatology may be malingered must also be taken into account.

We can assert that PTSD is the primary disorder in intimate partner violence cases. On a secondary level IPV is associated with trauma, depression, anxiety and other symptomatologies. Some authors have even established that this secondary trauma, observed in the absence of PTSD, cannot be attributed to the sequelae of a traumatic event. In spite of this, in Spanish forensic practice, IPV normally tends to be accepted as a feasible hypothesis, especially in the absence of other contradictory forensic evidence or a more accurate assessment.

The ways of assessing malingered in its various forms in forensic practice have been studied in depth. In our case, the most relevant aspect is the exaggeration of symptomatology, as it gives credence to the claims made by victims of intimate partner violence. It also enables the prosecution to establish a moncausal link between the alleged crime and the sequelae detected. In other words, the legal system itself acts as an external incentive, which is an indicator of malingered suspicion.

Studies have shown that the prevalence of malingered fluctuates depending on the offence type, occurring in approximately one
sixth of forensic cases\textsuperscript{19}, even though in cases of assault the figure rises up to a 50\%\textsuperscript{20}.

One of the main limitations of conducting a forensic assessment based on questionnaires relies in the characteristics of the population involved. The absence of specific instruments in this field only allow us to use it as complementary evidence\textsuperscript{21}, despite the efforts made to construct more sophisticated measurement systems\textsuperscript{10,12}.

It should also be noted that motivation plays a key role in the different uses of questionnaires depending on the context in which they are used (clinical or legal). In a legal context, for example, the answer a respondent gives may influence the ruling and may obtain scores that are more in line with his/her wishes and needs. If this tactic were successful, it would raise doubts about the tests’ validity and reliability in a forensic context\textsuperscript{22}.

All this means that courts often have to assess the psychological evidence from a clinical point of view and not a socio-legal one; that is, the disorders will be assessed secondarily solely, due to the difficulties faced when linking them to the reported crime or another stressful event. In addition, courts tend to underestimate or ignore the possibility that the victim can be malingering symptoms in order to gain an advantage in the judicial proceedings.

Our study had two objectives: first, to assess the ability of subjects who had not been victims of gender-based violence to mangle psychological distress; and second, to analyse whether the Beck Depression Inventory (BDI), State-Trait Anxiety Inventory (STAI) and Rosenberg Self-Esteem Scale, tests that are commonly used in clinical practice and whose results are presented as legal evidence in a forensic context, can discriminate properly real victims from malingers.

### Material and Methods

#### Participants

The sample was composed of 66 subjects (mean age 27.06 years, range 20-48), divided into two groups. The control group consisted of 36 female university students selected via a purposeful sampling process. None had been victims of IPV. The average age of the control group was 22.39 years (age range: 20-34).

The forensic experimental group was initially made up of 91 women (reduced to a final sample of 30 following the inclusion criteria described below, mean age 32.67 years, range 20-48), all victims of domestic violence.

#### Measures

All subjects answered three clinical tests regularly administered by public health professionals in forensic practice to determine psychological sequelae, which measure anxiety, depression and low self-esteem.

The State-Trait Anxiety Inventory (STAI)\textsuperscript{23} is composed of two subscales –state and trait– each with 20 items, enabling the self-assessment of anxiety as a transitory (S/A) or an underlying trait (T/A).

The inventory’s test-retest reliability coefficient is high (0.81) on the trait anxiety scale and rather low (0.40) on the state anxiety scale\textsuperscript{4}; the inventory also has a high internal consistency, with alpha coefficients of 0.91 and 0.94 These results are supported by a study conducted on patients suffering anxiety disorders and other pathologies\textsuperscript{24}.

The Beck Depression Inventory (BDI) was administered in the Spanish version\textsuperscript{25}, based on the US version\textsuperscript{26}. 
The BDI has been applied in very different cultures and countries and has repeatedly shown its usefulness as well as clinical and psychometric validity\(^27\). It has a test-retest reliability of 0.69 to 0.90, satisfactory validity of 0.62 to 0.66 and a reliability coefficient of 0.93, as measured by the split-half method\(^4,28,29\). Other recent studies\(^{29-31}\) support the inventory’s psychometric qualities and linguistic adaptatability.

Devised in 1965, the *Rosenberg Self-Esteem Scale* is a 10-item scale that assesses up to what degree people are satisfied with and accept themselves\(^32\).

The original version obtained a test-retest reliability of between 0.82 and 0.88, and internal consistency of various samples of 0.77 to 0.88. A number of studies support the appropriateness of its psychometric characteristics in other languages\(^33,34\). It has also shown an internal consistency of 0.87 and satisfactory test-retest reliability (0.72 for two months and 0.74 for a year)\(^34\). As for the instrument’s validity, the authors show a correlation between the Rosenberg Self-Esteem Scale and the interpersonal sensitivity dimension of the Symptom Checklist-90 Revised (SCL-90-R).

### Procedure and design

The study was conducted between 2008 and 2009. The tests were administered to both groups in the course of the first year, and in the second year the cases of violence in the experimental group were confirmed by the definitive legal sentences (after all appeals had been exhausted).

The control group was composed of legal psychology students from the University of Barcelona. On the second day of class they were asked to participate “in a research related to forensic practice in legal psychology”. After the tests were administered the students were informed of the study’s nature. They were asked to consent to the anonymous use of their data for research purposes, used in a manner consistent with the guidelines of the World Health Organization and the declaration of Helsinki, and were given the opportunity as well, to withdraw from the study if they wished to do so. Three subjects declined participation at the beginning, with no further drop-outs. Five participants were eliminated due to errors in the completion of the medical records either on the first or second test administration.

Each member of the control group answered each of the three tests (BDI, STAI and Rosenberg), on two occasions: First, they answered under the standard condition (honest answers) while on the second occasion, they answered under malingering condition (instructed to malinger answers). A counterbalance of the administration was performed.

The instructions they were given are the following:

>“Imagine you can obtain great benefits by giving false answers in the test and suggesting that you have been a victim of intimate partner violence”

Subjects had no free time between the first and second administration, nor were allowed contact with the other participants. No practice time was allowed assuming that participants had previous knowledge of the contents evaluated from previously conducted studies. The different tests were administered and corrected following the standard application procedures suggested by their authors.

The counterbalance technique was applied in order to neutralize the possible progressive error effect (the sum of positive and negative effects of the first tests on the later
ones) generated in those cases in which each subject undergoes all the experimental conditions. To avoid these progressive error effects, each subject or group of subjects underwent the treatments in a different order.

In the forensic experimental group, the tests were administered by the forensic teams at Girona Criminal Courts while the criminal hearings were in progress. Subjects were asked to give their consent in order for the results of the assessment to be used for research purposes, respect their right to privacy, and used in a manner consistent with the guidelines of the World Health Organization and the declaration of Helsinki. No records were rejected or eliminated on these grounds. As an external validity criterion, the only chosen cases were those in which –after the psychological assessment– a conviction that was not based exclusively or mainly on the proof of psychological damage as the State’s evidence, but in which it was only referred to as evidence for civil liability.

As a result, 38 subjects were eliminated during the first and second stages of the proceedings because psychological damage was stated as the prosecution’s proof. A further 20 subjects were eliminated after the court made a complete or partial judgement in favour of the defence, and three more because the records were not completed properly.

In those cases where further forensic assessment was required it was done so after the three tests of our study had been administered.

Version 15 of the SPSS was used to analyse the obtained data. The descriptive results of each group in each of the administered tests are given first.

In order to assure the normal distribution of variables, the Shapiro-Wilk test was performed. Due to the normal distribution of data, Student’s t-test for repeated measures was used to compare the real and feigned responses of the control group and Student’s t-test for independent variables was used to compare the control group and the experimental group.

Results

Descriptive details of the forensic experimental group

On the BDI, the forensic sample obtained a mean score of 36.9 (range = 19-56; SD = 8.29), equivalent to severe depressive symptomatology.

For state anxiety (STAI-S) the mean score was 42.47 (range = 26-57; SD = 8.03), situated in the 95th percentile (very high anxiety), while the trait anxiety (STAI-T) mean score was 28.23 (range = 18-40; SD = 4.84), equivalent to the 65th percentile (medium-high).

The mean score on the Rosenberg Self-Esteem scale was 18.30 (range = 8-25; SD = 5.02), a level considered discriminant and below the cut-off point of 25, therefore showing low self-esteem (table 1).
Table 1
Overall results for the sample

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Min – Max</th>
<th>Shapiro-Wilk test*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forensic Group (victims) (N = 30)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDI</td>
<td>36.90</td>
<td>8.29</td>
<td>19 – 56</td>
<td>ns</td>
</tr>
<tr>
<td>STAI-S</td>
<td>42.47</td>
<td>8.03</td>
<td>26 – 57</td>
<td>ns</td>
</tr>
<tr>
<td>STAI-T</td>
<td>28.23</td>
<td>4.84</td>
<td>18 – 40</td>
<td>ns</td>
</tr>
<tr>
<td>ROS</td>
<td>18.30</td>
<td>5.02</td>
<td>8 – 25</td>
<td>ns</td>
</tr>
<tr>
<td><strong>Control Group (students) (N = 36)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BDI real</td>
<td>5.03</td>
<td>5.27</td>
<td>0 – 23</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>BDI malingered</td>
<td>44.44</td>
<td>8.44</td>
<td>27 – 59</td>
<td>ns</td>
</tr>
<tr>
<td>STAI-S real</td>
<td>25.72</td>
<td>5.95</td>
<td>9 – 39</td>
<td>ns</td>
</tr>
<tr>
<td>STAI-S malingered</td>
<td>25.25</td>
<td>5.20</td>
<td>16 – 38</td>
<td>ns</td>
</tr>
<tr>
<td>STAI-T real</td>
<td>26.50</td>
<td>5.05</td>
<td>17 – 37</td>
<td>ns</td>
</tr>
<tr>
<td>STAI-T malingered</td>
<td>29.00</td>
<td>6.69</td>
<td>11 – 40</td>
<td>ns</td>
</tr>
<tr>
<td>ROSENBERG real</td>
<td>31.50</td>
<td>4.86</td>
<td>23 – 40</td>
<td>ns</td>
</tr>
<tr>
<td>ROSENBERG malingered</td>
<td>18.06</td>
<td>4.04</td>
<td>10 – 26</td>
<td>ns</td>
</tr>
</tbody>
</table>

*df = 36.

Describe the details of the control group

a) Real situation

The mean scores obtained in the three tests by the control group subjects were within the normal range. The mean BDI score was 5.03 (range = 0-23; SD = 5.27), indicating the absence of depressive symptomatology.

The mean STAI-S score was 25.72 (range = 9-39; SD = 5.95), indicating a medium-high level of state anxiety. The mean STAI-T score of 26.5 (range = 17-37; SD = 5.05) indicates normality.

Finally, the mean score of 31.50 (range = 23-40; SD = 4.86) on the Rosenberg Self-Esteem Scale was also considered normal.

b) Malingered situation

The scores obtained when the subjects feigned their responses were generally higher and indicated the presence of disorder, except in the case of the STAI-S score, which remained in a medium-high level.

The BDI score was 44.44 (range = 27-59; SD = 8.44), indicating extreme depression.

The mean STAI-S score was 25.25 (range = 16-38; SD = 5.20), which indicates a medium-high level of state anxiety. The mean STAI-T score of 29.0 (range = 11-40; SD = 6.69) indicates medium-high anxiety.

Finally, the mean score of 18.06 (range = 10-26; SD = 4.04) obtained on the Rosenberg Self-Esteem Scale denoted low self-esteem.
Comparison between real and malingered situations in the control group

The scale for measuring depression, the BDI, revealed clear differences between the situations: in the real situation subjects showed no depressive symptoms, but in the malingered situation they presented extreme depression. The statistical tests applied show significant differences between situations (Table 2).

There was a very small difference between anxiety scores (STAI-S) which were not statistically significant. The state anxiety level in both conditions was situated in the 65th percentile (medium-high).

As far as the trait anxiety (STAI-T) is concerned, when the subjects responded truthfully they were in the percentile 55 (normal), whereas in the feigned situation they moved to the 65th percentile (medium-high anxiety). The difference was not statistically significant (Table 2).

Finally, self-esteem based on true responses was normal, but low when the responses were malingered. The difference between the two situations was significant (Table 2).

<table>
<thead>
<tr>
<th>Table 2 Control group: comparison between malingered and real conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control group</td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>BDI real-malingered</td>
</tr>
<tr>
<td>STAIS real- malingered</td>
</tr>
<tr>
<td>STAIT real- malingered</td>
</tr>
<tr>
<td>ROSENBERG real- malingered</td>
</tr>
</tbody>
</table>

Comparison between the control group and the forensic experimental group

Comparing the forensic group with the control group in the real condition, the BDI showed a difference in mean scores of 31.87 points. In the malingered condition the difference was also significant, but the mean score in the control group was 7.54 points higher than in the battered women. In other words, the malingered subjects exaggerated to such extent that they exceeded the level of severe depression and reached extreme depression. The Student’s t-tests confirmed the statistical significance of these differences.

The results for state anxiety were inconsistent. The control group’s scores were significantly lower than those of the forensic group when they gave true responses: (25.72 vs. 42.47). However, the difference remained when they were asked to feign their responses, and their mean STAI-S scores were still lower than those of the forensic group (25.25 vs. 42.47). The differences were again significant.

Regarding trait anxiety, however, no significant differences were obtained in any of the comparisons. The control group’s mean score in the real situation was 26.50, which was very similar to that of the forensic group (28.23) and the difference was not statistical-
ly significant. Nor was there a large increase when the control group malingered its responses: the mean score of 29 did not differ significantly from that of the forensic group.

The scores on the self-esteem scale presented significant differences between the forensic and control groups in the real situation: 18.30 vs. 31.50 respectively. However, in the malingered situation, control scores were similar to those of the forensic group (18.06 vs. 18.30) and were not statistically significant (Table 3).

Table 3
Comparison of averages between the forensic group and the control group (real and malingered conditions)

<table>
<thead>
<tr>
<th>Test</th>
<th>CR vs. F</th>
<th>CM vs. F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDI</td>
<td>-18.201</td>
<td>3.649</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>STAI-S</td>
<td>-9.454</td>
<td>-10.448</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>STAI-T</td>
<td>-1.415</td>
<td>0.523</td>
<td>ns</td>
</tr>
<tr>
<td>ROSENBERG</td>
<td>10.818</td>
<td>-0.219</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

F = Forensic group; CR = Control group’s true responses; CM = control group’s malingered responses.

Discussion

Although the presence of gender-based violence often induces clinical symptomatology, the assessment of any of the constructs that may be studied should be standardized to ensure that the forensic conclusions are correct and that the clinical treatment is a suitable one\textsuperscript{35,36}. Generally, one would expect forensic subjects of domestic violence to obtain higher scores than normal subjects on pathology tests\textsuperscript{37}. Indeed, our study shows that, when responding truthfully, the normal population showed fewer symptoms.

When subjects were asked to malingering pathology associated with intimate partner violence, we expected that the forensic subjects, who knew the aim of the assessment, would have higher scores than normal population. This study shows how –even though unfamiliar with the tests administered– subjects were able to falsify scores in order to feign depression and loss of self-esteem, although not in the case of anxiety-trait.

Over half the items in the BDI refer to verbal or cognitive attitudes and symptoms, while only a small percentage refer to the affective traits of depression. For this reason some authors recommend that it should not be used by itself to assess depression\textsuperscript{38}. Our results show that this is one of the tests with the highest scores in the malingering condition, as it can easily be manipulated in cases of IPV. However, when subjects malingered their responses, the obtained scores were significantly higher than expected.
This may be partly due to the fact that the test was not designed to diagnose depression (understood from a nosological perspective) and because an assessment of depression should be made with different instruments applied concurrently and the diagnosis made via a clinical interview.

The study also showed that the Rosenberg Self-Esteem Scale was unable to detect when people were malingering the symptoms of IPV victims. This test, therefore, does not distinguish between real symptomatology induced by criminal victimization and symptomatology intentionally manipulated in order to obtain an advantage in the courts. The Rosenberg scale is the easiest one to manipulate: the feigned responses obtained identical scores to those of the clinical sample. In the case of the BDI scale, feigned responses can be detected because the subjects obtained extreme scores in the simulation condition, while the clinical sample obtained only moderate scores.

The STAI scale represents symptoms which may be more difficult for the general population to malingering in cases of intimate partner violence. For trait anxiety, both the truthful control and forensic group scores were high and the control group’s score did not increase when responses were malingered. However, for state anxiety, the control group’s scores were low and showed no change when the responses were malingered.

To sum up, the results show that the psychological sequelae shown by subjects malingering gender-based violence include depressive symptoms, low self-esteem, and average levels of trait and state anxiety. In contrast, the forensic sample presented severe depression, extremely high state anxiety, average trait anxiety and very low self-esteem.

Although all the tests used in the study are considered valid and reliable in a clinical context, due to the fact that the items clearly define the construct to be assessed, it is precisely for this reason that they are easier to manipulate, and this reduces their validity and reliability in a forensic context.

As mentioned above, the psychological motivations when responding assessment instruments vary depending on whether the assessment is for legal or clinical purposes. This difference has a significant effect on responses to the tests and therefore makes them less useful in a legal context.

The results of our study show that special attention is required when clinical tests are used for the forensic assessment of intimate partner violence victims. A more detailed analysis of the extent to which the tests designed for use in clinical practice can be manipulated when used in a forensic context to assess the sequelae brought on by intimate partner violence is needed.

The main limitations of our study should be overcome in future studies, especially those regarding the control group, thus legal psychology students have more knowledge about mental disorders in clinical and forensic settings, and therefore might be able to malinger better than the general population. Furthermore, the size of the sample should be increased and the results must be generalized to other culture groups. Finally, tests that contain control items should be administered as well in future studies in order to assess malingering with a greater precision.

References


Address for correspondence:
Miguel Ángel Soria
Departament of Social Psychology
University of Barcelona
Edifici de Ponent, Pg. Vall d’Hebron, 171
08035 Barcelona (Spain)
Tel. +34 933125189
Fax +34 934021366
E-mail: msoria@ub.edu