

# JOINING FORCES

VOLUME 5 ISSUE 3

RESEARCH NEWS YOU CAN USE

SPRING 2001

This issue of *Joining Forces* features the research recommendations from the initial report of the Defense Task Force on Domestic Violence (DTFDV). We hope that our summary of their research priorities will inform and help to inspire workers in the Army FAP to plan research.

Also included are abstracts of two recent research articles, one of a study of a Navy recruit population, and one of fatherhood as portrayed in the comic strips. Our regular statistics article returns with a discussion of effect size.

We always look forward to and appreciate your comments. Our e-mail addresses are on the next page.

## IN THIS ISSUE

### THE RESEARCH RECOMMENDATIONS OF THE INITIAL REPORT OF THE DEFENSE TASK FORCE ON DOMESTIC VIOLENCE

The National Defense Authorization Act of 2000 that established the task force (TF) required it to formulate a long-term strategic plan to assist the Department of Defense (DoD) in addressing domestic violence in the military. The initial report (sent to the Secretary on February 29, 2001) included a review of pending, completed, and

recommended DoD research relating to domestic violence on spouse abuse only (see [www.dtic.mil/domesticviolence/](http://www.dtic.mil/domesticviolence/)).

Their approach to problems of domestic violence included a review of scientific literature, in particular, a review conducted by the US Air Force of over 700 empirical studies of domestic violence. A total of 28 articles published since 1985 addressed domestic violence in the military. Most of these articles were descriptions of rates or trends or reviews of literature. Only three were studies of interventions and none addressed prevention; one addressed program evaluation and two were studies of methodology.

The military services had 14 studies pending at the time of the report. Several of these studies are descriptive, but others are much more ambitious, addressing such topics as a comparison of the prevalence of intimate partner abuse between active duty and civilian women, perceptions of the mandatory reporting of spouse abuse, and its medical and economic costs. There is one pending study on interventions and two pending evaluations of prevention programs. (Listings of completed studies are in Appendix II and III of the TF report).

Based on their background literature review and analysis, the TF recommended that the DoD: (1) partner with the Department of Justice and the Centers for Disease Control and Prevention to articulate the research agenda and

organize scientific community-wide requests for applications and peer review of proposals; (2) facilitate and encourage publication in peer reviewed journals; and (3) not be precluded from funding research into the causes, consequences, and interventions of domestic violence in the military through other service funding mechanisms such as Tri-Service Nursing or the Uniformed Services University of the Health Sciences (USUHS). Given this background, the TF posted seven research priorities.

#### **Priority 1: Differentiation of different types of abusers and abusive situations.**

In an attempt to move the FAP past the one-size-fits-all approach to intervention by developing new strategies for both victims and offenders based on the characteristics of each, they suggested:

- The development of a screening instrument to differentiate serious and chronic abusers from those who are unlikely to be a continuing threat.
- Longitudinal studies of the natural course of abuse to determine which abusers escalate their violence versus those who do not.
- Studies to determine whether typologies (such as those reported by Holtzworth-Munroe A. & Stuart, G. 1994. Typologies of male batterers: three subtypes and the differences among them. *Psychological Bulletin*, 116:476-497) are a good fit with military populations.



**Priority 2: Which interventions work and for whom?**

This priority addresses the need for interventions for both offenders and victims. For offenders, these recommendations aim to determine which specific interventions are most effective for which offender. These include:

- The duration of treatment.
- The mechanism by which an intervention works.
- The likelihood of recidivism by type of abuser.
- The length of the intervention’s effectiveness.

- How effective are intervention programs for victims?

**Priority 3: How well does the military approach to domestic violence work and where should it be modified?**

- What is the impact of mandatory reporting on the disclosure of abuse and the outcome of the process for victims and offenders?
- What information do commanders need at the various decision points? (What do they need to know? Do they know what actions to take?)
- Does the current definition of spouse abuse as a discrete, incident-based event facilitate or interfere with good case decision-making?
- Would a more dynamic definition lead to more reliable and accurate decisions by the case review committees?

- Related to priority 1, it asks for longitudinal surveys of people who were victims of violence or witnesses to parental violence to determine if such a history is related to involvement in adult domestic violence.

- If such a relationship does exist, this priority asks for the development of an intervention system for identifying and treating children of military families with currently identified domestic violence.

**Priority 6: Evaluate the knowledge and consistency of the actions of key FAP personnel (law enforcement, medical, chaplains, and FAP).**

This priority is aimed at increasing the effectiveness of all personnel functioning in the FAP.

- How do law enforcement personnel identify and label domestic violence incidents? Is the training on target, directed, effective, and consistent across installations and services?
- How knowledgeable are medical personnel and chaplains about domestic violence? What actions do they take, and how effective are these actions?
- How knowledgeable are FAP personnel about domestic violence and how do their knowledge and beliefs impact their effectiveness? What factors are most important in case-related decisions by FAP staff?

**Priority 7: Three topics were suggested in this category:**

- How effective is counseling for “low level” cases of domestic violence?
- What are the differences and similarities between the genders in the use of violence in intimate relationships?

**JOINING FORCES**

Editor-in-chief

James E. McCarroll, Ph.D.

Email: [jmccarroll@usuhs.mil](mailto:jmccarroll@usuhs.mil)

Editor

John H. Newby, DSW

Email: [jnewby@usuhs.mil](mailto:jnewby@usuhs.mil)

Joining Forces is a publication of the Community and Family Support Center and the Family Violence and Trauma Project at the Department of Psychiatry, Uniformed Services University of the Health Sciences, Bethesda, Maryland 20814-4799. Phone: (301) 295 - 2470



Specifically for victims:

- What are the best strategies to assist victims in dangerous situations?
- What interventions do victims prefer?

**Priority 4: What is the best estimate of the actual (population) prevalence of domestic violence in the military compared to the prevalence reported by counting cases?**

- How do the actual (e.g., population or “true” prevalence) and the prevalence of identified cases (the cases referred to FAP, reviewed by a CRC, and entered into a service central registry) vary across installations and services?
- Is any variation in prevalence related to the command climate or system policies or procedures?

**Priority 5: Which approaches to domestic violence prevention work and for whom?**

This set of priorities includes understanding the relationships between early (childhood) abuse and later involvement in adult domestic violence.



- What is the impact of the lack of confidentiality on disclosure and victim safety?

We hope that this summary of the recommendations of the DTFDV will help Army FAP personnel to undertake some research suggested, particularly as the Army FAP attempts to establish centers of excellence.

**FATHERHOOD  
(in the comic strips)**

In our never-ending search to find research articles of interest to the Army FAP, we present an abstract of a recent article (LaRossa R, Jaret C, Gadgil M, & Wynn, GR. 2000. The changing culture of father-hood in comic-strip families: A six-decade analysis. Journal of Marriage and the Family, 62:375-387) that ties in with some of our recent features of research on fatherhood.

The article presents some interesting data on how both fatherhood and motherhood are presented in the media, thus reflecting some aspects of our culture. It is also a good example of how to conduct one type of cultural assessment through literature review and analysis.

The authors analyzed 490 comic strips in the Atlanta Constitution published on Mother's Day and Father's Day from 1940-1999. They coded depictions of both mothers and fathers for incompetence, mocking, and nurturing and supportive parental behaviors. In their overall summary, they report-ed that while incompetence was rarely

portrayed, fathers were shown as more incompetent (10.9%) than mothers (5.7%). Fathers were mocked in 18%; mothers in 6% of the comic strips. Nurturing and supportive behavior toward children was split into two types. The first type (expressing affection, caring for, verbally encouraging, or comforting a child) showed 31.6% of the mothers displaying such behaviors compared to 23% of the fathers. The second type included praising, listening to, or teaching a child. In this category, mothers outscored fathers 42.9% to 36.4%.

Half-decade analyses of each of the categories showed a high degree of fluctuation with the differences between fathers and mothers oscillating from one decade to the next. Incompetent fathers were shown more in the late 1940s, early 1950s, and late 1960s. Mocked fathers were more common in the 1960s and early 1980s. Nurturing and supportive fathers were most evident in the late 1940s, early 1950s, and the 1990s.

Interwoven in the discussions of this article are references to societal changes in motherhood, fatherhood, and the nature of the family. Among these are the depictions of mothers in the 1940s as overprotective, the use of satire in the depiction of mocking, the fragility of the family in the 1990s, and the problematic nature of depicting the absentee father along with changes in the "new" father.

This is a complex study (there are no jokes), but interesting for both its content and methodology. It puts a slightly different light on mother and fatherhood than that usually seen in the scientific literature. Obviously, the extent that comics may illustrate current

culture-based behaviors, fears, or wishes leaves much to be understood about parental roles and how their children and others perceive them.



**CHILDHOOD ABUSE AND  
PREMILITARY SEXUAL  
ASSAULT IN MALE NAVY  
RECRUITS**

A recently published study presents new findings from the longitudinal study of Navy recruits (Merrill LL, Thomson CJ, Gold SR, & Milner JS. 2001. Childhood Abuse and Premilitary Sexual Assault in Male Navy Recruits. Journal of Consulting and Clinical Psychology, 69:252-261).

Please note that all the data reported here are on the PRE-MILITARY behavior of these recruits, not during active duty.

In addition to providing important new information about abuse history and its consequences in military personnel, this is a very well-written article that, by clear interpretation, walks the reader through complex statistical analyses.

The purpose of the research was to determine the effect of childhood physical abuse (CPA) and childhood sexual abuse (CSA) on the likelihood of rape. The subjects were samples of male Navy recruits selected over three years, 1994, 1996, and 1997. A total of 36-39% of the sample



reported CPA and 10-12% CSA. Ten to 12% reported that they had committed rape and 2-4% had attempted (but not committed) it.

They first examined whether CSA and CPA co-occur. CSA predicted CPA after controlling for demographic variables. CPA was a significant predictor of rape in all three samples. CSA was significant in two of the three, and it approached significance in the third sample. Their results also suggested that CPA and CSA were independent, additive predictors of rape. As they explained, this means that the likelihood of rape for victims of both types of abuse is roughly predictable as the sum of the increased likelihood of each. That is, the increased likelihood of rape for men who had experienced both CPA and CSA was 4-6 times more likely than the likelihood of rape by non-abused men.

In another set of analyses, they examined whether two potential variables, alcohol problems (AP) and number of sexual partners (NSP), were mediators of the relationships found between CPA and CSA and rape. (Ed. note: A mediator is a variable that acts as an intermediary, or intervenes, in a statistical relationship.) They pointed out that three steps are necessary to make a case for such mediation: (1) childhood abuse must predict AP and NSP, (2) NSP and AP must predict rape, and (3) the relationship between childhood abuse and rape must be eliminated or reduced when the effects of AP and NSP are controlled. (Ed. note: Mediation is not the same as confounding.)

CSA was no longer a predictor of rape after the mediators had been controlled. This indicated that AP and NSP were largely responsible for the effects of CSA

on rape. In contrast, CPA remained a predictor of rape after controlling for the mediators, but its effect was reduced. Thus, the relationship between CPA and rape was only partially mediated by AP and NSP.

Continuing their analyses to obtain as much information as possible, they conducted a path analysis to estimate the effect size of CSA and CPA on rape. (Ed. note: See accompanying statistics article on effect size). Their model included CSA and CPA as causal factors, AP and NSP as mediators, and rape as the dependent variable. Tests of the model indicated a good fit and together the predictors accounted for 10-14% of the variance in rape (the estimated effect size).

A final set of analyses tested whether alcohol use and high levels of sexual activity predicted sexual assault for men with different abuse histories. In other words, they tested for the presence of an interaction between abuse history and AP and NSP on rape. This was done because of the suggestion that men's sexual assaults may be due to situational factors such as alcohol intoxication. No such interaction was found. The negative results indicated little evidence that abuse history moderates the relationship of AP and NSP to rape.

So, what does it all mean? Men who experienced either CSA or CPA were more likely than non-abused men to commit rape. The authors highlighted two important contributions of their study. First, the effects of CSA and CPA on rape are independent and additive, not the result of an interaction. Second, the effects of CPA and, especially, CSA are mediated through AP and NSP.

The authors provided a thorough discussion of the limitations of the study and what they may mean. Among these were that replication of these results is necessary in other populations before they can be generalized. Because the data were based on self-reports, there was some likelihood of shared variance among the measures. There are memory constraints to self-reports, but the high level of negative behaviors makes an overestimate unlikely. This was a cross-sectional study and cannot determine causal relations. More study of the role of alcohol and number of sexual partners on family violence is required.

Among the strengths of the study are that the data were obtained from multiple samples, were consistent across samples, and were consistent with previous findings.

They suggested that future research should include additional mediators to help understand the processes by which childhood abuse may promote a cycle of violence.

### **STATISTICS** **Effect Size Measures**

In Volume 3, Issue 4, we discussed the interpretation of statistical significance. We made two points in this piece. First, that statistical significance is based on two factors in the data: sample size (the larger the sample the greater likelihood of significance) and the variability (the less variation the more the likelihood of significance). Second, a finding may be statistically significant, but not meaningful. That is, the actual contribution of the finding may be small. This is particularly true in



clinical work where it is important to have confidence that the procedure will make a difference in a person's life, and not just a statistical difference.

How does one tell if a research result (e.g., a new form of therapy) is powerful enough to produce clinically relevant change? We explore this issue through a review of effect size statistics. Although not new, such statistics are now often seen in the scientific literature, particularly in conjunction with evaluations of the effectiveness of therapies.

Evaluations of therapeutic effectiveness are important to the Army FAP as it attempts to develop outcome measures for the most effective treatment for family violence victims and offenders.

The thrust of recent research in psychotherapy has been toward identifying empirically supported therapies; that is, finding treatments that have been shown to work with certain patients. Three terms are frequently used in this type of research (Chambless & Hollon, 1998). These are efficacy (does the treatment work in a controlled setting?), effectiveness (does the treatment work in actual clinical practice?), and efficiency (is the treatment cost-effective with regard to other interventions?).

One method of attempting to provide answers to each of these three measures is through the use of a measure of effect size. Many professional journals have recently provided editorials about how to interpret effect size as well as other statistical topics such as tests of statistical significance, p-values, and confidence intervals. For example McClure (1999) recommended looking at effect

sizes in addition to traditional tests of significance.

In the most general terms, an effect is a difference between two populations that differ on some characteristic. For example, using standardized measures, you may find a difference in depression (the phenomenon of interest) in two groups of people, those with a history of abuse and those with no such history. This effect may be absolute (the difference between the means of the two groups on the standardized measure) or it may be relative (the mean of the exposed group divided by the mean of the unexposed group). One measure of an effect size is the magnitude of this difference. How large is the difference between the means when measured in standard deviation units? That is an example of one basic type of effect size measure. Another is that derived from a meta-analysis. A meta-analysis is a type of statistical procedure computed to obtain a measure of the results of a number of studies to determine if a phenomenon is present based on aggregated data. (Maxwell Smart might have said: "If you don't believe one study, would you believe five?") This type of measure is often presented at meetings to summarize the results of several studies where one type of therapy is compared with another or there are multiple comparisons on the same type of therapy.

Cohen (1988) gives guidelines for effect sizes based on the d-statistic (difference between means measured in standard deviation units) small, 0.20; medium, 0.50; large 0.80 (p. 40). (Cohen describes a small effect size as one that is similar to the difference in mean height between 15 and 16

year old girls.) A moderate effect size should be visible to the naked eye. In a large effect size, there should be little overlap between the two distributions. For example, one would expect a large effect size on depression between people who are medicated and those who have not been. These guidelines are often repeated in many publications; however, (1) they are arbitrary, and (2) the clinical relevance of a treatment effect cannot be deduced from an effect size (Scholten, de Beurs, & Bouter, 1999).

For example, investigators may use the same measure of depression in different samples of people. They obtain the mean differences in depression scores and divide by a pooled standard deviation obtained from the individual studies. (See Senra, 1995, for a formula for a pooled standard deviation.)

In a recent study comparing cognitive versus behavior therapy in the treatment of obsessive-compulsive disorder, effect sizes from multiple therapy groups were presented (McLean et al., 2001). An effect size of 1.62 was found for exposure therapy and 0.98 for cognitive-behavior therapy, both compared to controls who were put on a waiting list.

Cohen (1988) noted that a measure of effect size (1) does not imply causality, and (2) it is a measure of the degree to which the phenomenon under study is present in the population that was investigated. In other words, if there is no effect size, there is no effect.

In spite of its frequent use, effect size is not without its problems and its critics. For example, Greenlander (1998, p. 672) advises avoiding it because expressing



effects in standard deviation units can yield spurious results. Identical results can be made to appear different or even can reverse the order of the strength of results. He advocated expressing effects in meaningful units that are uniform across studies, not in standard deviation units. The reason for this rather strong warning is because the variability across studies is almost never uniform, especially in the behavioral sciences. Thus such standard deviation units may not be helpful for comparing effect sizes.

Scholten, De Beurs and Bouter (1999) demonstrated, using fictitious data, how an attempt to transform effect sizes of decreased blood pressure due to a new drug produced incorrect and confusing results. They noted that the translation of effect sizes into clinically meaningful units is hazardous and that the assessment of a treatment effect using only effect sizes is challenging.

How do you use an effect size statistic in analyzing the results of a study? First, it is difficult (unless you work consistently in this area of statistics) to make an intuitive interpretation of what an effect size means. We advise looking at the magnitude of changes in the basic measures, such as the differences between the means of the experimental and control groups and at the magnitude of variability on the actual measures of interest. If these figures make sense, you can probably tell if the actual result of the study is of importance to you. By examining the means and standard deviation, you can also derive your own estimate of an approximate effect size.

Second, look at the way that effect size has been computed and make sure it has been calculated

correctly and you understand what it means. As we have pointed out, there is more than one type of effect size, it may mean something different than what you think, and it may not be calculated correctly.

Third, make sure you distinguish effect sizes from other statistics such as p-values, odds ratios, and confidence limits. Not all of these statistics have the same interpretation.

### References

- Chambliss DL & Hollon SD. (1998). Defining empirically supported therapies. Journal of Consulting and Clinical Psychology, 66:7-18.
- Cohen J. (1988). Statistical power analysis in the behavioral Sciences. Hillsdale, NJ: Erlbaum.
- Greenlander S. (1998). Meta-analysis. In Rothman KJ & Greenland S. Modern Epidemiology (Second Edition). Philadelphia, PA: Lippincott-Raven.
- McClure PW. (1999). Determining the significance of significance: p-values, effect size, and clinical judgment. Journal of Hand Therapy, January-March, 1999: 41.
- McLean PD, Whittal ML, Söchting I, Koch WJ, Paterson R, Thordson DS, Taylor S, & Anderson KW. (2001). Cognitive versus behavior therapy in the group treatment of obsessive-compulsive disorder. Journal of Consulting and Clinical Psychology, 69:205-214.
- Scholten RJPM, de Beurs, & Bouter LM. (1999.) From effect size into number need to treat. Lancet, 354:598.

Senra (1995.) Measures of treatment outcome of depression: an effect size comparison. Psychological Reports, 76:187-192.

This newsletter was prepared for the U.S. Army Community and Family Support Center, Family Advocacy Program, under an Interservice Support Agreement between the Department of the Army, and the Department of Defense, Uniformed Services University of the Health Sciences, Department of Psychiatry.

This newsletter was prepared for the U.S. Army Community and Family Support Center, Family Advocacy Program, under an Interservice Support Agreement between the Department of the Army, and the Department of Defense, Uniformed Services University of the Health Sciences, Department of Psychiatry.