Comments On Tarrier et al's (1999) Study And The Investigation Of Exposure And Cognitive Therapy.

Grant J. Devilly

Department of Criminology,

University of Melbourne,

Parkville, Victoria 3010.

Australia.

and

Edna B. Foa

Center for the Treatment and Study of Anxiety.

Department of Psychiatry,

University of Pennsylvania,

3535 Market Street, Sixth Floor,

Philadelphia, PA 19104.

USA.

Running Head: Comments On Tarrier et al. (1999).

<u>Abstract</u>

This paper aims to briefly outline some of the concerns relating to the Tarrier et al. (1999) investigation comparing imaginal exposure (IE) and cognitive therapy (CT). Specifically, we offer Tarrier et al. the opportunity to operationally define and clarify the claim that more IE patients "worsened" during treatment. Equally, in light of Tarrier et al.'s low effect sizes in relation to past research we also highlight the need to utilize accountable treatment integrity checks.

In a recent article Tarrier, Pilgrim, Sommerfield, Faragher, Reynolds, Graham and Barrowclough (1999) reported on the relative efficacy of cognitive therapy and an exposure method in treating Post Traumatic Stress Disorder. Such research is of vital importance, the results of which will eventually translate into recommendation of treatment delivery. Therefore, it is important to critically evaluate all treatment outcome research with particular attention to studies in which the obtained results differ significantly from those of other research teams. Hopefully, this process will help clarify the reasons for the disparity.

Tarrier et al. (1999) found no difference between a cognitive intervention (CT) and imaginal exposure (IE) and noted in the discussion that "The only measure to suggest any superiority was that a significantly greater number of patients receiving IE worsened over treatment. However, this effect disappeared by follow-up" (p.17). The term "worsening" was also used in the abstract (p.13). Tarrier et al. arrived at this conclusion by identifying the number of patients who "showed a worsening on the CAPS [Clinician Administered PTSD Scale] total severity score ... by group" (p. 16). Chi Squared tests were then applied on this measure comparing the number of patients who had "worsened", presumably against those that were not identified as "worsening". This lead to the following statistic " χ^2 (1, N = 5.42), p = .02" (sic; p. 16) and the claim of group differences on rates of worsening.

This conclusion has important implications and therefore it merits a close examination with three questions in mind: first, were the appropriate data analyzed; Second, are the data reported in the paper able to support such a conclusion; third, even if the clarified operational definition of worsening withstands scrutiny, was IE given a fair test. Related to this last point we also consider the possible reasons for the inferior outcome of exposure therapy and for the inferred lack of compliance by the IE participants, as mentioned later in the discussion (p.17).

With respect to the first issue no argument was given for focusing solely on the CAPS when the Impact of Events Scale (IES) and the Penn Inventory for PTSD were also available and relevant. Secondly, the term "worsening" was not well defined. In a public discussion with one of us (EBF; Istanbul, 1999), Dr. Tarrier explained that he had used the term "worsening" to denote "lack of significant improvement". These two terms are usually held as very different constructs by researchers and clinicians alike. Worsening of patient status implies clinically significant deterioration. The determination of deterioration usually requires an operational definition that specifies the methods of measurement employed. Previous approaches of quantifying clinical change have been extensively discussed, the issue being of such importance as to merit a special section in a recent issue of this journal (June 1999). Clinical change typically involves the calculation of a reliable change index within a specific domain (Jacobson and Traux, 1991; Hsu, 1989), the delineation of clinically appropriate cutoffs, and the according classification of patients within these two regions of change (Jacobson, Roberts, Berns, McGlichey, 1999). Although other methods of assessment have been posited for different situations (e.g. Nietzel and Trull, 1988, for meta-analytic assessment of clinical significance), it is imperative that the method utilized within a research paper be operationally defined. For example, in the case of the IES, we know that there is a one-week test-retest reliability of 0.87 (Horowitz, Wilner and Alvarez, 1979) for the whole scale and a standard deviation of 13.7 for a normal population. Therefore, the standard error of measurement at any one time point for this scale would be 4.94 in a normal population ($S_E = 13.7\sqrt{1-0.87}$). The standard error of the difference between two administrations of this scale on a normal population would equal 6.99 ($\sqrt{2(S_E)^2}$). Research participants would, therefore, have reliably deteriorated if they had increased their IES score by at least 7 points between any two administrations. We also know that a normal population has a mean score of 13.1 on the IES and, hence, a score of above 41 (2 standard deviations above the mean) could be classified as being above a clinical cut-off. Therefore, a participant whose IES score increased by at

least 7 points and crossed from being under 41 at pre-treatment to scoring over 41 by post-treatment could be classified as having reliably and clinically deteriorated over treatment.

This is just one method of arguing for deterioration. Another method would use the standard deviation of the sample under inspection in deriving the reliable change index. In any event, only after such a method has been made explicit can one convincingly conduct χ^2 testing of the number of people who deteriorated from pre- to post-treatment, as conducted by Tarrier et al. However, one should a priori justify using only one assessment device (e.g. the CAPS) to compute reliable and clinical change when other instruments measuring the same construct (i.e. IES and Penn Inventory) were also utilized in the research. As can be seen, patients deemed to have "worsened" during IE or CT in Tarrier et al. may have been considered "unchanged" if alternate operational definitions of worsening had been employed and if tests of "worsening" had been applied to all PTSD scales. It is, therefore, critical to specify exactly what the operational definition was before examining the implications of the worsening rates. It is hoped that this exchange will provide Tarrier et al. the opportunity to clarify this most important issue.

The third issue we raise regarding the claim of higher rates of deterioration relative to CT is the much inferior outcome of imaginal exposure obtained in the Tarrier et al. study compared to other studies that used exposure therapy with civilian populations (e.g., Foa, Rothbaum, Riggs & Murdoch,1991; Marks, Lovell, Noshirvani, Livanou, & Thrasher, 1998; Foa, Dancu, Hembree, Jaycox, Meadows, & Street, 1999; Devilly and Spence, 1999). Such a disparity could be due to the fact that Tarrier et al. utilized only imaginal exposure rather than the combination of imaginal and in vivo exposure used in the other studies, and inconsistently spaced the delivery of the treatment sessions (points noted by Tarrier et al. in their discussion). Table 1 demonstrates the effect sizes obtained by Tarrier et al. in comparison to the various studies by the current authors and Marks' research group, all of which utilized high procedural integrity checks. As can be seen, Tarrier et al. consistently obtained

smaller effect sizes than all past research studies. It is unclear whether this inferior outcome is due to the omission of exposure <u>in vivo</u>, or to a lack of adherence to past protocols of imaginal exposure. It is our hypothesis that, with particular regard to avoidance symptoms, <u>in vivo</u> exposure adds to the efficacy of exposure interventions (e.g. Richards, Lovell and Marks, 1994) yet, in lieu of thorough documentation of the extent of this additive effect, one must ensure treatment integrity of the imaginal exposure.

An appropriate investigation of treatment integrity would have clarified if and how Tarrier et al.'s delivery of imaginal exposure differed from that of other researchers. For example, while Tarrier et al. note that the therapists kept the participants in the present tense, was this integrated into the session effectively? Did the therapist note 'hot spots' where appropriate and habituate the participant to these? How well were these tasks achieved, and so on? Unfortunately, the method of assessing "treatment fidelity" that was employed by Tarrier et al. provides no information about adherence to IE procedures as practiced by other researchers. Interestingly, just as the outcome of IE in the Tarrier et al. study is inferior to those found in other studies, so is the outcome of CT (e.g., Resick and Schnicke, 1992), raising again questions about deviations from the common use of CT.

Our questions regarding the interpretation of the results are not meant to detract from the importance of studying the efficacy of cognitive therapy for PTSD and comparing it with that of exposure therapy. Indeed, there is a dearth of research on the efficacy of CT for PTSD, and we need to conduct more studies using this therapy. However, while we have other concerns relating to the Tarrier et al. paper we feel that the above mentioned issues are of paramount importance and require clarification before the results can be interpreted without prejudice.

References

Devilly, G.J. & Spence, S.H. (1999). The relative efficacy and treatment distress of EMDR and a cognitive behavior Trauma Treatment Protocol in the amelioration of posttraumatic stress disorder.

Journal of Anxiety Disorders, 13 (1-2), 131-157.

Foa, E.B., Dancu, C.V., Hembree, E., Jaycox, L.H., Meadows, E.A., & Street, G.P. (1999). The Efficacy of Exposure Therapy, Stress Inoculation Training and Their Combination in Ameliorating PTSD for Female Victims of Assault. Journal of Consulting and Clinical Psychology, 67, 194-200.

Foa, E.B., & Kozac, M.J. (1986). Emotional processing of fear: Exposure to corrective information. Psychological Bulletin, 99, 20-35.

Foa, E.B., Rothbaum, B.O., Riggs, D.S., & Murdoch, T.B. (1991). Treatment of posttraumatic stress disorder in rape victims: A comparison between cognitive-behavioral procedures and counseling. Journal of Consulting and Clinical Psychology, 59, 715-723.

Horowitz, M., Wilner, M., and Alvarez, W. (1979). Impact of Event Scale: A measure of subjective stress. Psychosomatic Medicine, 41, 209-218.

Hsu, L.M (1989). Reliable changes in psychotherapy: Taking into account regression toward the mean. Behavioral Assessment, 11, 459-467.

Jacobson, N.S., Roberts, L.J., Berns, S.B., & McGlichey, J.B. (1999). Methods of defining and determining the clinical significance of treatment effects: Description, application, and alternatives.

Journal of Consulting and Clinical Psychology, 67, 300-307.

Jaycox, L.H., Foa, E.B., & Morral, A. (1998). The influence of emotional engagement and habituation on exposure therapy for PTSD. <u>Journal of Consulting & Clinical Psychology</u>, <u>66</u>, 185-192.

Marks, I., Lovell, K., Noshirvani, H., Livanou, M., & Thrasher, S. (1998). Treatment of Posttraumatic Stress Disorder by Exposure and/or Cognitive Restructuring. <u>Archives of General Psychiatry</u>, 55, 317-325.

Nietzel, M..T., & Trull, T.J., (1988). Meta-analytic approaches to social comparisons: A method for measuring clinical significance. Behavioral Assessment, 10, 159-169.

Resick, P.A., & Schnicke, M.K. (1992). Cognitive processing therapy for sexual assault victims. Journal of Consulting and Clinical Psychology, 60, 748-756.

Richards, D.A., Lovell, K., Marks, I.M. (1994). Post-traumatic stress disorder: Evaluation of a behavioral treatment program. Journal of Traumatic Stress, 7, 669-680.

Tarrier, N., Pilgrim, H., Sommerfield, C., Faragher, B., Reynolds, M., Graham, E., & Barrowclough, C. (1999). A randomized trial of cognitive therapy and imaginal exposure in the treatment of chronic posttraumatic stress disorder. <u>Journal of Consulting and Clinical Psychology</u>, 67, 13-18.

Waller, S. and Spates, C.R. (2000). Interventions for PTSD: A meta-analysis and quantitative review. <u>Third World Conference For The International Society For Traumatic Stress Studies</u>, March, Melbourne.

Table 1. Comparison of Effect Sizes in Past Research into PTSD Treated with Exposure Techniques.

Study & Follow-up Time Used In			ES	
Analyses	Condition	Measures	(Pre- to Post-tx)	Pre-tx to Follow-up
Foa et al., 1991. (3 Months)	PE (IE + <u>In Vivo</u>)	PSS-I (Total)	1.29	2.27
Foa et al., 1999. (6 Months)	PE (IE + <u>In Vivo</u>)	PSS-I (Total)	2.06	2.12
Devilly and Spence, 1999.	TTP (PE+SIT+COG)	PSS-SR (Total)	1.85	1.83
(3 Months)		IES (Total)	1.82	1.81
Marks et al., 1998 (6 Months)	PE (IE + <u>In Vivo</u>)	CAPS (Severity)	1.10	3.82
		IES (Total)	1.50	2.70
Tarrier et al., 1999. (6 Months)	IE	CAPS (Severity)	0.93	0.89
		IES (Intrusion)	0.82	0.92
		IES (Avoidance)	0.85	0.50

Note. Effect size (Cohen's *d* of the treated) was estimated by subtracting the post-treatment or follow-up mean from the pre-treatment score and dividing the result by the mean of the standard deviations from the two time points being compared. Such an approach has been found to be more conservative than only utilizing the pre-treatment standard deviation (Waller and Spates, 2000). PE = Prolonged Exposure; IE = Imaginal Exposure; PSS-I = PTSD Symptom Scale – Interview; TTP = Trauma Treatment Protocol; SIT = Stress Inoculation Training; COG = Cognitive Intervention; PSS-SR = PTSD Symptom Scale (Self-Response); IES = Impact of Events Scale; CAPS = Clinician Administered PTSD Scale.