Shorter communication

Psychometric properties of the Impact of Event Scale— Revised

Mark Creamer [a,*], Richard Bell [b], Salvina Failla [b]

[a] *Australian Centre for Posttraumatic Mental Health, ARMC Repat Campus, Waterdale Road, Heidelberg, Victoria 3081, Australia*
[b] *University of Melbourne, Melbourne, Australia*

Abstract

This study investigated the psychometric properties of the Impact of Event Scale—Revised (IES-R) in two samples of male Vietnam veterans: a treatment-seeking sample with a confirmed posttraumatic stress disorder (PTSD) diagnosis ($N = 120$) and a community sample with varying levels of traumatic stress symptomatology ($N = 154$). The scale showed high internal consistency (alpha = 0.96). Confirmatory factor analysis did not provide support for a three-factor solution corresponding to the three subscales of intrusion, avoidance, and hyperarousal. Exploratory factor analysis suggested that either a single, or a two-factor solution (intrusion/hyperarousal and avoidance), provided the best account of the data. However, correlations among the subscales were higher in the community sample than in the treatment sample, suggesting that the IES-R may be sensitive to a more general construct of traumatic stress in those with lower symptom levels. The correlation between the IES-R and the PTSD Checklist was high (0.84) and a cutoff of 1.5 (equivalent to a total score of 33) was found to provide the best diagnostic accuracy.
© 2003 Elsevier Ltd. All rights reserved.

Keywords: PTSD; IES-R; PCL; Veterans

1. Introduction

The Impact of Event Scale (IES; Horowitz, Wilner, & Alvarez, 1979) is probably the most widely used self-report measure in the field of traumatic stress. Published prior to formal

---

* Corresponding author. Tel.: +61-3-9496-4329; fax: +61-3-9496-2830.
  E-mail address: markcc@unimelb.edu.au (M. Creamer).

recognition of posttraumatic stress disorder (PTSD), the two subscales of the IES reflected Horowitz's views on the core phenomena of traumatic stress reactions: intrusion (B criteria in the DSM-IV PTSD diagnosis; American Psychiatric Association, 1994) and avoidance (C criteria). The scale asks respondents about the frequency with which each symptom has occurred over the past week. A scoring scheme with intervals of 0, 1, 3, and 5 was adopted for responses of "not at all", "rarely", "sometimes" and "often". Scores range from 0 to 35 for intrusion, 0 to 40 for avoidance, and 0 to 75 for the total IES.

Despite widespread use of the IES, the scale is now compromised since it does not cover the third major symptom cluster of PTSD, persistent hyperarousal. To address this deficit, a revised version of the scale (Impact of Event Scale—Revised; IES-R), including six hyperarousal items, was published in 1997 (Weiss & Marmar, 1997). The authors attempted to maintain comparability with the original IES and only minimal changes were made to the original intrusion and avoidance items. First, the original sleep item was split into two: "I had trouble staying asleep" (on the intrusion subscale) and "I had trouble falling asleep" (on the hyperarousal subscale). An additional question was added to the intrusion subscale to tap flashbacks ("I found myself acting or feeling like I was back at that time"), such that both subscales in the IES-R contain eight items. Six new hyperarousal items target sleep, irritability, concentration, hypervigilance, startle response, and physiological arousal. Instructions were modified such that respondents are asked about distress caused by the symptom, rather than the frequency. The response format was modified to a 5-point, 0–4 response format with equal intervals. Finally, the authors recommend that the scoring system be modified: rather than computing a sum of the item scores, they suggest using the mean of non-missing items.

Weiss and Marmar (1997) report psychometric data from two samples: emergency personnel exposed to a freeway collapse ($N = 429$ at Time 1 and $N = 317$ at Time 2) and workers from the 1994 Northridge earthquake ($N = 197$ at Time 1 and $N = 175$ at Time 2). The IES-R showed high internal consistency, with coefficient alphas ranging from 0.87 to 0.92 for intrusion, 0.84 to 0.85 for avoidance, and 0.79 to 0.90 for hyperarousal. Test–retest correlation coefficients ranged from 0.57 to 0.94 for intrusion, 0.51 to 0.89 for avoidance, and 0.59 to 0.92 for hyperarousal. A principal components factor analysis with varimax rotation revealed a strong single factor accounting for 49% of the variance. The authors suggest an explanation for that result may be that not all subjects were experiencing high, or even medium, symptom levels. Subscale correlations for that sample were 0.74 for intrusion with avoidance, 0.87 for intrusion with hyperarousal, and 0.74 for avoidance with hyperarousal.

Weiss and Marmar (1997) conclude their chapter by noting that, although the DSM criteria for PTSD are organized into three clusters of intrusion, avoidance, and hyperarousal, empirical validation of those clusters remains to be adequately documented. This is of interest not only in terms of psychometrics, but also in terms of the diagnostic formulation of PTSD. A further important issue is that of differential symptom patterns in individuals with low symptom levels compared to individuals with serious psychopathology such as PTSD.

Since those initial reports, the IES-R has been adopted as a measure of traumatic stress in several studies (e.g. Cusack & Spates, 1999; Pfefferbaum et al., 2000) and has been translated into several other languages including Japanese (Asukai et al., 2002), and Spanish (Baguena et al., 2001). Surprisingly little, however, has been published on the psychometric properties and construct validity of the scale. The current study investigated psychometric properties of the

IES-R in two samples of male Vietnam veterans: a treatment-seeking sample with a confirmed PTSD diagnosis and a community sample with a range of traumatic stress symptomatology.

2. Method

2.1. Paricipants and procedure

The community sample was recruited through an ex-service organization, with sub-branches distributing questionnaires to their members. A covering letter explained the nature of the research, assurances of confidentiality, and contact numbers for counseling. Participants returned the questionnaires anonymously in a pre-paid envelope. A total of 541 surveys were mailed to the sub-branches and 159 were returned. Clearly, this represents a volunteer sample and it is not possible to comment on the extent to which this group is representative of the total possible sample. The mean age of the community sample was 51.1 years (SD = 4.4) and 83% were married or living with a partner. Almost half (48%) were in paid employment.

The treatment sample comprised 124 consecutive admissions to three hospital-based PTSD treatment programs for Vietnam veterans. The IES-R was administered at the point of admission. The mean age of the treatment sample was 49.4 years (SD = 3.9) and 75% were married or living with a partner. Only 17% were in paid employment. Thus, the treatment sample was slightly younger than the community sample ($t(259) = 3.18$, $p < 0.01$) and much less likely to be working ($\chi^2(1) = 29.46$, $p < 0.001$). The difference in marital status between the two groups was not significant ($\chi^2(1) = 2.75$, ns).

2.2. Materials

In order to provide a measure of construct validity, veterans in both samples completed the PTSD Checklist (PCL; Weathers, Litz, Herman, Huska, & Keane, 1993). This 17-item scale was designed to assess the DSM-IV symptoms of PTSD and has shown good psychometric properties in Vietnam and Persian Gulf veterans (Weathers et al., 1993), as well as motor vehicle accident and sexual assault victims (Blanchard, Jones Alexander, Buckley and Forneris, 1996). A cut-off score of 50 is recommended as a good predictor of a PTSD diagnosis.

3. Results

Missing data resulted in the exclusion of nine cases (five from the community sample and four from the program sample), leaving a total of 274. Mean scores on each of the symptom measures are shown in Table 1. The IES-R demonstrated high internal consistency for the total scale (Cronbach's alpha = 0.96), as well as for the three subscales (intrusion: 0.94; avoidance: 0.87; hyperarousal: 0.91). Correlations among the subscales tended to be higher for the community sample (intrusion/avoidance = 0.81; intrusion/arousal = 0.87; avoidance/arousal = 0.84) than for the treatment sample (intrusion/avoidance = 0.53; intrusion/arousal = 0.82; avoidance/arousal = 0.52).

Table 1
Mean IES-R and PCL scores for treatment program and community samples

	Total sample ($N = 274$)		Treatment sample ($N = 120$)		Community sample ($N = 154$)	
	Mean	SD	Mean	SD	Mean	SD
IES-R total	2.17	1.00	2.64	0.69	1.82	1.05
Intrusion	2.17	1.09	2.72	0.77	1.75	1.11
Avoidance	1.90	1.00	2.30	0.80	1.59	1.03
Hyperarousal	2.55	1.14	2.99	0.85	2.21	1.22
PCL total	58.75	16.24	65.72	10.14	53.36	17.96
Intrusion	15.21	5.64	18.09	3.94	12.98	5.76
Avoidance	24.17	7.79	26.67	4.79	22.24	9.02
Arousal	19.37	4.70	20.96	3.37	18.14	5.19

3.1. Confirmatory factor analysis of the IES-R

A confirmatory factor analysis was carried out with Lisrel 8 using maximum likelihood esti-mation. Factor loadings for the model, shown in Table 2, are deceptively high. However, fitting the hypothesised model (with correlated factors) gave a chi-square (on 206 df) as 874.34 ($p < 0.001$); that is, the model did not fit the data. The less stringent assessment of fit, the

Table 2
Confirmatory factor analysis: factor loadings ($N = 274$)

IES-R item	Content	Scale	Intrusion	Avoidance	Arousal
1	Any reminders brought back feelings about it	In	0.86		
2	I had trouble staying asleep	In	0.74		
3	Other things kept making me think about it	In	0.85		
4	I felt irritable and angry	Hy			0.80
5	I avoided letting myself get upset...	Av		0.52	
6	I thought about it when I didn't mean to	In	0.83		
7	I felt as if it hadn't happened or wasn't real	Av		0.59	
8	I stayed away from reminders about it	Av		0.81	
9	Pictures about it popped into my mind	In	0.87		
10	I was jumpy and easily startled	Hy			0.85
11	I tried not to think about it	Av		0.82	
12	I was aware that I still had a lot of feelings...	Av		0.70	
13	My feelings about it were kind of numb	Av		0.70	
14	I found myself acting or feeling like I was back...	In	0.81		
15	I had trouble falling asleep	Hy			0.79
16	I had waves of strong feelings about it	In	0.92		
17	I tried to remove it from my memory	Av		0.84	
18	I had trouble concentrating	Hy			0.86
19	Reminders caused me to have physical reactions...	Hy			0.86
20	I had dreams about it	In	0.85		
21	I felt watchful and on guard	Hy			0.83
22	I tried not to talk about it	Av		0.83	

Adjusted Goodness-of-Fit Index (AGFI), was 0.73—also not satisfactory. The model obtained also showed extremely high correlation among the factors (intrusion–avoidance = 0.82; intrusion–hyperarousal = 0.93; avoidance–hyperarousal = 0.87) indicating strong relationships between these three dimensions. The modification indices (indicating specific poor fit) suggested that item 12 was particularly poorly fitting (although not responsible for the general poor fit of the model).

It has been suggested that the factor structure of both the original IES and the revised version may differ as a function of levels of symptom severity (Weiss & Marmar, 1997). This appeared to be true in the present instance, since a comparison of the correlation matrices from the two subsamples showed significant differences. However, when both samples were factor analysed with the same pattern of loadings (using the multi-sample function of Lisrel), comparable loadings were obtained for most items, indicating that this was not the source of the failure of fit. (A copy of the two solutions may be obtained from the authors.) Interestingly, while correlations between the intrusion and hyperarousal factors were high for both samples (0.94 and 0.93), correlations between intrusion and avoidance were higher for the community sample than for the clinical sample (0.86 and 0.62, respectively), as were correlations between hyperarousal and avoidance (0.93 and 0.65, respectively). Exploratory factor analysis was then undertaken in order to determine whether alternative models might equally well fit the data.

3.2. Exploratory factor analysis of the IES-R

A principal components analysis with varimax rotation was performed on the IES-R data and factor loadings are shown in Table 3. While observation of the scree plot suggested that a single factor provided the best account of the data, two factors had eigenvalues greater than one. Since a three-factor solution is conceptually consistent with the three subscales of the IES-R, and the eigenvalue of the third factor was only marginally less than one (0.99), loadings for a one-, two-, and three-factor solution are all shown in Table 3.

Consistent with the findings of Weiss and Marmar (1997), a single factor accounted for a substantial proportion (56%) of the variance. In the two-factor solution, this factor included all the intrusion and hyperarousal items, as well as one of the avoidance items (item 12: "I was aware that I still had a lot of feelings about it, but I didn't deal with them"). This is a complex question, since it has two components. While the latter—a failure to 'deal with' feelings fits conceptually with avoidance, the former—awareness of painful feelings—may be better interpreted as intrusion. Interestingly, item 12 was identified also in the confirmatory analysis as being a particularly poor fit. The second factor in the two-factor solution contained all the remaining avoidance items. In the three-factor solution, active and passive avoidance loaded on different factors, with the more passive, dissociative symptoms (numbing, derealization, emotional detachment) loading on the third factor. Interestingly, another type of dissociative experience (Item 14: flashbacks) also showed a secondary loading on this third factor.

3.3. Concurrent validity and diagnostic accuracy

The extent to which the IES-R is measuring the same construct as an alternative measure of traumatic stress was investigated by comparing IES-R with PCL total and subscale scores. The

Table 3
Single-, two-, and three-factor solutions (principal components, varimax rotation: $N = 274$)

IES-R item	Content	Scale	Single-factor	Two-factor		Three-factor		
16	Waves of strong feelings	In	0.87	0.82		0.81		
03	Other things making me think about it	In	0.80	0.80		0.79		
01	Reminders brought back feelings	In	0.81	0.78		0.77		
20	Had dreams	In	0.80	0.76		0.74		
09	Pictures in head	In	0.82	0.75		0.73		
02	Trouble staying asleep	In	0.72	0.73		0.72		
06	Thought about it when didn't mean to	In	0.80	0.74		0.72		
15	Trouble falling asleep	Hy	0.74	0.73		0.72		
19	Physical reactions	Hy	0.82	0.73		0.72		
21	Watchful and on guard	Hy	0.78	0.71		0.69		
04	Irritable and angry	Hy	0.79	0.70		0.69		
14	Acting or feeling like was back there	In	0.77	0.70		0.69		0.44
10	Jumpy and startled	Hy	0.80	0.67	0.44	0.65		
18	Trouble concentrating	Hy	0.81	0.65	0.49	0.62	0.43	
12	Aware of feelings, didn't deal with	Av	0.76	0.64	0.40	0.63		
11	Tried not to think about it	Av	0.72		0.73		0.78	
22	Tried not to talk about it	Av	0.69		0.71		0.75	
17	Tried to remove from memory	Av	0.73	0.41	0.70		0.74	
08	Stayed away from reminders	Av	0.70		0.76		0.68	
13	Feelings numb	Av	0.59		0.54			0.71
07	As if hadn't happened, wasn't real	Av	0.51		0.62			0.70
05	Avoided letting self get upset	Av	0.49		0.60			0.62
	Eigenvalue		12.30	12.30	1.35	12.30	1.35	0.99
	% Variance		55.90	55.90	6.15	55.90	6.15	4.52

Note: Loadings of <0.40 have been omitted.

IES-R total score was highly correlated with the PCL total score (0.84, $p < 0.001$). Not surprisingly (given the greater variation in scores), correlations were higher for the community sample than for the treatment sample (0.85 versus 0.70). Correlations for the whole sample for each of the subscales were: intrusion = 0.86; avoidance = 0.66; hyperarousal = 0.79 ($p < 0.001$ in each case).

Although neither the IES nor the IES-R were designed to make a categorical diagnosis, it is common for cut-off scores to be quoted in the literature. Thus, optimal cut-off scores for the IES-R were examined by using the recommended cut-off of 50 on the PCL. This analysis was possible only with the community sample since, by definition, the entire treatment sample had a diagnosis of PTSD. Against the PCL, the highest overall diagnostic power (0.88) was achieved with a cut-off of 1.5 (equivalent to a total score of 33) on the IES-R, providing a sensitivity of 0.91, a specificity of 0.82, positive predictive power of 0.90, and negative predictive power of 0.84.

4. Discussion

The current data, using community and treatment-seeking samples of veterans, largely provide support for the psychometric properties of the IES-R. The internal consistency of the total

scale, as well as the three subscales, was high, suggesting that items are tapping into a single construct.

The confirmatory factor analysis using the three subscales of intrusion, avoidance, and hyperarousal as a model, despite high factor loadings, did not provide a good account of the data. The subsequent exploratory analyses suggested that either a single, or a two-factor solution (intrusion/hyperarousal and avoidance), may be more parsimonious. This finding is consistent with the few other published studies on the factor structure of the IES-R. Baguena et al. (2001) reported a two-factor structure they labeled as 'intrusion–hyperactivation and avoidance', while Asukai et al. (2002) report a three-factor solution with the first factor comprising intrusion and hyperarousal items, the second comprising avoidance items, and the third comprising numbing items as well as sleep, irritability, and concentration items.

This difficulty in providing strong psychometric support for the intrusion, avoidance, and hyperarousal subscales is not limited to the IES-R, with several studies proposing four- (Amdur & Liberzon, 2001; King, Leskin, King, & Weathers, 1998) or two-factor models (Taylor, Kuch, Koch, Crockett, & Passey, 1998). One possible explanation for these findings lies in the nature of the samples, particularly with regard to the extent and chronicity of any psychopathology. It seems from the current data, as well as from previous analyses, that there may be less differentiation between the core constructs in trauma survivors with lower symptom levels. This may reflect more general 'traumatic distress' rather than a clearly defined syndrome such as PTSD. This suggestion is supported by the finding in the current data that correlations between intrusion, avoidance, and hyperarousal were higher for the community sample than for the clinical sample. In the final analysis, this is more a question for PTSD as a clinical construct than it is for the psychometric properties of the IES-R.

In terms of construct validity, the IES-R correlated reasonably well with a measure more closely tied to PTSD symptoms, the PCL. At 0.66, the correlations for the avoidance subscale between the IES-R and the PCL were on the low side. This is, perhaps, not surprising since the IES-R questions differ more on this subscale from the PTSD criteria than on intrusion or hyperarousal. Notably, the IES-R places greater emphasis on dissociative avoidance strategies. While caution must be exercised in using the IES-R as a categorical measure, a cut-off of 1.5 (or 33 for a full scale score) provided optimum diagnostic accuracy against the PCL. This is only slightly more conservative than cut-off of 30 (an item mean of 1.4) proposed by Asukai et al. (2002).

In summary, the current data generally support the IES-R as a useful instrument in the assessment of traumatic stress. Although the factor structure is not consistent with the subscales (which represent the PTSD symptom clusters), this finding is common across several other PTSD measures and probably relates more to the validity of the DSM conceptualization of the condition than to the psychometric properties of the measurement instruments. The primary disadvantage of the scale is that the items are not directly tied to the DSM diagnostic criteria, making direct comparisons difficult. On the other hand, as diagnostic criteria tend to change with each new edition of the manuals, there are advantages to a scale that is concerned more with measuring the core constructs of intrusion, avoidance, and arousal that characterize traumatic stress than with adhering to somewhat arbitrary diagnostic criteria.

Acknowledgements

References

Amdur, R. L., & Liberzon, I. (2001). The structure of posttraumatic stress disorder symptoms in combat veterans: a confirmatory factor analysis of the impact of event scale. *Journal of Anxiety Disorders*, *15*, 345–357.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders.* (4th ed.). Washington, DC: American Psychiatric Association.

Asukai, N., Kato, H., Kawamura, N., Kim, Y., Yamamoto, K., Kishimoto, J., Miyake, Y., & Nishizono-Maher, A. (2002). Reliability and validity of the Japanese-language version of the Impact of Event Scale—Revised (IES-R-J): four studies of different traumatic events. *Journal of Nervous and Mental Disease*, *190*, 175–182.

Baguena, M., Villarroya, E., Belena, A., Amelia, D., Roldan, C., & Reig, R. (2001). Propiedades psicometricas de la version espanola de la Escala Revisada de Impacto del Estressor (EIE-R) [Psychometric properties of the Spanish version of the Impact of Event Scale—Revised (IES-R)]. *Analisis y Modificacion de Conducta*, *27*, 581–604.

Blanchard, E. B., Jones Alexander, J., Buckley, T. C., & Forneris, C. A. (1996). Psychometric properties of the PTSD Checklist (PCL). *Behaviour Research and Therapy*, *34*, 669–673.

Cusack, K., & Spates, C. R. (1999). The cognitive dismantling of Eye Movement Desensitization and Reprocessing (EMDR) treatment of posttraumatic stress disorder (PTSD). *Journal of Anxiety Disorders*, *13*, 87–99.

Horowitz, M., Wilner, N., & Alvarez, W. (1979). Impact of Event Scale: a measure of subjective stress. *Psychosomatic Medicine*, *41*, 209–218.

King, D. W., Leskin, G. A., King, L. A., & Weathers, F. W. (1998). Confirmatory factor analysis of the clinician-administered PTSD Scale: evidence for the dimensionality of posttraumatic stress disorder. *Psychological Assessment*, *10*, 90–96.

Pfefferbaum, B., Seale, T. W., McDonald, N. B., Brandt, E. N., Rainwater, S. M., Maynard, B. T., Meierhoefer, B., & Miller, P. D. (2000). Posttraumatic stress two years after the Oklahoma City bombing in youths geographically distant from the explosion. *Psychiatry—Interpersonal and Biological Processes*, *63*, 358–370.

Taylor, S., Kuch, K., Koch, W. J., Crockett, D. J., & Passey, G. (1998). The structure of posttraumatic stress symptoms. *Journal of Abnormal Psychology*, *107*, 154–160.

Weathers, F. W., Litz, B. T., Herman, D. S., Huska, J. A., & Keane, T. M. (1993). *The PTSD Checklist (PCL): reliability, validity, and diagnostic utility.* Paper presented at the 9th annual conference of the ISTSS, San Antonio.

Weiss, D. S., & Marmar, C. R. (1997). The Impact of Event Scale—Revised. In J. P. Wilson, & T. M. Keane (Eds.), *Assessing psychological trauma and PTSD: A handbook for practitioners* (pp. 399–411). New York: Guilford Press.