The Epidemiology and Impact of Workplace Violence in Pennsylvania Teachers and Paraprofessionals

Request for Office of Management and Budget Review and Approval for Federally Sponsored Data Collection

Section B

Project officer: Hope Tiesman, MSPH, PhD
Epidemiologist
National Institute for Occupational Safety and Health
Division of Safety Research
1095 Willowdale Road, MS 1811
Morgantown, WV, 26505

Phone: 304-285-6067
Fax: 304-285-6235
E-mail: htiesman@cdc.gov
June 2009

Table of Contents

B.1   Respondent Universe and Sampling Methods

There are 501 school districts throughout the 67 counties of Pennsylvania (PDE Website, 2007). In the 2005 and 2006 school year, there were 150,613 professional personnel, of which, 82% (n=123,395) were classroom teachers (2007). There were 16,593 education workers in the 'Coordinate Services' category, also known as paraprofessionals (2007). This could include instructional aides, administrative supportive staff, library support staff, and 'all other support staff'.

The state of Pennsylvania is served by two educational unions, the National Education Association (NEA) and the American Federation of Teachers (AFT). Teachers and paraprofessionals hold membership in one of these unions based on the part of the state they reside in. Those in either the Pittsburgh or Philadelphia metropolitan area can join local affiliates of the AFT. Teachers and paraprofessionals in the rest of the state belong to the Pennsylvania State Education Association (PSEA).

The Philadelphia School District has approximately 10,000 teachers and 99% are members of the Philadelphia AFT (PA-AFT) (Personal Communication, 2008). There are approximately 6,000 paraprofessionals and 99% are members of the PA-AFT (Personal Communication, 2008). A random sample of the 16,000 teachers and paraprofessionals will be pulled using administrative databases, provided by the union. One-thousand and three-hundred participants will be selected at random without replacement (Table 1). The sample of participants will be distributed using proportional allocation, which will take into account the population size distribution. We expect an 80% response rate from the 1,300 mailed surveys, resulting in sample size of 1,040. From those returned surveys, we further expect that 3% (n=40) will have incomplete data. Taking into account this response rate and the fact a small proportion of responses may have missing data, we are anticipating a final sample size of 1,000 persons from the Philadelphia AFT (Table 1).

The Pittsburgh school district is represented by the Pittsburgh AFT (PFT). Approximately 95% of classroom teachers and 90% of paraprofessionals are union members (Personal Communication, 2009). There are approximately 2,700 classroom teachers and 750 paraprofessionals active in the PFT union (Personal Communication, 2009). A random sample of the 2,700 teachers and paraprofessionals will be pulled using administrative databases, provided by the union. Six-hundred and fifty participants will be selected at random without replacement (Table 1). The sample of participants will be distributed using proportional allocation, which will take into account the population size distribution. We expect an 80% response rate from the 650 mailed surveys, resulting in sample size of 520. From those returned surveys, we further expect that 3% (n=20) will have incomplete data. Taking into account this response rate and the fact a small proportion of responses may have missing data, we are anticipating a final sample size of 500 persons from the Pittsburgh AFT.

Teachers and paraprofessionals in Pennsylvania school districts outside of Pittsburgh and Philadelphia are represented by the Pennsylvania State Education Association (PSEA). Approximately 85% of classroom teachers and 65% of paraprofessionals are union members (Personal Communication, 2009). As of early 2009, the PSEA's database has approximately 120,000 teachers and 35,000 paraprofessionals (Personal Communication, 2009). A random

sample, stratified by urban-rural status (urban, rural, and suburban) will be pulled using administrative databases, provided by the union. Approximately 1,500 participants will be pulled from each of three separate strata: urban, rural, and suburban, resulting in a total sample size of 4,500 (Table 1). This sample is different than the sample pulled from the Pittsburgh and Philadelphia school districts for two reasons. Firstly, while the samples from the Pittsburgh and Philadelphia districts are located in one urban location, the sample from the PSEA will come from a variety of urban, rural, and suburban locations. The sample of participants will be distributed using proportional allocation, which will take into account the population size distribution. We expect an 80% response rate from the 4,500 mailed surveys, resulting in sample size of 3,600. From those returned surveys, we further expect that 10% (n=100) will have incomplete data. Taking into account this response rate and the fact a small proportion of responses may have missing data, we are anticipating a final sample size of 3,500 persons from the rest of the state.

There will be 6,450 surveys mailed to potential participants in Pittsburgh, Philadelphia, and the rest of Pennsylvania. We anticipate an 80% response rate based on prior paper-pencil surveys of teachers in this state, resulting in approximately 5,000 returned surveys (Table 1).

Table 1. Sample Size Estimates

Sampling Unit	Total Universe	Pulled Sample	Sample after 80% Response Rate	Sample After Removing Missing Data
Philadelphia	16,000	1,300	1,040	1,000
Pittsburgh	3,450	650	520	500
Rest of State				
Urban	45,000	1,500	1,200	1,166
Rural	45,000	1,500	1,200	1,167
Suburban	45,000	1,500	1,200	1,167
TOTAL	154,450	6,450	5,160	5,000

Participation in the study will be offered to all individuals given that correct contact information is available. However, this participation will be voluntary with no incentives for participation beyond self-motivation and no negative impacts for declining to participate. Unions have agreed to use a variety of methods to help increase participation rates of their respective union members. This could include flyers, posters, and ads in local union papers and print documents.

The survey packet, to be mailed to the home address of the potential participant, includes a description of the study and the measures that will be taken to protect their confidentiality should they choose to participate in the study (Appendix C). This text has been written to provide potential respondents with the information required in an informed consent form, but we are requesting a waiver of written informed consent. A returned survey will be deemed to be the subject's consent to participate. This research project has passed CDC's Human Subjects Review Board protocol.

The letter will emphasize that it is the choice of the participant to participate and the confidentiality protections that will apply should they choose to participate. The letter confirms their willingness to participate by completing and retuning the enclosed survey. A returned survey will be deemed to be the subject's consent to participate.

B.2   Procedures for the Collection of Information

Teachers and paraprofessionals will be identified from the Pennsylvania State Education Association (PSEA), Philadelphia AFT (PA-AFT), and the Pittsburgh AFT (PFT) union member databases. Every identified potential participant will be contacted three times and given every opportunity to complete the questionnaire. The sample will be divided into three primary sampling units (PSU's). The first primary sampling unit will be the Philadelphia School District. A random sample of 1,300 participants will be drawn from this PSU. Union members will be considered eligible if they are listed in the current union database. The second PSU unit will be the Pittsburgh School District. A random sample of 650 participants will be drawn from this PSU. The third PSU will be all other school districts in Pennsylvania covered by the PSEA. This third PSU will be further stratified into districts that reside in one of three strata: urban, suburban, or rural. A random sample of size 1,500 will be drawn from each of the three strata for this PSU. Assuming an 80% response rate, and taking into account the fact a small proportion of responses may have data missing for some items on the survey, we are anticipating actual sample sizes to be approximately 1,000 for PSU 1, 500 for PSU 2, and 3,500 for PSU 3 with approximately equal samples sizes within each strata for PSU 3, for a total sample size of approximately 5,000 completed surveys.

It is acknowledged that our study population contains only those teachers and paraprofessionals enrolled in an educational union and is therefore not equally representative of all teachers and paraprofessionals in Pennsylvania. However, at the present time, these databases offer the best potential for capturing such a diverse worker population (teachers, blue-collar paraprofessionals, and white-collar paraprofessionals). Since approximately 90% of teachers and 65% of paraprofessionals are enrolled in an education union, we feel that this population bias will be minimized. Given the general scarcity of etiologic research on the topic of workplace violence in the education field, this proposed project will be an important step in providing a framework for future research.

Each union involved with this research project has their own unique database and the survey methodology will be dependent on their specific database's capabilities. Each union will independently develop a database file for this research project. Each union will randomly assign each member in their database a unique ID. The unions will then strip the database files of the name and address, leaving only the union-assigned unique ID and other pertinent socio-demographic information (gender, occupation). These files will be sent to DSR via Fed Ex on a pass-word protected CD or transmitted via a secure FTP site.

Using these stripped database files; DSR staff will pull a random sample of eligible participants, from each union. Participants will be selected at random without replacement. We will have SAS, or another statistical software system, generate a random number list. We will then take

each database file and based on the row number in the access or excel file, pull the sample based on the random number list. This same process will be replicated for each for the three unions. The sample of participants will be distributed using proportional allocation, which will take into account the population size distribution. DSR staff will then assign their own unique ID to these potential participants.

DSR will send back to the unions the revised database file of the selected participants along with the union's original unique ID and DSR's assigned unique ID. The unions will be able to access this cross-walk between their unique ID, DSR's unique ID, and the participants personal contact information, but neither DSR nor the survey contractor will have access to this. From here, each union will independently prepare mailing labels for their respective participants.

At this time, DSR staff will also send to the survey contractor a list of the DSR-assigned unique ID's broken down by the three unions (PSEA, PA-AFT, and PFT). The survey contractor will develop, print, and assemble the initial questionnaire packet. This will include the paper-and-pencil questionnaire embedded with the DSR unique ID, a cover letter, and a self-addressed and stamped return envelope. This packet will be sealed in a flat envelope with the DSR unique ID embedded on the paper envelope. The appropriate questionnaire packets will be mailed via Fed Ex to the union offices located in Pittsburgh, Harrisburg, and Philadelphia.

DSR staff will travel to the Pittsburgh AFT office and manually affix home address labels to the sealed questionnaire packets. This work is to be done only at the PFT office in Pittsburgh, under union leadership. Once complete, the PI will perform a final quality control measure to verify that the name and address label of the potential participant corresponds to the NIOSH unique ID embedded in the mailing envelope. At this time, only the PI will be given access to the union member's contact information. This is strictly to verify that the questionnaire packet is going to the correctly assigned DSR unique ID. After this quality control measure is complete, the PFT questionnaire packets will be mailed out.

DSR staff will travel to Philadelphia and duplicate the methods used to prepare the PFT mailings including affixing name and address labels onto corresponding mailing envelopes. This work will take place in the PA-AFT office located in Philadelphia. A final quality control measure will be performed by the PI. While DSR staff members are preparing the Pittsburgh and Philadelphia AFT questionnaire packets, the PSEA will perform these same steps with PSEA staff dedicated to this research project. Dr. Dan Mercer, contact person at PSEA, will perform similar quality control checks before the final mailing of the PSEA questionnaire packets.

The survey contractor will coordinate, record, and track all incoming questionnaires and correspond with DSR, PSEA, PFT, and PA-AFT as needed. Approximately two weeks after the initial mailing, the survey contractor will send reminder postcards to all three unions. The unions will then send these reminder postcards to all eligible participants, regardless of whether they have returned the questionnaire or not. If needed, DSR staff will travel to Pittsburgh and Philadelphia to assist in this task, though it is presumed that the unions will perform this task with union staff members.

Approximately four weeks after the initial mailing, the survey contractor will e-mail the union contacts (Dr. Dan Mercer, Dee Philips, and John Tarka) a list of the non-responders using the DSR unique IDs. The unions will prepare mailing labels for these non-responders. DSR staff will then repeat the steps listed above, traveling first to Pittsburgh and then to Philadelphia. PSEA will perform these steps with PSEA staff as before.

This survey methodology will allow DSR to randomly select potential participants and follow-back with non-responders without keeping a record or database of union's member's names and home addresses. The PI and DSR staff will see the name and address of participants as labels are affixed to questionnaire packet envelopes; however, this task will be performed at the union offices; therefore NIOSH will have no written or otherwise recorded list of union members.

The survey software used by most survey and research contractors is Teleform by Cardiff software. Forms can be designed in its design application and once these forms are completed, both handwritten information and tick box responses can be scanned into the software system. From here, forms can then be read, evaluated, verified and exported to end databases such as Access, Excel, and SAS.

The paper-and-pencil questionnaire will be converted into a Teleform design by the survey contract firm. As the questionnaires are returned to the contractor, they will be logged and scanned into Teleform. The scanning process is usually operated by a survey professional that performs quality control steps along the way. Upon completion of the data collection phase, the contractor will provide DSR staff with a cleaned data file in either an Excel or SAS database. This data will be sent to DSR via Fed Ex on a pass-word protected CD.

Once received by DSR, the data will be stored on a password-protected computer and in secure files in locked DSR offices. Because of the inherent cost of these data and their intrinsic value to researchers, upon completion of the intended research, the data will be maintained as "active" files for a period of up to five years. Subsequently, the data will be maintained as archived protected data files for a period of up to 15 years. Final disposition of the data will be handled in accordance with federal recordkeeping requirements. All NIOSH (DSR) staff who handle this data will be required to keep their confidentiality training up-to-date.

We will describe and quantify the physical, non-physical, and electronic workplace violence incidents in teachers and paraprofessionals. Descriptive, univariate, and multivariate analyses will be conducted using the survey procedures available in the Statistical Analysis System version 9.1 (SAS Institute, Cary, NC). The survey procedures in SAS allow for the complex design we are employing to be accounted for PSU's, strata, and weights. Weights will be calculated as the inverse of the probability of selection in the study. Data to calculate the weights for the population under study will be obtained from the Common Core of Data (NCEHS, 2009). Further adjustments to the weights will be made to account for non-response by strata and PSU. If a response bias is found by specific demographics of the study population, further adjustments to the weights to account for this non-response bias will also be made. All statistical tests will be two-sided and the Type I error will be set at 5%.

Prevalence proportions (rates) will be calculated for physical, non-physical, and electronic workplace violence (WPV) incidents. Rates of WPV event frequencies will be calculated using Proc SURVEYFREQ. Percentages of WPV with corresponding standard errors will be calculated by major demographic categories such as occupation (teacher/paraprofessional), gender, years of experience (5-year categories), status (full-time/part-time), rural-urban status (urban, suburban, and rural), type of school (elementary, middle, or high school), and school size (small, medium, and large). Teachers and instructional aides for special education populations will be analyzed separately. Recall bias by demographics for the 12 month recall period will be assessed and adjusted with the methodology developed by Landen and Hendricks (1995).

Risk factors for WPV will be modeled through Proc SURVEYLOGISTIC. Univariate odds ratios with corresponding confidence intervals for each potential risk factor will be calculated. A final logistic model, controlling for all potential significant confounding factors, will be fit to obtain multivariate adjusted odds ratios for specific risk factors.

We will evaluate the effects of physical, non-physical, and electronic WPV on job satisfaction and quality of life with Proc SURVEYREG. For the primary analysis, the main exposure of interest will be WPV; physical, non-physical, or electronic. Two separate outcomes will be examined for this specific aim; job satisfaction and quality of life.

Since we are interested in comparing the prevalence of WPV between our PSU's and strata, precision estimates for each PSU and strata are presented in Tables 1 and 2 based on previous estimates of physical and non-physical WPV rates in educators (Gerberich et al, 2007). These precision estimates are derived from the formula $\pm 1.96[p(1-p)/n]^{1/2}$, where p is the proportion being estimated and n is the respective sample size.

Table 2. Precision Estimates for each PSU in the Sample Based on a Non-Physical WPV Rate of 39 Incidents per 100 Persons per Year

PSU	Sample Size	Precision
Philadelphia	1,000	±3%
Pittsburgh	500	±4%
Rest of Pennsylvania	3,500	±2%
Stratum urban, suburban, rural	1,166	±3%

Table 3. Precision Estimates for each PSU in the Sample Based on a Physical WPV Rate of 8 Incidents per 100 Persons per Year

PSU	Sample Size	Precision
Philadelphia	1,000	±2%
Pittsburgh	500	±2%
Rest of Pennsylvania	3,500	±1%
Stratum urban, suburban, rural	1,166	±2%

Power estimates for evaluating risk factors to WPV to detect a rate ratio of at least 1.5 by varying percentages of the levels of the risk factor present in the population are presented in Table 3. These calculations assume that the level of Type I error is 0.05, the alternative hypotheses are two-sided, and the total sample size is 5000. The calculations were performed with the EPI-Info software package and assume a simple randomly selected sample.

Table 4. Power Estimates to Detect a 1.5 Rate Ratio by Varying Percentages or Risk Factors Present in the Population

	Rate of 39 per 1,000	Rate of 8 per 100
% Risk Factor	Power	Power
10%	100%	79%
20%	100%	94%
50%	100%	98%
80%	100%	86%
90%	99%	54%

All power calculations performed here assume a simple random sample and do not account for the finite population correction factor. Due to the complex design of the sample, some power will be lost when accounting for the survey design in the analysis of the data compared to what is presented here. However, adjustments for the finite population we are studying in the analysis will improve the power compared to what is presented here. Both of these factors in the analysis should be minor in terms of the power of the study and will generally offset each other.

B.3   Methods to Maximize Response Rates and Deal with Non response

Based on prior survey research by our stakeholders, we estimate that we will achieve an 80% response rate using our proposed paper-and-pencil questionnaire. PSEA routinely conducts survey research with their union members and has achieved response rates between 50% and 80%. The University of Minnesota achieved an 84% response using the questionnaire we have modified for use in this study. Given this background and the positive reaction we have received from union leaders, union members, and state leadership, we are anticipating an 80% response rate. We plan on taking the following steps to encourage participation in this research study and assure achieving this response rate:

1. The state-based unions will inform their constituents of this study via documents, newsletters, and meetings. This information dissemination effort may help to influence potential respondents to participate.

2. Stakeholders involved with this research project have indicated that this topic is of strong interest to educational unions. They believe that K-12 teachers will find this line of research to be salient and will be inclined to relay their personal experiences with workplace violence in hopes of developing interventions or change current policies and practices.

3. Participants will receive at least two copies of the survey packet, in addition to a reminder postcard. The mailing procedures for this research project will follow the Dillman Total Design Method as closely as financially possible. This method can achieve high response rates from historically difficult subjects (Hoddinott & Bass, 1986). This methodology includes the mailing of a pre-survey letter, specifically constructed cover letter, several carefully timed follow-up mailings (Dillman, 1978).

4. The questionnaire has been designed to be as non-burdensome as possible. This includes ordering the questions in a logical sequence and asking only those questions that are needed for analysis purposes.

The possibility for non-response bias is present in any study using self-report survey data. Ultimately, we will compare the socio-demographic information we have on respondents and non-respondents using union databases. This will provide insight on any potential response bias.

Finally, it is acknowledged that this study population contains only those teachers and paraprofessionals enrolled in a state-based educational union. Thus, the population surveyed is not equally representative of all teachers and paraprofessionals in Pennsylvania; however there is no known database that retains contact information on both teachers and paraprofessionals in the state of Pennsylvania. Since a very high percentage of teachers and paraprofessionals are enrolled in an education union, we feel that this population bias will be minimized. Also, we will use the National Center for Education Statistics' Common Core of Data (CCD) to compare the socio-demographics of our sample with the socio-demographics for the state as a whole (http://www.nces.ed.gov/ccd/). This population bias will be acknowledged in all presentations and publications. Given the general scarcity of etiologic research on the topic of WPV in the education field, this proposed project will be an important step in providing a framework for future research.

B.4   Test of Procedures or Methods to be Undertaken

To measure potential risk factors for workplace violence in teachers and paraprofessionals, we will use a modified version of survey questionnaire developed by the University of Minnesota's Center for Violence Prevention and Control's for the "Minnesota Educators Study" (Gerberich et al, 2007). The original survey tool was developed and pilot-tested at the University of Minnesota, under the direction of Dr. Susan Goodwin Gerberich (2007). This survey was used to measure specific risk factors for work-related physical violence in a cohort of 6,469 K-12 licensed teachers in the state of Minnesota (2007). We have modified the tool for use in this study and population.

We added several questions on job satisfaction and overall quality of life to this questionnaire. These questions came from two sources: the CDC core, "Healthy Days Measures" and the NIOSH Quality of Work Life Questionnaire (CDC, 2005 and NIOSH, 2009). The CDC measures population health-related quality of life using four questions called the "Healthy Days Measures" (2005). These questions have been part of the state-based Behavioral Risk Factor

Surveillance System (BRFSS) sample since 1993 (2005). There is a wide variety of publications on the validity, reliability, and responsiveness of these questions (2005).

In 2000, NIOSH developed questions to assess the quality of work life for the 2002 General Social Survey (NIOSH, 2009). The General Social Survey is a biannual, nationally representative, personal interview survey of U.S. households conducted by the National Opinion Research Center (2009). NIOSH selected 76 questions that dealt with a wide assortment of work organization issues, such as hours of work, workload, worker autonomy, layoffs and job security, job satisfaction/stress, and worker well-being (2009). Data collected in 2002 on 2,765 persons is available for comparison purposes (2009). We selected eight questions from this national survey to use in our "Workplace Violence in Pennsylvania Teachers and Paraprofessionals" survey. Questionnaire numbers 58, 59, 60, 61, 62, 63, 64, and 65 on our survey correspond with item numbers 5.74, 5.75, 5.79, 5.70, 5.71, 5.72, 5.73, and 5.64, respectively on the national survey.

Since no tool or set of questions on electronic WPV existed, experts at the CDC's Division of School and Adolescent Health (DASH) were asked to help develop these. DASH performed their own review process on these additional questions and they were added to the original questionnaire. DASH performed pilot testing for comprehension of these new questions with members of their own staff.

We provided our stakeholders, union partners, state education leaders, and district personnel a copy of this modified questionnaire. We then further refined the questionnaire based on their and responses, suggestions, and feedback. Finally, the questionnaire was pilot tested for length and comprehension with five NIOSH (DSR) employees and further modifications were made. Consequently, we scrutinized each question individually to confirm that it would provide useful information for analyses in terms of content.

B.5   Individuals Consulted on Statistical Aspects and Individuals Collecting and/or Analyzing Data

The following individuals will be involved in the design, collection and analysis of the data obtained in this study:

Hope M. Tiesman, M.S.P.H, Ph.D. - Project Officer/Epidemiologist, Analysis and Field Evaluations Branch, Division of Safety Research, NIOSH, Morgantown WV, 304-285-6067, fto9@cdc.gov will be involved in the design, collection and analysis of data.

Harlan Amandus, Ph.D. – Branch Chief, Analysis and Field Evaluations Branch, Division of Safety Research, NIOSH, Morgantown WV, 304-285-5913, hea1@cdc.gov will be involved in the design, collection and analysis of data.

Dan Hartley, Ed.D. – Epidemiologist, Analysis and Field Evaluations Branch, Division of Safety Research, NIOSH, Morgantown WV, 304-285-5812, dsh3@cdc.gov will be involved in the design of the data collection.

Scott Hendricks, M.P.H. – Statistician, Analysis and Field Evaluations Branch, Division of Safety Research, NIOSH, Morgantown WV, 304-285-6000, sah5@cdc.gov will be involved in the design and analysis of data.

Lunette Utter – Data Collection Specialist, Analysis and Field Evaluations Branch, Division of Safety Research, NIOSH, Morgantown WV, 304-285-6001, lku1@cdc.gov will be involved in the collection of the data.