

MEASURING ACTUAL REDUCTION OF RISK TO CHILD ABUSE: A NEW APPROACH

GEORGE E. FRYER, JR., M.A., M.S.W.

Instructor of Family Medicine and Pediatrics and Coordinator of Research,
University of Colorado School of Medicine, Denver

SHERRYLL KERNS KRAIZER, M.S.

Director and President of Health Education Systems, Inc., Palisades, NY

THOMAS MIYOSHI, M.S.W.

Kempe National Center, Department of Pediatrics, University of Colorado School of Medicine, Denver

Abstract—Previous efforts to measure the effectiveness of child abuse prevention programs have relied on proximate measures presumed to be predictive of actual skills. This paper presents documentation that brings those assumptions into question, describes actual observation and measurement of behavioral change in children before and after prevention education, and correlates that behavioral response with more traditional measures of effectiveness. Unique to the evaluation was the staging of an actual situation in which each of the children had an opportunity to leave the school building with a stranger. Each simulation was videotaped and conducted in such a way that the children remained unaware of the fact that they had been tested. In addition, tests of language development, self-esteem and knowledge of prevention and safety concepts were administered before and after participating in the Children Need to Know Personal Safety Training Program [1]. Several findings have significant value for future examinations and programming. The effectiveness of a primary prevention program based on age-appropriate, experiential and interactive instruction was empirically documented. Traditional instrumentation which elicits written or verbal responses to cognitive questions about safety may be misleading in assessing children's vulnerability. Higher self-esteem before instruction and higher knowledge/attitude scores after instruction were found to be predictive of a reduction in vulnerability. While this is a significant series of findings, some children did not achieve the objectives of the prevention program. These results suggest further possibilities for evaluation and some direction for improving prevention education.

Résumé—Jusqu'à ce jour quand on voulait quantifier l'efficacité d'un programme préventif de maltraitance, on se fondait surtout sur une appréciation approximative et immédiate qui était censée avoir une valeur prédictive quant aux compétences réelles acquises grâce au programme. Le présent article apporte des données qui font douter de la validité de telles estimations: ces données résultent de l'observation d'un réel changement de comportement des enfants soumis à un programme éducatif de prévention avant et après l'application du dit programme: le changement de comportement est mis en relation avec des estimations plus traditionnelles de l'efficacité. Ce qu'il y avait d'unique dans cette évaluation, c'était la mise en place d'une situation réelle dans laquelle chacun des enfants testés avait l'occasion de quitter le bâtiment de leur école accompagné par un étranger. Cette simulation était enregistrée sur bande vidéo et conduite de telle sorte que les enfants impliqués ne se doutaient pas qu'ils étaient en train d'être soumis à un test. En plus avant et après la participation des enfants au programme de prévention intitulé "Les enfants doivent connaître les règles de la sécurité personnelle," des tests du développement du langage, de l'appréciation de soi-même et de la connaissance de concepts de prévention et de sécurité étaient administrés aux enfants. Plusieurs des observations faites pourront servir à de futures évaluations et programmations. On a ainsi obtenu par une méthode empirique une évaluation de l'efficacité d'un programme de prévention primaire fondé sur une instruction appropriée à l'âge, experiential et interactive. L'instrumentation traditionnelle qui conduit à obtenir des réponses orales ou écrites à des ques-

tions formelles sur la sécurité personnelle pourraient bien induire en erreur en ce qui concerne l'évaluation de la vulnérabilité des enfants. Une idée de soi-même élevée avant l'instruction spécifique et des scores élevés quant aux connaissances des attitudes après l'administration du programme d'instruction se sont révélés être de bons index prédisant une réduction de la vulnérabilité. Cela est vrai en général mais quelques enfants n'ont pas du tout atteint l'objectif du programme de prévention et les résultats observés suggèrent aussi d'autres possibilités d'évaluer et d'améliorer les programmes éducatifs visant à prévenir la maltraitance.

INTRODUCTION

THERE IS A RICH HISTORY of innovative and well-conceived attempts to demonstrate the value of programs for preventing the abuse of children. But assessing the effectiveness of preventive interventions is a difficult undertaking. Formidable obstacles to objective evaluation, upon which compelling arguments of the benefits derived from programming could be predicated include unreported abuse, matters associated with confidentiality, the ethics of attempting to sample child behavior in actual abuse situations, and a variety of others. Most evaluators have therefore elected to focus on proximate measures which are "presumably linked to desired ultimate outcomes. . . . Unfortunately, there is often little research evidence that the purported relationships hold" [2]. The relevance of traditional measurement of a child's understanding of concepts, with the assumption of a relationship to behavior resistant to victimization has recently been challenged [3, 4]. This paper presents documentation of the effectiveness of the prevention program, Children Need to Know Personal Safety Training Program [1], obtained by observing and measuring the behavioral change of children in simulated situations and correlating that behavioral response with more traditional measures of program effectiveness.

METHODS

Students enrolled in kindergarten, the first and second grades of a midtown Denver elementary school took part in the study. Twenty-four each were randomly assigned to the experimental and control groups tested. Of these, 21 control group and 23 experimental group children were present for school on both testing days and were included in the study. A pretest-posttest control group design, otherwise known as the classical experimental design, was employed. It controls for all factors which jeopardize internal validity [5] and is the most widely used of the true experimental designs [6]. The respective groups were not related/matched; children were not paired on predetermined criteria [7], although their memberships were very similar in attributes. But for the purpose of comparisons made for before and after test results, children were matched against themselves at two different points in time [7].

Experimental group children participated in an 8-day block of child abuse prevention instruction consisting of 20-minute presentations each day. Control group subjects received no instruction, but were scheduled to take part in the program in the near future. The 8-session classroom program emphasized discovering and clarifying existing misconceptions children have about their personal safety. For example, participating children were not told that strangers are bad or dangerous, but merely "people you don't know" and an everyday part of life.

The program further attempted to establish simple, concrete rules and specific criteria for the application of those rules. The children were asked to follow the rules only when they were not with caretaking adults. In fact, children were encouraged to meet and interact with people they had not met before when in the presence of caretaking adults. The rules were as follows:

1. Stay an arm's reach away.
2. Don't talk or answer questions.
3. Don't take anything.
4. Don't go anywhere.

In a situation such as one simulated in the school setting, therefore, children, when by themselves, would be expected to follow the rules and not interact if approached by a stranger.

The program provided extensive opportunities for individual children to exercise those rules and concepts through role-playing. We hypothesized that children who can parrot concepts may not be able to actually use those skills in real life situations. Role-playing gives children an experiential base for future consideration, thoughtfulness and confident decision making when they find themselves in a situation associated with risk for child abuse.

Immediately before the classes, the Peabody Picture Vocabulary Test (PPVT), a non-verbal, multiple choice examination, "useful in measuring extensiveness of receptive vocabulary and (an instrument that) may serve as a screening device for children with a limited expressive vocabulary or for children who are verbally inhibited in a testing situation," was administered [8]. It measured not actual risk to abuse, but rather the ability of the children to master cognitively the concepts that were presented.

The students were subjected to two additional tests both prior to and following the instruction: the Harter Perceived Competence Scale for children [9] and the Children Need to Know Knowledge Attitude Test [1]. The Harter Perceived Competence Scale for Children is a self-report instrument strong in psychometric properties, designed to measure the competence of children across one general worth and three specific skill domains [9]. It is a commonly applied instrument measuring self-esteem, a construct sometimes theorized as being related to abuse and otherwise clinically important [9]. The Children Need to Know Knowledge-Attitude Test [1], a 20-item examination, was administered to measure the children's cognitive awareness and understanding of issues associated with risk and prevention.

Unique to the evaluation was the staging with each child of an actual situation the day before and the day after the classroom program, by which the child's response reflected the degree of vulnerability to abuse. This simulation, enacted only after extensive discussion and collaboration with parents and officials of the school and school district, was comprised of the child's encounter with a member of the research team in an isolated setting in the school. The research assistant, posing as a stranger, requested the child's assistance in the performance of a task which entailed their leaving the school building together. Acquiescence would clearly create the potential for an act of abuse.

Two scenarios were constructed. In each, the would-be stranger requested the child's assistance. In the pretest, he asked the child to accompany him to his car to help bring treats in for his son's birthday party. In the other simulation, the posttest, the child was similarly asked to come to the stranger's car to bring in puppets to be used in a puppet show: "Hello, I'm presenting a puppet show here at the school today. I have puppets and other neat things outside in my car. Will you come out and help me bring them inside?" If the child agreed, he/she was told that the stranger would come for him/her later for help. After all children had been through the simulation and had returned to their classrooms, they were told that the stranger had gotten help from the school's main office and would therefore not need their assistance.

Each simulation was carried out in a way that could not be expected by the children, who were unaware that they were, in fact, being tested. They were not informed after the fact that they had had a simulated encounter. In both instances, teachers had fabricated

an excuse to have the child leave the classroom and move to the designated area for the confrontation with the stranger. Collaborating office staff, classroom teachers, and school nurses took steps to insure that no one else was present in the hallway in which the simulation occurred and that the child could have departed the building with the stranger unseen by school personnel and/or other students. Thus, the major criterion to judgment of risk to the child's becoming victimized was his/her compliance/noncompliance with the stranger's request and other specific behavior in response.

A hidden camera and wireless microphone produced an audiovisual record of the encounter which was later reviewed and scored by research team members. A straightforward pass-fail rating was awarded the child based on performance during the simulation. It connoted simply whether or not the student agreed to the stranger's request. Interrater reliability among the four evaluation team members was 1.0 (total reliability). There was no disagreement concerning any child's expression of willingness to leave the school building with the simulator. The entire simulation process had been previously piloted with a class of kindergarten children.

SPECIAL CONSIDERATIONS

Even before discussing findings attained, the authors emphasize what we believe to be the extreme importance of attentiveness to the emotional well-being of children participating in the study. Extreme care was taken in the creation of these simulations to protect the children's emotional well-being and to assure that the simulations were perceived as being well within the norms of everyday occurrences. Formal meetings were held with administrators of the Denver Public Schools to insure that they were fully cognizant of the protocols to be employed and to entertain any questions or concerns.

Parents were sent a letter detailing explicitly the format, process and content of the program. A positive permission form was used to assure that parents had judged participation appropriate for their child. Finally, a parent seminar was held where the concept and planned protocol were fully explained and all questions answered prior to initiation of the evaluation.

In designing the simulation, the school setting was selected because it is a protected and controlled environment. The "strangers" were conservative in appearance. The specific statements of both simulations were plausible and nonthreatening.

All personnel who took part in testing and staging the simulations were trained in social work and/or child development and had experience working with children of this age. A social worker observed on a television monitor each child's simulation and was prepared to terminate the simulation and attend to the child at the first indication of anxiety.

After the simulation, each child spent the next 30 minutes in a one-to-one meeting with a member of the research team. This allowed ample opportunity to express any fear or anxiety as well as to report the encounter.

Logistically, this made the process difficult, time-consuming, and labor intensive. But in our judgment this considerable allocation of resources is necessary and we should not be made complacent by the fact that none of the 44 children required the assistance for which provision had been made.

Only 1 of the 44 children expressed even minor anxiety about the simulation. This took the form of the child's straightforward request that a member of the research team investigate the activities of the stranger in the hallway. A researcher spoke with the child later that day and then stringently observed the child during the next two weeks. In this way we ascertained that there was no lingering effect or disturbing memory of the event. We

later learned from teachers and other school personnel that in the past this child sometimes had become anxious in situations not considered stressful by others.

The diligent procedures implemented in an effort to curb anxiety on the part of the children are essential to this process and should not be omitted nor less emphasized in its replication. The authors welcome contact with others contemplating similar delicate research.

RESULTS

There was a small difference in the scores of the two study groups on the pretest simulation (See Table 1). About half failed. Ten of the 21 (47.6%) control group children agreed to accompany the stranger vs. 13 of the 23 (56.5%) experimental group members. During the posttest, again 10 of 21 (47.6%) control group students present failed compared with 5 of the 23 (21.7%) experimental group children. Chi-square for the posttest difference between groups was statistically significant, $p = .05$. The improvement from pretest to posttest simulation outcome for the experimental group was also significant, $p = .02$.

The average scores for the 21 children who passed the pretest simulation were virtually the same on each of the four Harter component scales as those for the 23 who failed. Those who successfully completed the pretest simulation scored slightly higher on the PPVT and knowledge-attitude test but differences did not approach statistical significance when t -tested. Pretest scores on each of the three tests administered were very nearly the same for the two study groups. The various measures were not predictive of performance in the simulation nor revealing of differences in group membership. When at possible risk, neither child perceptiveness, self-esteem, nor ability to verbally state appropriate action, as measured by the knowledge-attitude test, were associated with vulnerability to abduction and abuse as indicated by performance during the first simulation.

But among the experimental group members, those who ultimately passed the second simulation were statistically significantly superior to others in that study group on both the composite pre-Harter self-esteem test score, $p = .05$, and that for the post-intervention knowledge-attitude test, $p = .02$. These indices did discriminate outcome. In the absence of intervention, self-esteem and knowledge of concepts were not useful in determining risk. But children with high esteem before intervention who improved through participation their mastery of concepts and techniques exhibited behavior much more preventive to victimization than did other children in the same group. Esteem would appear to be a desirable pre-condition to the intervention, enabling assimilation of the information presented as then measured by score on the knowledge-attitude test. Mastery of the crucial concepts contained in the knowledge-attitude test when founded on participation in experiential instruction laden with role-playing life-like scenarios, in turn, equates with real reduction in risk.

The existence of this critical path is further verified by examination of data for the 13 experimental group children who failed the pretest simulation. Eight of the 13 passed the second simulation. The 8 who passed had scored much higher on the pretest Harter self-

Table 1. Simulation Outcomes

| Group | Pretest simulation | | Posttest simulation | |
|--------------|--------------------|------------|---------------------|------------|
| | Passed | Failed | Passed | Failed |
| Experimental | 10 (43.5%) | 13 (56.5%) | 18 (78.3%) | 5 (21.7%) |
| Control | 11 (52.4%) | 10 (47.6%) | 11 (52.4%) | 10 (47.6%) |

esteem test (composite score) than had the 5 children who again failed (81.4 vs. 67.8). The 5 repeated failures actually improved their composite scores on the self-esteem test approximately twice as much as the 8 children who passed (10 vs. 5.3 points, respectively) the second simulation. But the 8 children who passed enhanced their knowledge-attitude test scores by 42.9% more (6 vs. 4.2 points) than experimental group children who did not pass the pretest simulation.

The discriminant power of these two independent variables, their utility in distinguishing between the two groups that differed in outcome on the second simulation, was examined. Discriminate analysis, a previously described [10-12] multivariate technique, was employed. Discriminant functions, linear combinations of variables, produce scores which separate groups and thereby facilitate classification. The categories for the purpose of this analysis were (1) passed and (2) failed the second simulation. The function calculated with just the pretest Harter self-esteem score and posttest knowledge-attitude test score as discriminating variables correctly classified 12 of the 13 outcomes (92.3%) of experimental group children who had failed the first simulation. Thus, these two variables were found to be extremely predictive of outcome.

DISCUSSION

Discriminate analysis is usually based on this kind of comparison of the characteristics of cases for which success or failure is known [13]. It enables the investigator to go beyond mere portrayal of group differences and association among variables, to foretelling outcomes with accuracy superior to that achievable in the absence of information. But the small number of cases examined through this research gives cause to seek future opportunities to validate these results using larger study groups.

Earlier we noted that the experimental group performed statistically significantly better on the posttest simulation than (1) it did on the pretest simulation ($p = .02$), as well as significantly better than did the control group on the posttest simulation ($p = .05$). The experimental group had not done as well as the control group on the pretest simulation. Within the control group there was modest directionally balanced movement (three from pass to fail and three from fail to pass).

This suggests certain possibilities. First, the forms of the test (the two simulations) are not perfectly parallel. Second, young children cannot be expected to respond with total consistency to situations of this type. And third, professional assessment of risk, particularly in cases in which the child's response was ambivalent or his actions contradictory, remains a difficult task. Yet, especially validating of the comparison as well as convincing of the effectiveness of the program is the fact that change in the experimental group was unidirectional. Eight of 13 who failed the initial simulation passed the subsequent simulation. But none of the 10 children who passed the first simulation failed the second. From this observation outcome appears not to be an arbitrary matter and the instrumentation devised for use in this study appears to be sound and consistent with purpose.

This result is of major importance. One might hypothesize that there is no urgent need to teach children who demonstrate, by means of performance in a simulation such as the ones implemented in this study, the ability to cope with the advances of a stranger. But the failure on the posttest simulation of three children who had previously passed, constitutes a basis for exercising caution in so concluding. Had it not been for receipt of instruction, certain of the 10 experimental group children who had passed the first simulation might have been expected to fail the later one. The program served in that respect to maintain and reinforce appropriate behavior.

tar
wa
by
ag
tio
on
in
sel
tio
a r
inc
tio
wh
par
unc
anc
dur
bet
tha
ver
ide
pro
gra
shc
Ack
this
1.
2.
3.
4.
5.
6.
7.
8.
9.
10.
11.
12.
13.

CONCLUSION

The results of this evaluation, using innovative methods and criteria, are clearly important and merit brief restatement. The effectiveness of the child abuse prevention program was empirically documented. The need for this and similar programming was evidenced by the fact that most of the children participating in the simulation prior to intervention agreed to leave the safe environment of their school building with a total stranger. Traditional instrumentation which elicits written or verbal response to cognitive questions was on balance not helpful in assessing a child's reaction to a stranger's request in a situation in which the child was at risk for abduction and abuse. But higher scores on the pretest self-esteem examination and higher scores on the posttest knowledge-attitude examination were predictive of a reduction in vulnerability to abuse and abduction as measured by a refusal to go with the stranger in the simulation. Only after the intervention did those indices take on meaning. Central to the entire scheme of analysis is the concept of simulation, a situation believed to be actual by participating children, and a standard against which the validity of other measures of risk to abduction and abuse may be judged.

These findings suggest exciting new possibilities. But because the number of children participating in the study was small, its replication is in order. Additional work is now underway in the development of alternative/parallel forms of nonthreatening simulation and better ways of assessing the intent and resolve of children based on their behavior during testing.

This work gives cause to evaluate again our child abuse prevention efforts using still better methods clearly more suited to the task. In using this methodology, the benefits that accrue from participation in one of this country's most widely used child abuse prevention programs have, in effect, been documented. Yet, a small group of children was identified who did not achieve outcomes which constitute the primary objectives of the program. These results suggest general directions to explore in attempting to refine programming to provide all children with protective skills which they may confidently use, should need arise, without harboring unfounded fear of their environment.

Acknowledgment—The authors wish to thank the Denver Public Schools for their cooperation and support of this research.

REFERENCES

1. KRAIZER, S. K. *Children Need to Know: Personal Training Program*. Health Educations Systems, Inc., New York (1981).
2. WEISS, C. H. *Evaluation Research*, p. 37. Prentice-Hall, Englewood Cliffs, NJ (1972).
3. CONTE, J. R., ROSEN, C., SAPERSTEIN, L. and SHERMACK, R. An evaluation of a program to prevent the sexual victimization of young children. *Child Abuse & Neglect* 9:319-328 (1985).
4. SWAN, H. L., PRESS, A. N. and BRIGGS, S. L. Child sexual abuse prevention: Does it work? *Child Welfare* 4:395-405 (1985).
5. GRINNELL, R. M. *Social Work Research and Evaluation*. F. E. Peacock, Itasca, IL (1985).
6. CAMPBELL, D. J. and STANLEY, J. C. *Experimental and Quasi-experimental Designs for Research*. Houghton Mifflin, Boston (1963).
7. SHARP, V. F. *Statistics for the Social Sciences*. Little, Brown, Boston (1979).
8. SATTLER, J. M. *Assessment of Children's Intelligence and Special Abilities*, p. 272. Allyn and Bacon, Boston (1982).
9. HARTER, S. The perceived competence scale for children. *Child Development* 53:87-97 (1982).
10. COOLEY, W. W. and LOHNES, P. R. *Multivariate Data Analysis*. Wiley, NY (1971).
11. TATSUOKA, M. M. *Multivariate Analysis*. Wiley, NY (1971).
12. VAN DE GEER, J. P. *Introduction to Multivariate Analysis for the Social Sciences*. W. H. Freeman, San Francisco (1971).
13. NORUSIS, M. J. *SPSS/PCPLUS Advanced Statistics*. SPSS, Chicago (1986).