Do Fairly-Decided Maltreatment Determinations Significantly Reduce Recidivism? A Quasi-Experimental Evaluation of a System-Level Intervention Implementation

Amy M. Smith Slep¹, Richard E. Heyman¹, Danielle M. Mitnick¹, Michael F. Lorber¹, Sara R. Nichols¹ and Daniel F. Perkins²

Abstract
Two studies examined the impact of the implementation of the Field-tested Assessment, Intervention-planning, and Response (FAIR) system, a system-level intervention for determining whether allegations of family maltreatment meet threshold for abuse or neglect, on alleged recidivism. Data were collected at the 10 U.S. Army installations with the largest family maltreatment caseloads. Participants were family members who had an allegation of family maltreatment (i.e., child maltreatment or partner abuse) during one of the two study periods. Data were collected when Family Advocacy Program staff used the then-in-place system (Case Review Committee) and later the FAIR

¹New York University, USA
²The Pennsylvania State University, University Park, USA

Corresponding Author:
Amy M. Smith Slep, Family Translational Research Group, New York University, 137 East 25th Street, 6th Floor, New York, NY 10010, USA.
Email: amy.slep@nyu.edu
system. In Study 1, cases were followed for 6 months following the initial maltreatment allegation to measure the occurrence of subsequent allegations of any type. Additionally, at five installations, alleged victims of partner abuse were recruited into a study (Study 2) in which they anonymously reported on intimate partner violence via telephone. In Study 1, the advantage for the FAIR condition was concentrated in cases with unsubstantiated initial determinations; the mean relative risk reduction for recidivism was 0.48. In Study 2, FAIR extended median time to recidivism by approximately 170%. These results replicate and extend earlier findings that employing the FAIR system can result in decreased family maltreatment re-offense.

**Keywords**
child maltreatment, partner abuse, child welfare, fairness, decision-making, implementation, FAIR system

Family maltreatment, comprising child and partner physical, emotional, and sexual abuse, as well as child neglect, is a public health burden, costing billions and causing acute and long-term physical, mental, and behavioral impacts (Foran et al., 2013). Systems responsible for addressing maltreatment have had only modest impacts with one exception: The FAIR (Field-tested Assessment, Intervention-planning, and Response) system—a multi-component approach (i.e., a system-level intervention) involving assessment, training, criteria, and computerized decision making—has evidence of reliability (Heyman & Slep, 2006) and validity (Heyman & Slep, 2019). As will be summarized below, FAIR was found to reduce re-offense in a quasi-experimental study. Questions remained, however. First, does this reduction in re-offense generalize to other settings? Second, was this decrease in maltreatment under the FAIR system due to (a) reduced subsequent maltreatment or (b) reduced reporting of maltreatment? The current research addresses these questions with two studies: a record review and a prospective longitudinal study of anonymous victim reports.

FAIR was developed through a program of research with the United States Air Force Family Advocacy Program (FAP) to develop criteria and procedures for reliably and validly making partner abuse and child maltreatment (comprising abuse [physical, psychological, and sexual] and neglect) determinations. FAP is the military equivalent of civilian Child Protective Services but has responsibility for both child maltreatment and partner abuse prevention, assessment, and intervention. FAIR was developed and tested in a series of field trials and a dissemination trial (Heyman & Slep, 2006, 2009; Slep &
Heyman, 2006) where the criteria and procedures were refined. These studies established that the system worked well in real-world conditions (e.g., agreement between local decisions and master reviewers >90%), even when disseminated broadly with non-volunteer sites. In addition, FAIR resulted in substantially lower rates of re-offense (Snarr et al., 2011). On the basis of these findings, FAIR’s criteria were adapted for the International Classification of Diseases, 11th Edition (ICD-11; World Health Organization, 2021) and the Diagnostic and Statistical Manual of Mental Disorders–5th Edition (DSM-5; American Psychiatric Association, 2013). The FAIR system has also been adopted by Alaska’s child welfare system (Mitnick et al., 2022).

FAIR’s results regarding reliability and consistency in decision making is unique in healthcare and child welfare. First, most behavioral disorder determinations have poor inter-rater agreement (Rettew et al., 2009). In contrast, FAIR resulted in 90% agreement with master reviewers (Heyman & Slep, 2006, 2009). Second, determinations under FAIR are perceived as fair to alleged offenders and victims by stakeholders (Heyman et al., 2010). Third, FAIR shows evidence of negligible systematic bias (Heyman et al., 2016). Finally, compared with the military’s previous system, FAIR cut 1-year re-offense rates for those who met criteria in half (Snarr et al., 2011), suggesting a tertiary preventive effect of a fairer and more consistent system.

The preventative effect of the FAIR system is perhaps the most important evidence supporting its further dissemination. This effect was found in a quasi-experimental study of the Air Force comparing cases nested within sites during the final window of the use of their prior system and comparing that to cases nested within sites during a comparable window after switching to FAIR (Snarr et al., 2011). Although not directly testable, we interpret the preventative impact as a function of all stakeholders perceiving the FAIR system to result in more fair determinations. We believe that this, along with the clarity of the definitional criteria in FAIR, made it easier for commanders and others to present the behavioral standards clearly to service members and families and for them to send more consistent messages about the need to not cross the line that definitions draw. This multiparty clarity of communication is a key element of the coordinated community response model employed by the military. However, it cannot be ruled out that another factor, such as decreased reporting of maltreatment to authorities, contributed to the reduction in re-offense because only official allegations were examined in that study (Snarr et al., 2011).

Study 1 replicated and extended the evaluation of recidivism in the U.S. Air Force by Snarr et al. (2011). The study was conducted in the U.S. Army, where the definitional criteria used by FAIR had been implemented; however, FAIR’s recommended training, structured assessments, and re-organized committee determination had not been adopted. In this study, we
compared the Army’s existing system (Case Review Committee [CRC]) with the fully implemented FAIR system at 10 large Army installations. We hypothesized that cases processed through the fully implemented FAIR system, compared with those processed through the CRC system, would have lower rates of alleged re-offense. Following Snarr et al. (2011), we further explored the moderation of these effects by victim age and substantiation status of the initial allegation. In Study 1, recidivism was solely based on official allegations. To address this concern in Study 2, we supplemented initial official partner abuse records with anonymous victim reports of subsequent intimate partner physical, psychological, and sexual abuse.

**Study 1**

Permission to collect data was obtained from the Army’s Human Research Protection Office, and New York University Institutional Review Board approved the study protocol.

**Method**

**Participants.** There were 14,611 FAP incident records during the study period from 10 Army installations with the largest FAP caseloads. FAP cases include parent-child and partner physical, emotional, and sexual abuse, as well as child neglect. All Army sites directed by Army leadership to participate in the research project (and full FAIR implementation) did so. Records randomly selected for coding by research staff (8,231 records nested in 4,673 families) were used for the present analysis; all records that fell in the designated time frame were randomly ordered using a random number generator, and records were coded in that order as long as the study period permitted.

**Design.** Utilizing a quasi-experimental design, recidivism was compared in the CRC (5,143 incidents between November 2016 and August 2017) versus FAIR (3,088 incidents between October 2017 and July 2018) implementation phase conditions. In the CRC phase, maltreatment decisions were evaluated via Case Review Committee, the system in place in the Army at the outset of the study. In the CRC, assessing social workers presented their own cases, commanders did not vote, and the committee members outside of the FAP were not trained in the definitions. Discussion included incident-related information and treatment recommendations. Additional details can be found in Heyman et al. (2022). In the FAIR phase, maltreatment decisions were evaluated via the updated committee structure and procedures and in meetings that separated substantiation decisions from clinical treatment...
discussions. The committee was chaired by an installation leader, a supervising social worker presented cases after review with the assessing social worker (who was not present), all members were trained in the definitions and procedures and passed a quiz certifying mastery, and each involved Soldier’s commander was trained and voted along with the rest of the committee. At all Army installations, CRC implementation immediately preceded FAIR implementation. Because we needed 6 months to pass from the meeting date to allow for recidivism to occur, and due to study timeframe cutoff, the CRC phase allowed for more records to be reviewed. Because records were reviewed in random order in each phase, this does not introduce bias.

**Measures.** All data were drawn from official FAP records. To protect anonymity, a very limited set of data lacking characteristics (e.g., race/ethnicity) that might, in combination, inadvertently identify individuals was provided to the research team.

**Alleged initial maltreatment and recidivism.** The presence/absence of parent-child and partner physical, emotional, and sexual abuse, as well as child neglect, was judged from official allegations received by FAP personnel. Upon receiving an initial allegation, FAP intake workers categorize it into one of the above categories using domestic abuse definitions as specified in the Decision Tree Algorithm (Department of Defense, 2016). To operationalize recidivism, maltreatment was coded in the 6-months following initial reports for whether new allegations of any type were recorded (yes/no; 1/0). The 6-month follow-up period was chosen as the longest period possible within the practical constraints of the study length.

**Initial substantiation determination.** Each initial allegation in the study period was categorized as meeting (1) versus not meeting (−1) criteria for maltreatment.

**Initial maltreatment severity.** Each initial allegation was categorized on a 4-point ordinal scale: 0 = did not meet criteria, 1 = mild, 2 = moderate, and 3 = severe, based on the Incident Severity Scale (Erlanger et al., 2022).

**Victim age group.** Victims were categorized as child (−1) or adult partner (1).

**Military status.** The alleged perpetrators’ and victims’ military statuses were coded as service member (1) or civilian (−1).

**Gender.** The alleged perpetrators’ and victims’ genders were coded as male (1) or female (−1).
Analytic Method

Regression models were conducted using Mplus version 8 (Muthén & Muthén, 2017). The nesting of participants within families was handled via robust pseudo maximum likelihood estimation method (“type = complex” in the Mplus analysis specifications), which utilizes a sandwich estimator to adjust parameters’ SEs for the non-independence of clustered observations (Asparouhov & Muthén, 2005).

We first evaluated differences in the CRC and FAIR phases to screen for potential confounds on available incident characteristics: offender and victim gender and military status (service member vs. civilian), allegation substantiation, maltreatment severity, maltreatment victim age group (child vs. adult partner), and maltreatment type (emotional abuse, physical abuse, sexual abuse, and neglect). In each case, nine effects coded (i.e., 1 vs. −1) dummy variables for installation (reference installation is the 10th) were included as covariates. These analyses are summarized briefly in the “Results” section and more fully presented in the Supplemental Material to this article.

To test the main effect for phase, we evaluated a logistic regression model in which alleged recidivism (1 = present; 0 = absent) was regressed on effects coded variables for phase (−1 = CRC; 1 = FAIR), as well as covariates representing installation and the two variables identified in the confound screening (maltreatment type and victim age group).

To test the Phase × Victim Age Group × Allegation Substantiation interaction, we evaluated a logistic regression model that included all lower order main effects, as well as two-way and three-way interaction terms among phase, victim age group, and allegation substantiation. The covariate set described for the main effect model was also included in this model. The interaction was probed graphically and simple slopes (SS) were calculated per Preacher et al. (2006).

Due to missing data, Ns varied slightly from analysis to analysis, ranging from 8,192 to 8,213 incidents nested in 4,657 to 4,666 families.

Results

Descriptive statistics are presented in Table 1 for the incident characteristics and in Table 2 for the recidivism rates by initial substantiation decision per condition. The sample-wide recidivism rate was 10.9%.

Of the variables evaluated as potential baseline confounds, there were significant differences by phase for maltreatment victim age group and emotional abuse (Table 3; Supplemental Tables S3 and S4). Compared to the
CRC phase, there were proportionately more adult partner than child victims and an overrepresentation of emotional abuse incidents in the FAIR phase. Accordingly, variables representing installation (see “Analytic method”) were supplemented with effects coded variables for age category (1 = adult;
−1 = child) and maltreatment type (emotional, physical, and sexual abuse; reference category is neglect; 1 = present, −1 = absent) as covariates in the following tests of main and interactive effects.

Table 2. Descriptive Statistics for Substantiation and Recidivism (Study 1).

| Initial Allegation Substantiation | Recidivism |
|----------------------------------|------------|
|                                  | No (%)     | Yes (%)   | Overall |
| CRC Unsubstantiated              | 86.94      | 13.06     | 53.27   |
| CRC Substantiated                | 89.37      | 10.63     | 46.73   |
| CRC Overall                      | 88.08      | 11.92     |         |
| FAIR Unsubstantiated             | 92.38      | 7.62      | 51.49   |
| FAIR Substantiated               | 89.91      | 10.09     | 48.51   |
| FAIR Overall                     | 91.18      | 8.82      |         |

Note. CRC = Case Review Committee; FAIR = Field-tested Assessment, Intervention-planning, and Response.

Table 3. Baseline Difference Tests for Incident Characteristics (Study 1).

| Baseline Variable                     | B     | SE   | p     | OR    |
|---------------------------------------|-------|------|-------|-------|
| Male offenderb                        | .040  | 0.038| .289  | 1.041 |
| Male victimb                          | −.007 | 0.037| .839  | 0.993 |
| Service member offenderc              | .036  | 0.037| .338  | 1.037 |
| Service member victimc                | .031  | 0.035| .368  | 1.031 |
| Allegation substantiationd            | .028  | 0.034| .411  | 1.028 |
| Maltreatment severitye                | .010  | 0.030| .728  |       |
| Maltreatment victim age groupf        | .086  | 0.031| .006  | 1.090 |
| Emotional abuseg                      | −.081 | 0.038| .033  | 0.922 |
| Physical abuseg                       | .035  | 0.028| .200  | 1.036 |
| Sexual abuseg                         | −.030 | 0.088| .732  | 0.970 |
| Neglectg                              | .016  | 0.035| .653  | 1.016 |

Note. Covariates (installation) are excluded from this table (see Supplemental Tables S2–S5); FAIR = Field-tested Assessment, Intervention-planning, and Response; CRC = Case Review Committee.

a1 = FAIR, −1 = CRC; b1 = male, 0 = female; c1 = service member, 0 = civilian; d1 = substantiated, 0 = unsubstantiated; e0 = did not meet criteria, 1 = mild, 2 = moderate, and 4 = severe; f1 = child, 0 = adult partner; g1 = present, 0 = absent.

−1 = child) and maltreatment type (emotional, physical, and sexual abuse; reference category is neglect; 1 = present, −1 = absent) as covariates in the following tests of main and interactive effects.
The main effect of phase was statistically significant per the results of logistic regression (Table 4). Based on this model, the predicted probabilities of alleged recidivism were lower in the FAIR (.072) than the CRC (.097) phase.

The Phase × Victim Age Group × Allegation Substantiation interaction was significant as well (Table 5; Figure 1). SS analyses showed that the lower recidivism rate in the FAIR, relative to the CRC, phase was observed for both child (SS = −0.558 [95% CI: −0.894, −0.222], SE = 0.171, p = 0.001, OR = 0.572, relative risk reduction [RRR] = 0.653) and adult (SS = −0.194 [95% CI: −0.373, −0.015], SE = 0.091, p = .034, OR = 0.824, RRR = 0.298) victims for incidents with unsubstantiated maltreatment allegations. In contrast, for incidents with substantiated maltreatment allegations, the recidivism rate was not significantly affected by phase in child (SS = 0.173 [95% CI: −0.107, 0.454], SE = 0.143, p = .226, OR = 1.189, RRR = −0.385) or adult (SS = −0.148 [95% CI: −0.345, 0.050], SE = 0.101, p = .144, OR = 0.862, RRR = 0.233) victims. The predicted probabilities of alleged recidivism are displayed in Figure 1.

**Study 2**

Permission to collect data was obtained from the Army’s Human Research Protection Office, and New York University Institutional Review Board approved the study protocol.

**Methods**

**Participants.** FAP social workers identified Soldiers or partners (N = 88) with allegations indicating victimization by intimate partner physical, psychological, or sexual abuse, referring these individuals to an on-site researcher. The Study 2 participants were distinct from those of Study 1.

**Design.** Utilizing a quasi-experimental design, partner abuse recidivism was compared in the CRC versus FAIR implementation phases (see Study 1 description). Four of the sites in the full study were assigned to this protocol. Of the 88 participants, n = 46 were recruited during the CRC phase; n = 42 were recruited during the FAIR phase.

**Measures.** Alleged partner abuse recidivism was assessed via victim-report using an automated telephone assessment system administered by Northern Illinois University. All responses involved pushing “1” for “yes” and “2” for “no” and were anonymous. The automated assessment included eight items measuring partner abuse (emotional, physical, or sexual) based on the
definitions of partner abuse in FAIR. Participants were asked to call weekly
for 4 weeks; then monthly for 6 months. They were paid $10 to hear about the
study and an additional $10 was uploaded to a prepaid debit card each time
they made a call to the assessment system (up to $110). Partner abuse, of any
type, was scored as present (1) or absent (0) across all waves of assessment.

To protect anonymity, no additional data (e.g., demographics) were pro-
vided to the research team.

**Analytic Method**

Survival and logistic regression analyses were conducted with Mplus ver-
sion 8, using the robust pseudo maximum likelihood estimation method
described for Study 1. To test the effect of phase on time to recidivism, a
Cox proportional hazards model (Singer & Willett, 2003) was estimated

| Predictor       | B    | SE   | p    | 95% CI (B) | OR  |
|-----------------|------|------|------|------------|-----|
| **Covariates by category** |      |      |      |            |     |
| **Installation** |      |      |      |            |     |
| One             | -1.520 | 0.404 | .000 | -2.313 | -0.728 | 0.219  |
| Two             | .444  | 0.202 | .028 | 0.047 | 0.840 | 1.559  |
| Three           | .421  | 0.214 | .050 | 0.001 | 0.841 | 1.523  |
| Four            | .432  | 0.215 | .044 | 0.011 | 0.853 | 1.540  |
| Five            | .364  | 0.214 | .089 | -0.055 | 0.784 | 1.439  |
| Six             | .294  | 0.213 | .167 | -0.123 | 0.711 | 1.342  |
| Seven           | .417  | 0.221 | .059 | -0.015 | 0.849 | 1.517  |
| Eight           | .373  | 0.237 | .115 | -0.091 | 0.837 | 1.452  |
| Nine            | .108  | 0.219 | .621 | -0.321 | 0.537 | 1.114  |
| Victim age group | .182 | 0.070 | .009 | 0.044 | 0.319 | 1.200  |
| **Maltreatment type** |      |      |      |            |     |
| Emotional abuse | .044  | 0.084 | .599 | -0.121 | 0.209 | 1.045  |
| Physical abuse  | .051  | 0.077 | .513 | -0.101 | 0.203 | 1.052  |
| Sexual abuse    | -.071 | 0.135 | .601 | -0.336 | 0.194 | 0.931  |
| **Primary predictor** |      |      |      |            |     |
| Phase           | -.161 | 0.073 | .027 | -0.304 | -0.018 | 0.851  |

aTen is reference installation; bneglect is reference maltreatment type; cphase is coded CRC
(−1) and FAIR (1).
in which time to recidivism was regressed on phase. To test the effect of phase on the presence versus absence of recidivism, a logistic regression model was estimated in which abuse was regressed on phase (CRC = 1; FAIR = 2).

Table 5. Alleged Recidivism: Phase \( \times \) Victim Age Group \( \times \) Allegation Substantiation Interaction (Study 1).

| Predictor | \( B \) | \( SE \) | \( p \) | Low | High | OR |
|----------|--------|--------|-------|-----|-----|----|
| Covariates by category |         |        |       |     |     |    |
| Installation\(^a\) |        |        |       |     |     |    |
| One      | −1.505 | 0.406  | 0.000 | −2.300 | −0.710 | 0.222 |
| Two      | 0.444  | 0.205  | 0.031 | 0.041 | 0.846 | 1.559 |
| Three    | 0.416  | 0.216  | 0.054 | −0.007 | 0.839 | 1.516 |
| Four     | 0.463  | 0.217  | 0.033 | 0.038 | 0.888 | 1.589 |
| Five     | 0.380  | 0.217  | 0.080 | −0.046 | 0.806 | 1.462 |
| Six      | 0.276  | 0.215  | 0.200 | −0.146 | 0.697 | 1.318 |
| Seven    | 0.416  | 0.223  | 0.062 | −0.021 | 0.853 | 1.516 |
| Eight    | 0.394  | 0.239  | 0.099 | −0.075 | 0.864 | 1.483 |
| Nine     | 0.130  | 0.222  | 0.558 | −0.305 | 0.566 | 1.139 |
| Maltreatment type\(^b\) |        |        |       |     |     |    |
| Emotional abuse | 0.062  | 0.086  | 0.473 | −0.107 | 0.230 | 1.064 |
| Physical abuse  | 0.066  | 0.077  | 0.390 | −0.085 | 0.218 | 1.068 |
| Sexual abuse    | −0.051 | 0.136  | 0.707 | −0.318 | 0.216 | 0.950 |
| Primary predictors |        |        |       |     |     |    |
| Phase\(^c\) | −0.181 | 0.079  | 0.022 | −0.337 | −0.026 | 0.834 |
| Victim age group\(^d\) | 0.201  | 0.077  | 0.009 | 0.050 | 0.351 | 1.223 |
| Allegation substantiation\(^e\) | 0.034  | 0.061  | 0.579 | −0.085 | 0.153 | 1.035 |
| Phase \( \times \) victim age group | 0.011  | 0.059  | 0.856 | −0.105 | 0.126 | 1.011 |
| Victim age group \( \times \) allegation substantiation | −0.115 | 0.060  | 0.055 | −0.232 | 0.003 | 0.891 |
| Phase \( \times \) allegation substantiation | 0.194  | 0.061  | 0.001 | 0.075 | 0.313 | 1.214 |
| Phase \( \times \) victim age group \( \times \) allegation substantiation | −0.171 | 0.060  | 0.005 | −0.290 | −0.053 | 0.843 |

\(^a\)Ten is reference installation; \(^b\)neglect is reference maltreatment type; \(^c\)phase is coded CRC (−1) and FAIR (1); \(^d\)victim age category is coded as child (−1) and adult partner (1); \(^e\)allegation substantiation is coded as unsubstantiated (−1) and substantiated (1).
Figure 1. Phase × Victim Age Category × Allegation Substantiation interaction (Study 1). SS for the effect of phase (CRC vs. FAIR) are statistically significant for both child (p = .001) and adult (p = .034) victims for incidents with an unsubstantiated baseline allegation. Neither SS is significant for incidents with substantiated baseline allegations.

Note. SS = simple slopes; CRC = Case Review Committee; FAIR = Field-tested Assessment, Intervention-planning, and Response.
Results

Time to recidivism was significantly lower for cases determined via CRC versus FAIR, \( B = -0.337, SE = 0.064, p < .001, 95\% CI: -0.462 \text{ to } -0.212 \) (Figure 2). Median time to recidivism was 33.48 (CRC) versus 90.10 (FAIR) days.

The likelihood of recidivism at any time was also significantly lower in the FAIR than in the CRC phase, \( OR = 0.361 \) (\( SE = 0.033, p < .001, 95\% CI: 0.311 \text{ to } 0.432 \)). Based on predicted probabilities, the recidivism rate was 69.55\% in the CRC phase and 45.21\% during in the FAIR phase; \( RRR = 0.35 \).

Discussion

Results replicate and extend prior findings (Snarr et al., 2011) that the FAIR system for making maltreatment determinations results in notable tertiary preventative effects (i.e., 65\% reduction in recidivism in child maltreatment and 30\% to 35\% reduction in adult maltreatment among unsubstantiated incidents). Snarr et al. (2011), using base-level U.S. Air Force FAP partner and child maltreatment allegation data, could not isolate whether the decrease in reporting under the FAIR system was due to (a) reduced subsequent maltreatment or (b) reduced reporting of maltreatment. The current study included anonymous victim reporting; findings suggest that the FAIR system does, indeed, reduce risk of subsequent maltreatment incidents.

Whereas the Snarr et al. (2011) study found that the reduction in recidivism was driven by the families with previously substantiated cases, the current study finds this reduction in families with previously unsubstantiated cases. It seems important to note that the original substantiation rates in the Snarr et al. (2011) sample (56\%, which moved to 47\% after the shift to the FAIR) differed from this sample (47\%, which moved to 48.5\%). In addition, in this study, the same definitions of maltreatment were used in both conditions, and this was not the case in Snarr et al. Therefore, the patterns of substantiation changes and the slightly different interventions may be able to explain the different interactions found in each study, although inconsistency does make the interaction less interpretable.

Military communities’ responses to maltreatment use a coordinated community response approach, which translate ecological theories (Bronfenbrenner, 1979) into action. Such theories suggest individuals’ behaviors are impacted by nested social systems (e.g., couple/family, workplace, installation, larger civilian community, subculture, culture) at multiple levels. Prior studies have found that both Air Force (Heyman et al., 2010) and Army (Heyman et al., 2022) unit commanders and first sergeants are more likely to attend FAIR-system meetings and perceive the system as fair. Thus, it is plausible that the expected positive
impact on the coordinated community response elicited by including the chain of command directly in the maltreatment decision-making process resulted in lower recidivism.

No other field-tested maltreatment definitional criteria are in the literature, and other systems for maltreatment determinations have notable shortcomings. The FAIR system appears to improve upon standard practice based on prior research (Heyman & Slep, 2006; Slep & Heyman, 2006; Snarr et al., 2011) and the findings in this report. Although designed for use in the U.S. military, the FAIR system has been ported over to one state child welfare system with some evidence of success (Mitnick et al., 2022). Translation to very large maltreatment systems might require tailoring of the FAIR system. However, this study took place in the 10 largest FAPs at U.S. Army installations, which care for nearly 50,000 Soldiers and their families. These settings are likely as demanding as many local civilian child welfare settings. Furthermore, the structure of FAIR includes training in assessment and the structured system, criteria, and computerized decision-making tools, which could reasonably be expected to improve social work practice in non-military settings.

As with all studies, this study has several limitations. First, this study, as with the prior evaluation, used a quasi-experimental design. Participants did not choose, or have awareness of, the condition to which they belonged, thus

Figure 2. Recidivism survival curves by condition (Study 2).

Note. CRC = Case Review Committee; FAIR = Field-tested Assessment, Intervention-planning, and Response.
avoiding a major threat to the validity of causal interpretations in quasi-experimental designs: selection effects (Cook et al., 2002). Furthermore, we controlled for potential confounds that were available in the dataset. However, at all sites, the CRC condition preceded the FAIR condition, thus other threats cannot be decisively ruled out (e.g., history). Yet, in the absence of a specific identifiable threat to the causal validity of our findings, we believe the differences in recidivism were most plausibly caused by the change from the CRC to the FAIR system. Second, the sample size in Study 2 was relatively small, likely due to participation demands, and it may not be representative of all victims. Third, further, we did not have access to information about participants’ racial and ethnic identity or many other demographic factors, so we cannot speak on generalizability of our findings across those subgroups nor rule out the possibility that participants in the two groups had demographic differences that might confound group differences in recidivism, although this seems unlikely. Fourth, because these data are anonymous, we cannot evaluate whether Study 2 participants’ initial maltreatment cases differed from those not invited to participate or those who chose not to participate. However, the consistency of the findings to those of Study 1, despite the small size, make the generalizability of its findings likely. Finally, these results cannot be considered indicative of the potential prevention impacts of a FAIR implementation in civilian contexts without the necessary evaluation. The military’s coordinated community response includes the workplace as an important element in fostering change. Because civilian systems do not have this subsystem of the family’s ecology, implementing FAIR there involves less leverage. Alternatively, if the active ingredient in FAIR is the family’s understanding of the criteria’s bright line indicating maltreatment, it may be that workplace representatives merely emphasized this line, something that could be undertaken by those in the civilian system with whom families interact. Further research is necessary to understand these pathways.

To date, the FAIR system has been implemented across the U.S. military and Alaska (Mitnick et al., 2021). Given the current studies’ replication that the FAIR system reduces subsequent incidents of maltreatment, wider dissemination of the FAIR system to make substantiation determinations in U.S. and international jurisdictions should be considered.

**Conclusion**

Clear and consistent decisions about whether alleged incidents of child maltreatment or intimate partner abuse are founded contribute to the perceived fairness of the process and support bright-line distinctions about what constitutes maltreatment, which in turn, support healthy community norms around parenting and
intimate relationships. Although not all families who encounter formal systems are able to learn and apply the distinction between maltreating and non-maltreating behaviors, some can. FAIR reduces re-offense without added time or expense to the system (see Heyman et al., 2022) or the families once implemented. Thus, the cost effectiveness for reduction in injury, mental health and developmental consequences, and mortality is potentially substantial.

The FAIR system is implemented via policy throughout the U.S. Department of Defense and the state of Alaska. Standardized trainings, automation systems, and quality assurance systems have been developed and are being routinely used in most settings, with no support from the system developers. This suggests the potential for sustainability is high. Research to date, including that presented here, suggests the FAIR system would represent a notable improvement over current practice for making substantiation decisions.

Acknowledgments

This project would have been impossible without the coordination and collaboration with U.S. Army colleagues (Army Medical Command Family Advocacy Program: Dr. Rene Robichaux, Dr. Jennifer Humphries, Ms. Lucinda Geeslin, Lt. Col. Tanya Juarez, Capt. Ryan Black; Army Installation Command Family Advocacy Program: Ms. Cynthia Taylor, Lt. Col. Ricky Martinez, Col. Steve Lewis; Army Medical Department Center and School: Col. Anthony Cox, Lt. Col. Erica Dijoseph; Office of the Deputy Army Chief of Staff, G-9: Dr. Najah Barton; Installation Command G-9: Ms. Kimberly Williams; Army regional consultants: Warren Aoki, Michele Barber, Dwayne Elder); NYU research coordinators embedded at the 10 Army posts: Jamie Anderson, David Ascher, Sara Bauman, Lisa Bell, Nicole Calhoun, Alayna Colburn, Kaleena Cook, Demetra Cropp, Amber Crozier, LaRita Dalton, Guadalupe Drazkowski, Nadia Haughton, Heather Henderson, Aimee Henson, Caila Holm, Michale Jaurigue, Andrew Phelps, Alexis Proctor, Carol Rickmon, Maria Rodriguez, Diana Rodriguez, Chenille Smith, Heather Solley, Dominic Stewart, William Stucker; and the FAP personnel, Soldiers and Families, and CRC and FAIR members at each of the 10 installations.

Data Availability

These data were collected as part of a Cooperative Agreement with the U.S. Army. The data that support the findings of this study are available from U.S. Army Family Advocacy Program, but restrictions apply to the availability of these data, such that private information is protected and Human Research Protections Office protections are maintained, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of U.S. Army Family Advocacy Program.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interests with respect to the authorship and/or publication of this article.
Disclaimer
The views expressed are solely those of the authors and do not reflect the official policy or position of the U.S. Army.

Funding
The author(s) disclosed receipt of the following financial support for the research and/or authorship of this article: US Department of Agriculture 2015-48783-24394.

ORCID iD
Danielle M. Mitnick [ID] https://orcid.org/0000-0003-4958-2788

Supplemental Material
Supplemental material for this article is available online.

References
American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596
Asparouhov, T., & Muthén, B. (2005). Multivariate statistical modeling with survey data [Conference session]. Federal Committee on Statistical Methodology (FCSM) Research Conference.
Bronfenbrenner, U. (1979). The ecology of human development: Experiments by nature and design. Harvard University Press.
Cook, T. D., Campbell, D. T., & Shadish, W. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.
Department of Defense. (2016). DODM 6400.01, Volume 3 Family Advocacy Program (FAP): Clinical Case Staff Management (CCSM) Incident Determination Committee (IDC).
Erlanger, A. C. E., Heyman, R. E., & Slep, A. M. S. (2022). Creating and testing the reliability of a family maltreatment severity classification system. Journal of Interpersonal Violence, 37(7–8), NP5649–NP5668. https://doi.org/10.1177/0886260520961866
Foran, H. M., Beach, S. R. H., Slep, A. M. S., Heyman, R. E., & Wamboldt, M. Z. (Eds.). (2013). Family problems and family violence: Reliable assessment and the ICD–11. Springer Publishing Co.
Heyman, R. E., Collins, P. S., Slep, A. M. S., & Knickerbocker, L. (2010). Evidence-based substantiation criteria: Improving the reliability of field decisions of child maltreatment and partner abuse. Protecting Children, 25, 35–46.
Heyman, R. E., & Slep, A. M. S. (2006). Creating and field-testing diagnostic criteria for partner and child maltreatment. Journal of Family Psychology, 20(3), 397–408. https://doi.org/10.1037/0893-3200.20.3.397
Heyman, R. E., & Slep, A. M. S. (2009). Reliability of family maltreatment diagnostic criteria: 41 site dissemination field trial. *Journal of Family Psychology, 23*(6), 905–910. http://dx.doi.org/10.1037/a0017011

Heyman, R. E., & Slep, A. M. S. (2019). Relational disorders and beyond. In B. H. Fiese, M. Celano, K. Deater-Deckard, E. N. Jouriles, & M. A. Whisman (Eds.), *APA handbook of contemporary family psychology: Family therapy and training* (pp. 19–34). American Psychological Association. https://doi.org/10.1037/0000101-002

Heyman, R. E., Slep, A. M. S., Lorber, M. F., Xu, S., & Hillenbrand, T. M. (2016). Quality assurance testing of the Air Force Central Registry Board [Unpublished technical report]. Family Translational Research Group, New York University.

Heyman, R. E., Slep, A. M. S., Mitnick, D. M., Nichols, S. R., Cracknell, K. M., Tiberio, S. S., Kim, S., & Perkins, D. F. (2022). Evaluation of two approaches for responding to allegations of family maltreatment in the U.S. Army: Coordinated community response impacts and costs. *Military Medicine, 187*(7–8), e987–e994. https://doi.org/10.1093/milmed/usab115

Mitnick, D. M., Slep, A. M. S., Heyman, R. E., Lively, D., & Perkins, D. F. (2022). *The FAIR System for child maltreatment substantiation determinations: An evaluation of Alaska’s statewide implementation* [Manuscript submitted for publication]. Family Translational Research Group, New York University.

Muthén, L. K., & Muthén, B. O. (2017). *Mplus user’s guide* (8th ed.). Muthén & Muthén.

Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics, 31*(4), 437–448. https://doi.org/10.3102/10769986031004437

Rettew, D. C., Lynch, A. D., Achenbach, T. M., Dumenci, L., & Ivanova, M. Y. (2009). Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. *International Journal of Methods in Psychiatric Research, 18*(3), 169–184. https://doi.org/10.1002/mpr.289

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195152968.001.0001

Slep, A. M. S., & Heyman, R. E. (2006). Creating and field-testing child maltreatment definitions: Improving the reliability of substantiation determinations. *Child Maltreatment, 11*(3), 217–236. https://doi.org/10.1177/1077559506288878

Snarr, J. D., Heyman, R. E., Slep, A. M. S., Malik, J., & United States Air Force Family Advocacy Program. (2011). Preventive impacts of reliable family maltreatment criteria. *Journal of Consulting and Clinical Psychology, 79*, 826–833. https://doi:10.1037/a0025994

World Health Organization. (2021). *ICD-11 for mortality and morbidity statistics*. Retrieved September 24, 2021, from https://icd.who.int/dev11/1-m/en#/

**Author Biographies**

Amy M. Smith Slep, PhD, is a clinical psychologist, professor, and co-director of the Family Translational Research Group at New York University. Her research focuses on
anger, conflict, and violence in families and its prevention. She is interested in leveraging implementation science to further system improvements to prevent violence.

**Richard E. Heyman**, PhD, is professor in the Department of Cariology and Comprehensive Care at New York University. He earned a BS from Duke University and a PhD in clinical psychology from the University of Oregon. He’s a licensed psychologist. Dr. Heyman has received over 60 grants/contracts from major U.S. funding agencies on a variety of topics, from dental fear to social determinants of health to couples communication to community-level prevention of family maltreatment, substance problems, and suicidality. Dr. Heyman has published over 200 scientific articles/chapters on these topics. At the core of Dr. Heyman’s research is translating basic knowledge into prevention and treatment and on improving adoption of evidence-based practices.

**Danielle M. Mitnick**, PhD, is a clinical assistant professor at New York University College of Dentistry and the Director of Family Clinical and Prevention Research at the Family Translational Research Group. She is a licensed clinical psychologist. Her research interests include promotion of health within families and preventing family maltreatment.

**Michael F. Lorber**, PhD, is a senior research scientist with the Family Translational Research Group in the College of Dentistry at New York University. His primary research interests are centered on externalizing behaviors—their form, development, etiology, consequences, and prevention—from infancy through adulthood, primarily in relational (e.g., parent-child; couple) contexts.

**Sara R. Nichols**, PhD, is a developmental and clinical psychologist and an adjunct assistant professor at New York University College of Dentistry. Her research interests include early social development, developmental psychopathology, and implementation of evidence-based interventions.

**Daniel F. Perkins**, PhD, is a professor of Family and Youth Resiliency and Policy at the Pennsylvania State University. As Principal Scientist of an applied research center, the Clearinghouse for Military Family Readiness, he translates science into action through science-based programs and technical assistance to professionals serving military/veterans families. His work involves implementation, hybrid evaluations of interventions, and implementation science.