

**The author(s) shown below used Federal funds provided by the U.S. Department of Justice and prepared the following final report:**

**Document Title:           Protecting Children Online: Using Research-Based Algorithms to Prioritize Law Enforcement Internet Investigations, Technical Report**

**Author(s):                   R. Gregg Dwyer, M.D., Ed.D., Michael Seto, Ph.D., Dana DeHart, Ph.D., Elizabeth Letourneau, Ph.D., Tracy McKee, M.S., Robert Moran, Ph.D.**

**Document No.:           250154**

**Date Received:           August 2016**

**Award Number:           2011-MC-CX-0002**

**This report has not been published by the U.S. Department of Justice. To provide better customer service, NCJRS has made this federally funded grant report available electronically.**

**Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.**

**Protecting Children Online:  
Using Research-Based Algorithms  
to Prioritize Law Enforcement Internet Investigations**

**Technical Report**

Prepared By:

R. Gregg Dwyer, M.D., Ed.D.

Michael Seto, Ph.D.

Dana DeHart, Ph.D.

Elizabeth Letourneau, Ph.D.

Contributors:

Tracy McKee, M.S.

Robert Moran, Ph.D.

Submitted to the Office of Juvenile Justice and Delinquency Prevention  
May 2016



Community and Public Safety Psychiatry Division  
Department of Psychiatry and Behavioral Sciences  
College of Medicine  
Medical University of South Carolina  
29-C Leinbach Drive  
Charleston, SC 29407

**Research Team**

**Principal Investigator**

R. Gregg Dwyer, M.D., Ed.D.

**Co-Principal Investigators**

Michael Seto, Ph.D.

Elizabeth Letourneau, Ph.D.

**Co-Investigators**

Dana DeHart, Ph.D.

Robert Moran, Ph.D.

Kevin Armstrong, B.S.

**Consultant**

Tracy McKee, M.S.

**Research Assistants**

Anna Culpepper

Kathleen Whitehouse

### **Acknowledgements**

Special thanks are extended to Lyle Browne for his extensive work assisting with coding of qualitative data. Thanks also to our partners on state and local Internet Crimes Against Children (ICAC) Task Forces who provided data to be utilized for this study. Although multiple Task Forces participated by providing data, the South Carolina ICAC Task Force also served as the recipient and clearinghouse for data from all the other Task Forces for which we thank them.

### **Disclaimer**

This project was supported by Grant No. 2011-MC-CX-0002 awarded by the Office of Juvenile Justice and Delinquency Prevention, Office of Justice Programs, U.S. Department of Justice with support of the South Carolina Office of the Attorney General. Points of view in this document are those of the authors and do not necessarily represent the official positions or policies of the U.S. Department of Justice or the South Carolina Office of the Attorney General.

### **Abstract**

There is increasing public and professional concern about Internet sexual offending, as reflected in increasing law enforcement cases and clinical referrals. While all instances of Internet offending against minors require intervention, the number of cases and the overarching goal of protecting children require law enforcement to prioritize cases. This project used data from 20 Internet Crimes Against Children task forces across the United States -- offender characteristics, crime characteristics, and online behavior -- to develop empirically-based recommendations to assist law enforcement in prioritizing: (1) cases involving production of child pornography over possession/distribution; (2) cases involving online luring for the purpose of meeting the minor to commit sexual offenses, over luring restricted to online behavior such as sexual chat or exchanging pornographic images; and (3) cases involving offenders who have committed contact sexual offenses against children over cases involving offenders with no known history. The research builds on a previous OJJDP-funded project by increasing the number of task forces and thereby sample size in order to develop practical recommendations, and adding a new component – a geographical analysis of cases – that will assist decision makers in the allocation of training and resources across the United States in order to combat online sexual exploitation and abuse of children.

## Background

### Statement of the Problem

There is increasing public and professional concern about Internet sexual offending involving children, as reflected in increasing law enforcement cases and clinical referrals for possession, distribution, and production of child pornography, online sexual solicitations of children, and conspiracy to commit these kinds of crimes (Motivans & Kyckelhan, 2007; United States Sentencing Commission, 2012). The National Child Exploitation Threat Assessment (U.S. Department of Justice, 2010) reported an overall increase in cases of child sexual exploitation in the U.S., with a specific increase by 230% in the number of complaints of online enticement of children from 2004 to 2008. Although changes in complaint processing by Internet Crimes Against Children Task Forces may account for some of this increase, there were 20,562 complaints reported during those four years. Mitchell and colleagues (2014) reported that approximately 1 in 11 youth experienced an unwanted sexual solicitation in the past year, and approximately 1 in 25 experienced an aggressive sexual solicitation. For some, online solicitation includes both online and offline exploitation or abuse.

The number of online offending cases exceeds law enforcement resources (e.g., Canwest News Service, 2009; Long et al, 2016; Seto, 2014). Faced with more cases than personnel or time to work them, law enforcement personnel must triage and prioritize their resources. But how should they do so? All Internet sex crimes involving children are serious and require attention. Given the importance of pursuing those who pose an immediate risk to children, one strategy is to work cases involving production of child pornography before those involving possession/distribution; cases of *contact-driven* solicitations over solicitations restricted to online behavior alone (*fantasy-driven*); and cases of Internet offenders known to have committed contact sexual offenses against children.

This study specifically addressed this resource challenge by systematically analyzing data from a sample of Internet Crimes Against Children (ICAC) Task Force cases representing a cross section of the United States, representing all geographic regions. This study expanded on a prior Office of Juvenile Justice and Delinquency Prevention (OJJDP) grant funded data collection of ICAC cases assembled to develop a typology of Internet offenders against children. In so doing this project identified variables that can be readily available to law enforcement officers at the field investigation stage and that have potential for use to distinguish each higher-priority group: (1) producers of child pornography; (2) contact-driven solicitations; (3) Internet offenders who have also committed contact sexual offenses.

There are numerous potential risk factors that might distinguish between these groups, but many are available only after a case is well underway or concluded. For example, it is likely that individuals under investigation who have been previously diagnosed with Antisocial Personality Disorder are more prone to produce child pornography, attempt contacts, and to directly offend against children, than their non-disordered counterparts, but such diagnostic information is typically unavailable until after arrest, when a clinical evaluation is conducted in preparation for a legal proceeding. Such after-the-arrest information is of little use to investigators who must triage cases early in the investigative process. In contrast, a criminal history check can be performed as soon as a suspect is positively identified and may reveal a prior offense related to the current investigation. In addition, the content and process of an offender's online activity with an actual minor or an undercover officer's persona is an integral part of the investigation and will be collected from the very beginning. These types of data might offer unique indicators of risk for harm and can be readily converted to a practical investigation priority guide (see Seto & Eke, 2015, on the development of an investigator-friendly risk

assessment tool for child pornography offenders, the Child Pornography Offender Risk Tool or “C-port”).

Foundational research conducted by this study’s team members to address these questions was in existence before the start of this study. This included research on the motivations of child pornography offenders (Seto et al., 2006), the explanations given by child pornography offenders for their crimes (Seto et al., 2010), the likelihood that Internet offenders have already committed contact offenses against children (Seto et al., 2011), the likelihood that Internet offenders will commit contact offenses in the future (Seto et al., 2011), and the factors that predict recidivism (Eke, Seto, & Williams, 2011; Seto & Eke, 2005, 2015).

### **Supporting Theoretical and Practical Background**

As reviewed in this section, Internet crimes against children differ along a variety of important offender characteristics, which may be relevant in developing strategies to identify and triage more serious cases. These characteristics pertain to offender motivation, types, histories, and risk factors.

*Offender motivations.* Many, but not all, Internet offenders are motivated by a sexual interest in children. Seto and colleagues have, for example, reported that the majority of child pornography offenders evidenced greater sexual arousal to children than to adults; indeed, the relative preference for child stimuli demonstrated by child pornography offenders exceeded that of contact offenders (Seto et al., 2006). In a separate study, 33-50% of child pornography offenders admitted they were sexually interested in children or in child pornography (Seto et al., 2010). This relationship is sufficiently robust to support the inclusion of child pornography use in the proposed revision of the criteria for a diagnosis of Pedophilia (see Seto, 2010; Seto, Stephens, Lalumière, & Cantor, 2016). Not all child pornography offenders have pedophilia, however; some offenders report indiscriminate sexual interests, an “addiction” to pornography, and



curiosity (Seto et al., 2010). The method for obtaining access to sexual material may also be a discriminating factor; Wolak et al. (2011) found that individuals who obtain child pornography through peer-to-peer software have larger, more explicit collections of abusive imagery, and may therefore pose a greater risk for future contact offending. Identifying indicators of child sexual interest among Internet offenders has practical value for prioritization of cases.

***Solicitation offender types.*** Research has identified differences between fantasy driven and contact-driven solicitation offenders. Briggs et al. (2011) have suggested that fantasy driven offenders engage in online activities such as sexual chat or exchange of pornographic images that is gratifying in and of itself, without a desire for physical contact with the victims. Contact-driven offenders, on the other hand, engage in online activities in order to arrange real-world meetings. Briggs et al. (2011) found that contact-driven offenders have fewer online sessions before attempting to meet victims. This suggests that early indicators that might distinguish contact-driven from fantasy-driven solicitation offenders to include the speed with which the offender attempts to arrange an in-person meeting and how much of the conversation is focused on discussing meeting offline versus online sexual interactions such as sending pornography, requesting sexually explicit images from the minor, or suggesting mutual masturbation via webcam.. This distinction between fantasy- and contact-driven solicitations has not yet been replicated.

***Contact offending history.*** In a recent meta-analysis, Seto et al. (2011) identified 21 samples of Internet offenders with information about their contact offending histories. On average, one in eight had an official criminal record for contact sexual offending. In the six samples with self-report data, 55% admitted to a history of contact sexual offending. Identifying the factors that distinguish those who have contact offenses from those who do not would advance our understanding of risk of initial sexual offending and recidivism. Advance warning of

an impending contact offense based on factors identifiable during online contact would be important for effective early responding and the protection of children. Online behavior suggestive of poor self-control, such as simultaneously contacting multiple minors or bringing up sexual topics very quickly during online interactions, may be germane (Seto & Hanson, 2011).

***Risk of Sexual Recidivism.*** Seto et al. (2011) also reviewed recidivism rates available from nine samples of Internet offenders, followed for an average of a little over three years, with a range from 1.5 to 6 years. Approximately 1 in 20 (4.6%) committed a new sexual offense of some kind (whether contact or noncontact) during follow-up. These recidivism rates belie the idea that all Internet offenders pose a high risk of committing contact offenses in the future. But it is clear that Internet offenders vary in risk; some pose a relatively high risk to reoffend sexually and some will attempt to perpetrate contact sexual offenses. Thus, an important task for law enforcement is to identify those higher-risk individuals for investigation prioritization.

Content analysis of letters by stalking offenders to victims has been useful in determining whether stalkers are likely to attempt contact (Dietz et al, 1991a; Dietz et al, 1991b). It follows that similar comparisons might be made between contact-driven and fantasy-driven solicitation offenders using parameters such as the number of sexually explicit words in the chat, the total number of chat exchanges, the amount of time taken to initiate sexually explicit chat, sexual words in the offender's screen name, the use of a webcam, and the quantity of offender falsehoods.

***Risk Factors.*** Research is beginning to emerge on the factors that predict recidivism among Internet sex offenders. These factors appear to be the same as factors identified in decades of research on conventional sex offenders. For example, classic criminological factors such as offender age at time of first arrest and prior criminal history predict recidivism, and failure on prior conditional release are related to risk of recidivism (Seto & Eke, 2005; Eke, Seto &

Williams, 2011). We also have shown that other intuitively sensible factors such as admission of sexual interest in children and the relative interest in boys versus girls, as reflected in child pornography and other child-related content, predict sexual reoffending (Seto & Eke, 2015). In particular, the ratio of boy to girl content predicts child pornography recidivism (Seto & Eke, 2015), just as contact offenders who victimize boys are more likely to reoffend than those who victimize girls (Seto, 2008). Online offenders who exhibit a sexual interest in boys (seeking child pornography content depicting boys or soliciting boys online) are therefore expected to be at greater risk of recidivism. Other investigators have reported that lower education, being single, possessing non-Internet child pornography (e.g., magazines), prior sexual offender treatment (likely a proxy for prior sexual offending history), and absence of depictions of adolescent minors within the child pornography cache predicted recidivism among child pornography offenders (Faust et al., 2009). It bears noting that our study, and the research cited on risk factors, pertain to adult offenders and not to adolescent offenders who, by definition will be young, will have younger victims, and will be single.

### **Prioritizing Cases for Law Enforcement**

Based on the foregoing literature review, several variables appeared to be good candidates as potential early indicators for law enforcement to prioritize adult Internet sex crime cases. These include:

***Offender demographic.*** Potentially relevant factors in this category include offender age, marital status, whether the offender resides with children, and whether the offender works in a child-focused occupation and/or volunteers in child-focused activities (e.g., teacher, sports coach, youth faith-based group);

***Criminal record.*** Potentially relevant factors in this category include the offender's prior criminal history, specific types of prior offenses (e.g., violent offenses), and whether there is any indicator of failure on conditional release (Seto & Eke, 2005; Eke et al., 2011);

***Online activity.*** Potentially relevant factors in this category include the ratio of boy to girl content and sexual interest in boys (Seto & Eke, 2015), the nature of online chat (i.e., fantasy- vs. contact-driven; Briggs et al., 2011) and the method for obtaining access to sexual material (e.g., peer-to-peer software versus commercial sites);

***Child pornography content.*** Variables include the number, age and gender of children depicted in the abusive images; the ratio of child images to other forms of pornography; the ratio of younger (prepubescent) to older children; and the ratio of boys to girls depicted in the images;

***Online time.*** Potentially relevant factors in this category include when online contact with victims was made (e.g., time of day and day of the week); frequency of online contact; and duration per session (frequency and duration may serve as proxy measures of an offender's drive or compulsiveness);

***Sites visited.*** Potentially relevant factors in this category include number of different sites used by the offender and whether the sites an offender has visited suggest interest in a particular age group or a particular theme (e.g., indicative of pedophilia and/or other paraphilic sexual interests); sexual or not);.

***Online name.*** Potentially relevant factors in this category include whether or not an offender's online name is sexually suggestive and whether an offender uses multiple aliases across different sites;

***Number of victim targets.*** Potentially relevant factors in this category include the number of child victims being simultaneously targeted. We recognize that multiple victims are more

likely to be identified when other law enforcement officers are posing as children or when a citizen report has been made;

***Chat content.*** Potentially relevant factors in this category include the speed with which sexual content is introduced or solicitation made; the speed with which an offender suggests a move to private setting; the introduction of questions about the victim's family or parent figures, the offender's request for victim's physical description, questions designed to elicit the victim's sexual knowledge or sexual experience; the offender's request for online sexual behaviors; the offender's request for/introduction of webcam activity;

***Use of other communication methods.*** Potentially relevant factors in this category include whether the offender utilizes multiple methods for contacting a victim, including email, texting, and telephone.

## **Study Goals & Program Objectives**

### **Goals**

OJJDP identified the importance of predicting which subjects “of an online child exploitation investigation” pose the greatest “risk to harm children” (USDOJ, 2011). As a result of the increasing number of cases, “law enforcement officers need tools to identify which suspects and cases should be given priority in order to identify and rescue victims” (USDOJ, 2011). This study addressed these OJJDP priorities using data from ICAC task forces in the US (the risk assessment tool by Seto & Eke, 2015, was developed in Canada in parallel to the conduct of this project). The overall goal of this study was to develop methods for law enforcement personnel to use in real time, during investigations. The project built on the OJJDP funded research of Dwyer, DeHart, Moran (the principal investigator and co-investigators respectively of this proposal) and others by adding task forces and thus sample size. Geographical mapping of cases was conducted to identify if there is a greater prevalence of

certain types of online offending across the country, to facilitate the allocation of law enforcement resources and investigative efforts and to determine if there is meaningful geographic variation attributable to demographics, state laws/policies, and other broad factors.

The researchers designed the study to be conducted in partnership with ICAC Task Forces, obtaining data on closed cases, across a 3-year span of time, from a minimum of 50 ICAC Task Forces (representing each U.S. state). At the time of the proposal, more than one-third of the available ICAC Task Forces were already collaborating with Dwyer on a then in-progress project and they planned to continue with this current project. Recruitment of the remaining Tasks Forces was done with a final total of 20 ICAC Task Forces participating over a 4-year (extended from the original proposal) study period: Alaska, Arkansas, Georgia, Los Angeles Police Department, Maine, Maryland, Nebraska, New Mexico, Oklahoma, Oregon, Pennsylvania, Polk County (FL), Sacramento, San Diego, San Jose, Seattle, South Carolina, Texas, Utah, and Wisconsin. The distribution of sites provides for a cross section of the United States with different geographic regions (Southeast, Northwest, etc) represented.

By partnering with ICAC Task Forces to obtain data on hundreds of closed ICAC cases, the goal was for the study to be able to identify early and broadly available indicators of Internet offender risk and identify criteria for law enforcement to prioritize three types of investigations: (1) those involving production of child pornography over cases involving possession alone; (2) cases of contact-driven solicitations over fantasy-driven solicitations; and (3) Internet offenders who have already or are currently directly victimizing children over those who have no history of contact offending. This study addressed the OJJDP Child Protection Research Program goals of (a) identifying factors that reliably indicate whether subjects of online/Internet offense investigations pose great risk of harm to children, (b) prioritizing law enforcement resources for higher risk cases, and (c) encouraging research-law enforcement partnerships.

Data collection from multiple sites and use of both quantitative and qualitative analyses was required to create further understanding and hence practical applications. The specific study aims were: (1) Primary applied aim: Develop an algorithm(s) combining the variables that most accurately distinguish higher priority cases. (2) Primary scientific aim: Expand the empirical foundation of a behavioral typology of online offenders. (3) Ancillary scientific aim: Obtaining a broad geographic representation of ICACs would support evaluation of regional difference in prevalence of higher risk offenders across the three case types. Such regional differences have been found in related areas of research, such as online pornography use (Edelman, 2009). Geographical variations would be expected because of state differences in sociodemographic factors (e.g., access to broadband internet), state online offending laws and policies, and geographic variations in sexual offending related resources such as the availability of specialists to assist those who are concerned about their sexual interests and/or online sexual activities (McGrath et al, 2009; <http://bit.ly/fkLYTb>). (4) Ancillary applied aim: The final aim was to develop and implement a training program. This was to be in conjunction with our ICAC partners and with a commitment from the National Center for Missing and Exploited Children to work with the study team. The intent was to develop a field-based training curriculum for deploying the empirically based algorithms for prioritizing higher risk cases with the trainings developed specifically for ICAC and other law enforcement personnel who investigate online sex offender cases. As a result of the limited depth of data across the task forces who supplied case information, it was determined that there was insufficient material to warrant a training course, so instead specific recommendations are provided here for ICAC and other investigators and for those responsible for prevention initiatives.

### **Program objectives**

The goals were addressed by completing the following objectives: (1) Collection of ICAC crime data: offender demographics and criminal history; victim (victim persona) demographics including age, sex, relationship to offender and education; offending behavior, initial charge(s) and result of adjudication. (2) Inferential statistical analysis for significant associations among offender, victim and crime details to identify types of offenders. (3) Collection of Internet chat texts between offenders and child victims or undercover officers portraying minors. (4) Content analysis of chat text, using qualitative methods, to identify patterns associated with higher priority cases. (5) Description of a sample of female offender cases. The focus was on male ICAC cases, as past research has shown that most (typically 99%+) Internet offenders are male and most of what is known about pedophilic disorder and sexual offending is based on studies of men (Seto, 2008). However, it was planned that if a sufficient sample of female Internet offenders could be obtained by combining all participating ICAC task forces, then descriptive and qualitative analysis of female Internet offenders would be reported. This will be among the first of its kind and help determine how and to what extent responses to this population need to differ.

### **Performance measures**

Each of the proposed study aims was linked to specific objectives that could be evaluated via discernible, measurable deliverables subject to ongoing progress monitoring. Progress towards obtaining, organizing, classifying and analyzing data (Goals 1-5) was summarized within OJJDP bi-annual interim reports and this final report. The investigators will be presenting findings at professional conferences and will also develop peer-reviewed publications describing the discrimination of higher versus lower priority cases. Manuscripts will be submitted to refereed professional journals for publication. Policy recommendations and supporting research is presented herein for lay audiences and will be broadly disseminated to local, state and national



stakeholders representing ICAC and other law enforcement agencies and national victims' rights groups via OJJDP.

The original plan was for training to be provided at national conferences, offered as a component at Department of Justice National Advocacy Center (NAC) training courses for investigation of Internet crimes against children, and as a stand-alone course conducted on each coast, the New England region and the Midwest to reach the ICAC task force personnel not attending conferences or NAC trainings. As noted, there was not enough material to warrant courses, but rather only practical recommendations, as set forth in this report.

## **Methods**

### **Design and Implementation**

The study design involved a retrospective record review of closed ICAC cases. Data were extracted, entered into a secure database and analyzed using commercially available software. Each step was conducted and supervised by qualified research personnel. Protocols were in place to ensure confidentiality and other legal and ethical requirements were met. The database was stored on a secure server at the Medical University of South Carolina (MUSC). Access to the server was password protected and limited to necessary study personnel. All methods were approved by the PI's institutional review board at MUSC.

The study leveraged an ongoing project that included a law enforcement partnership with more than one third of the available ICACs. All data were collected from existing records that were created for non-research purposes. ICAC personnel conducted de-identification of cases files after geocoding and before providing data to the researchers. The variables sought were those typically found in law enforcement incident reports; investigative reports; arrest warrants; indictments; victim interviews and statements; subject interrogation reports, transcripts or statements; forensic examinations of computer media; Internet chat, email and text exchanges

(printed copies). The principal investigator had collected these variables for the majority of subjects in a single state (South Carolina) pilot study, demonstrating the feasibility of collecting this depth and breadth of data provided the other ICAC sites collected and maintained similar files. This level of data collection continued during this study for that site and approached the same level for another site, but unfortunately the majority of participating ICAC sites provided data with significantly less depth and breadth. Given data were collected from existing file systems, the research team could only process what was provided.

Data were extracted from available criminal justice system reports, which varied in content and format across task forces. Some ICACs has electronic records, for example, whereas data extract from other ICACs required reading police incident and investigative reports, court and probation documents and forensic evaluation results. As such the research team conducted training for the database entry personnel to ensure understanding of how to extract, code and enter data into the database. Given some variability in terminology across collected data, the term “offender” is used in this report to refer to all study subjects regardless of final adjudication status.

### **Quantitative Analysis**

The original design to address the key research questions was to use a series of nested discriminant function analyses to identify key variables (offender characteristics, criminal history, online activity; when applicable, child pornography content and/or victim characteristics) that distinguish the critical group of cases from other relevant cases: (1) production of child pornography over cases involving possession alone; (2) cases of contact driven solicitations over fantasy-driven solicitations; and (3) Internet offenders who have already or are currently directly victimizing children over those who have no known history of contact offending. Significant discriminative variables from each set of variables would then be combined in an omnibus

discriminant function analysis, to determine overall classification accuracy. This analytic strategy has been successfully used in the development of sex offender risk measures (Quinsey et al., 2006) and other decision-making algorithms. However, these planned analyses were not possible as a result of significant problems with missing data; multivariate analyses such as discriminant function analysis require relatively complete data from large samples in order to produce robust results. Because these analyses were not possible, we instead conducted and report simpler group comparisons, distinguishing child pornography and solicitation offenders; fantasy- and contact-driven solicitation offenders (based on whether a meeting was attempted or completed); child pornography with offenders who committed both child pornography and solicitation offenses; and solicitation offenders with contact plus online offenders.

A relational database was designed and tested extensively prior to any data being stored from the study. The core of a good relational model is a process called normalization (Schwartz et al, 2012). Normalization ensures optimal performance with minimal resources, ensures that data is easily accessible for reporting, and removes duplication, which can lead to conflicting data.

The database consists of 16 tables and a customized data entry form for each table. Each form was customized to restrict data entry to predetermined choices. For instance, Gender was restricted to Female and Male choices only. Careful thought concerning data typing was used to improve the database design's efficiency while enforcing integrity. The primary purpose of data typing is to choose a data type that will give all the versatility needed and no more. On the one hand, if you use a 4-byte integer to store a value that you know will never get beyond 100, you are wasting 75% of your space. The cost of these redesigns is often problematic, so careful consideration was given at the beginning.

Data types also enforce integrity. By choosing an integer type, you ensure that letters will never be placed in this field. By choosing a date type, you will never put an invalid date in this field. All tables were indexed on the unique identity used for links between tables. This allowed for quick queries and reports. Indexing is the cheapest and easiest way to improve database performance. Indexing allows the database to perform well now and into the future. This design protected from errors and minimized the costs related to mistakes.

The database was documented with a detailed data dictionary for reference by the investigators and other approved researchers. Backup was automatically performed each night, with copies maintained locally. Each week offsite (MUSC) backups were performed so that in the event the storage site was compromised, the study could continue. Recovery procedures included the ability for access up to 30 days after individual backups, if needed. The database was password protected and stored on a secure university maintained server. Security threats to the databases are minimized using a good security model and MUSC has a very secure overall model given the need to maintain personally identifiable patient information. All data removed from the database for analysis was thoroughly cleaned prior to release by the database manager. Analyses were performed with SAS 9.4 or SPSS 23, with descriptive and analytical statistics detailed and summarized.

### **Geographic Analysis**

An ancillary aim of examining geographic differences was included, given that local laws and policies can influence offender behavior as well as identified cases. Also, the distance an offender is willing to travel to meet a victim might be a fruitful indicator of risk (e.g., offenders willing to cross several state lines to meet a victim might pose a greater risk to children than offenders who remain in-state for their offending behaviors).

***Geocoding the offender data.*** Raw data for GIS analysis was collected in a MS Access database containing only the offense code, address of the offender, and the agency that supplied the data. The data were exported to a format required by the US Census Geocoder tool and then geocoded. There were 1,615 records and 1,593 distinct addresses first run through the US Census Geocoder. Of those, 1,410 records (87.4%) of the records were successfully geolocated using the US Census Geocoder. An additional 179 (11.1%) records were geolocated using Google's geocoder and 26 (1.6%) of the records did not have enough information to successfully geocode.

***Preparing the US Census data.*** 2010 US Census block files and summary files for the 36 states represented by the offender database were downloaded from the US Census FTP site. A MS Access database was created for each state containing a query template to export the fields desired for this study. The summary files were extracted and loaded into the appropriate database. Using the query template, comma delimited files were then exported from each database.

***Geoprocessing.*** State by state, the comma delimited file containing the US Census summary file data needed for the study was loaded into GIS and joined to the corresponding block file based on the US Census defined GEOID field. The Select By Location tool was then used to select blocks in the newly joined block file using the point file of geocoded offender addresses. Selected blocks were copied into a GIS polygon file that would eventually contain the 1,542 census blocks represented by the 1,587 geocoded offender addresses. The geocoded point file of the offender addresses was then spatially joined with the GIS polygon file containing all of the census blocks of interest using the Join Attributes by Location tool. The attributes from the resulting joined GIS file were then exported into a new MS Access database for analysis.

## **Qualitative Analyses**

From the total study dataset, 251 cases (19%) had chat, email, or social network files that informed qualitative analyses. Received case file content was scanned to PDF and uploaded to a secure file server for coding and analysis using MaxQDA software (VERBI GmbH Berlin, Germany). Using MaxQDA, text passages were marked and tagged with commentary or codes. The passages, codes, and commentaries were then sorted into hierarchies and grouped into categories.

The lead qualitative analyst (co-investigator DeHart) conducted first-cycle coding on a subset of 20 cases with provisional codes developed based on sexual offending literature (e.g., "grooming", "sexual interests," "request to meet;" Kontostathis, Edwards, Leatherman, 2010; Olsen et al., 2007). In addition, open coding was performed to further delineate data and provide analytic leads for exploration (Saldana, 2009). This 20 case subset was provisionally grouped into "families" representing possible typologies based on chat content (e.g., interested in children, wanting to meet for sex, online sex only). This process enabled identifying codes that could potentially differentiate various types of offenders, such as those specifically seeking children (as opposed to those seeking a sexual interaction regardless of the target's age). Opportunistic offenders, who were seeking an in-person meeting, kept chats brief and moved quickly to scheduling a personal meeting. "Online-only" offenders would more typically engage in or seek masturbation during the chat.

To investigate these leads, a second coder applied selected codes (e.g., "exposes self," "attempts to schedule," "real-time sex") to all 200 chat logs. Second-cycle axial coding was then conducted to differentiate and organize codes (Saldana, 2009). The analysts communicated throughout the coding process to address discrepancies, clarify concepts, and refine codes based on consensus (Hill, Knox, Thompson, Williams, Hess, & Ladany, 2005; Sandelowski & Barroso,

2003). As Forman and Damshroder (2008) have explained, such an approach supports construct development and validity.

**Case classification.** Coders performed a quantitative step by using dummy codes for identifying the presence or absence of specific event categories (e.g., engaging in real-time cybersex, requesting a meeting) and event timing (e.g., time-to-onset for introducing sexual content) and exporting the dataset into SPSS.

Next, using the qualitatively identified subset of key codes, the analyst used exploratory cluster analyses to determine which codes would aid in developing offender groupings. The following six key binary codes that were applied to cases were used in the cluster analysis: (1) whether the offender exposed himself sexually via photos or video; (2) whether the offender sought sexually themed photos of the victim; (3) whether the offender showed the victim third-party pornography; (4) whether the offender engaged in or encouraged real-time masturbation; (5) whether the offender attempted to schedule a meeting (i.e., specific times and places); and (6) whether the offender mentioned child-specific or incest topics in the chat.

Although grouping varied with the clustering methods used (e.g., two-step, hierarchical, k-means), the one revealed most frequently was that real-time masturbation and scheduling most impacted group classification. Using qualitative methodology to assess the exploratory clusters lead to classifying cases as one of three types: engaged in or encouraged real-time masturbation but didn't attempt meeting (*cybersex*); attempted scheduling to meet, but no real-time masturbation (*schedulers*); and online masturbation and scheduled meeting (*cybersex/schedulers*). This typology is similar to that of Briggs et al. (2011) differentiation of fantasy-driven (focus is on cybersex interactions) and contact-driven solicitation (focus is on meeting offline).

Using thumbnail case descriptions, an analyst sorted cases into three categories and that process led to adding a fourth category, that was a further distinction of the scheduler group in which chats included a third party (e.g., pimp, family member) and the goal was child sex trafficking (*buyers*). Several cases were dropped from the final set based on either a pre-existing relationship between offender and victim (19 cases) or lack of both masturbation and scheduling (2 cases). This left the groupings of the final sample set as cybersex-only offenders (n = 48), schedulers (n = 44), cybersex/schedulers (n = 64), and buyers (n = 23).

### **Miscellaneous Data Topics**

*Issues of Data Protection.* Data was extracted from closed ICAC cases involving persons arrested for sexual offenses against children (or child personas of undercover law enforcement personnel) that originated via use of the Internet. The study did not involve contact with human participants. Data selection, safety and monitoring were addressed as following: (1) Case Selection Criteria: All closed cases resulting in conviction were considered for inclusion. ICAC task forces provided their data to the South Carolina (SC) Attorney General's office as a central collection point. The research team partnered with the SC Attorney General's office (specifically the ICAC Task Force commander and staff) to enable the participating ICAC Task Forces (TFs) to supply their files without having to redact sensitive undercover operational methods information or conduct de-identification of study subject data, which would be time and resource consuming potentially preventing participation. Another advantage of a centralized site for removal of this information before being provided to the research team was enabling oversight of consistency of methods; (2) Data safety. After extraction, all data was maintained on a secure, password-protected university server. Data was electronically transferred via the same means used to transfer legal case data from ICACs to the SC Attorney General's office. All presentations and publications will include aggregate data only so that no single case will be



identifiable. (3) Data monitoring. This was a retrospective record review, therefore direct contact with subjects did not occur and a formal data safety and management board (DSMB) was not warranted. Nevertheless, strict controls were put into practice and all study procedures were approved and monitored by the principal investigator's institutional review board (IRB).

Identification and management of relevant ethical issues was performed as follows: (1) Institutional Review Board (IRB): No data collection or analysis by study personnel was conducted until approval had been received from the MUSC IRB and co-investigator IRBs, where required. (2) Financial conflicts: There were no financial relationships between the study investigators and OJJDP. There are no commercial sponsors of this study. (3) Equity Interest in a Sponsoring Company or Company's Product: There were no pertinent equity interests to declare.

## **Results**

Findings are addressed based on methodology and each of the associated study aims.

### **Primary applied aim: Algorithm Development**

The study design was created with the goal of being able develop an algorithm using collected data for identifying offenders most likely to attempt an in-person meeting with a child first solicited online. The depth and breadth of quantitative data concerning offenders, victims to include undercover law enforcement officer personas of victims, and offense behavior varied to such an extensive degree across data sites that this was not possible at a national, generalizable level. The original sought depth and breadth of data as obtained during the pilot from a single ICAC Task Force did not meet expectations. As a result, the use of more sophisticated multivariate analysis was not possible.

Demographic and descriptive statistics are available and reported below. More success was experienced with chat room data; although still limited in scope of cross sectional

representation, but nonetheless sufficient for drawing conclusions and making recommendations as set forth herein.

### **Primary scientific aim: Expansion of Empirical Foundation for Online Offenders**

#### **Behavioral Typology**

Quantitative database: The database consisted of 1,341 usable cases. The majority of cases (57%) came from just two states, SC and UT, reflecting state differences in cases and in collaboration with the project. Because the data came from many different states, it was common to have missing variables.

The cases were divided into four categories: Child Pornography (child pornography) offenses, Child Solicitation offenses, Both (solicitation & child pornography) offenses and Contact (online & contact) offenses.

**Table 1. Distribution of cases across states**

<b>State</b>	<b>Child Pornography</b>	<b>Solicitation</b>	<b>Both</b>	<b>Contact</b>	<b>Total</b>
AK	91	27	1	15	134
CA	28	7	7	0	42
FL	1	33	2	0	36
GA	32	14	2	4	52
NE	84	14	7	14	119
OK	35	5	0	1	41
SC	112	217	8	1	338
UT	377	63	20	2	462
WA	16	7	2	0	25
Other	48	38	2	4	92
<b>Totals</b>	<b>824</b>	<b>425</b>	<b>51</b>	<b>41</b>	<b>1341</b>

































whether the offender had more content depicting boys so that later evaluators are aware is relevant to risk of sexual recidivism as well as psychiatric diagnosis and treatment planning.

Differences in online behaviors such as exposing oneself sexually or asking for sexual photos or videos from minors, and the speed with which such requests can be made, are relevant to investigators because it suggests their early interactions with suspects can quickly reveal whether someone is likely to be interested in scheduling a meeting offline. Investigators need to be alert to this behaviors as clues of the type of offender they are likely interacting with. For example requests for images or online exposure are more likely going to stay with online activities rather than move to an in-person meeting.

Those seeking to meet engaged in significantly shorter exchanges than other types. Thus the longer the interaction, the more likely the solicitation behavior will remain online only. This is also an important consideration for prevention planning in that those seeking an in-person contact will be quick to arrange such meetings.

We believe these recommendations can aid investigators in directing the focus of online undercover operations. Doing so without purpose and structure may lead to losing contact with potential offenders, missing data of evidentiary value, or expending limited resources with little chance of fruitful outcome. It is notable that offender classifications covaried with the investigative techniques and sample sites. For example, 65% of trafficking cases were identified from undercover officer posted ads and 100% of cybersex and 97% of cybersex/schedulers were identified from chats. These differences reveal that different investigative methods will result in detecting different types of offenders. Although not a focus of this study, it was noted during qualitative analysis that some investigators varied on whether they encouraged or discouraged in-person meetings and/or real-time sexual activity online. Future research should explore how

these differences impact online offending behavior and which approach or approaches prove most beneficial for detecting higher risk offenders.

More study is needed regarding the behavior of those engaging in internet-based offending with child victims. This study adds to existing research on types of online offenders, their varying sexual interest in prepubescent vs postpubescent children, and their online offending behavior. This study also identified new avenues and questions to pursue. It also revealed the difficulty in conducting a retrospective file-based study using data that were not originally collected for research purposes.

## References

- Briggs, P., Simon, W. T., & Simonsen, S. (2011). An exploratory study of internet-initiated sexual offenses and the chat room sex offender: Has the Internet enabled a new typology of sex offender? *Sexual Abuse: A Journal of Research and Treatment, 23*, 72-91. Doi: 10.1177/1079063210384275
- Canwest News Service (2009). Information was retrieved September 2, 2009 from <http://bit.ly/ecfibF>
- Cohen, J. (1988). *Statistical power analyses for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Dietz, P. E., Mathews, D. B., Van Duyne, C., Martell, D. A., Parry, C. D. H., Stewart, T., Warren, J., Crowder, J. D. (1991). Threatening and otherwise inappropriate letters to Hollywood celebrities. *Journal of Forensic Sciences, 36*, 185-209.
- Dietz, P. E., Mathews, D., Martell, D., Stewart, T., Hrouda, D., Warren, J. (1991). Threatening and otherwise inappropriate letters to members of the United States Congress. *Journal of Forensic Sciences, 36*, 1445-1468.
- Eke, A. W., Seto, M. C., & Williams, J. (2011). Examining the criminal history and future offending of child pornography offenders: An extended prospective follow-up study. *Law and Human Behavior, 35*, 466-478. <http://dx.doi.org/10.1007/s10979-010-9252-2>
- Faust, E., Renaud, C., & Bickart, W. (2009, October). *Predictors of re-offense among a sample of federally convicted child pornography offenders*. Paper presented at the 28<sup>th</sup> annual conference of the Association for the Treatment of Sexual Abusers, Dallas, TX.
- Forman, J., & Damschroder, L. (2008). Qualitative content analysis (pp.39-62). In L. Jacoby & L. Siminoff (Eds.) *Empirical research for bioethics: A primer*. Oxford, UK: Elsevier.
- Hill, C. E., Knox, S., Thompson, B., Williams, E., Hess, S., & Ladany, N. (2005). Consensual Qualitative Research: An Update. *Journal of Counseling Psychology, 52*(2), 196-205.

Kontostathis, A., Edwards, L., & Leatherman, A., (2010). Text mining and cybercrime. In M.

Berry and J. Kogan (Eds.) *Text Mining: Applications and Theory*, pp.149-164.

Long, M., Alison, L., Tejeiro, R., Hendricks, E., Giles, S. (2016). KIRAT: Law enforcement's prioritization tool for investigating indecent image offenders. *Psychology, Public Policy, and Law*, 22(1): 12-21.

Leatherman, A. (2009). Luring language and virtual victims: Coding cyber-predators' online communicative behavior. Technical report. Collegeville, PA: Ursinus College.

Marcum, C. D., Higgins, G. E., & Ricketts, M. L. (2010). Potential factors of online victimization of youth: An examination of adolescent online behaviors utilizing routine activity theory. *Deviant Behavior*, 31(5), 381-410.

Mitchell, K. J., Finkelhor, D., & Wolak, J. (2007). Youth Internet users at risk for the most serious online sexual solicitations. *American Journal of Preventive Medicine*, 32(6), 532-537.

Mitchell, K., Jones, L., Finkelhor, D., & Wolak, J. (2014). *Trends in unwanted sexual solicitations: Findings from the Youth Internet Safety Studies*. Durham, NH: Crimes Against Children Research Center.

Motivans, M., & Kyckelhahn, T. (2007). Federal prosecution of child sex exploitation offenders, 2006. *Bureau of Justice Statistics Bulletin* (Report No. NCJ 219412). Washington, DC: Bureau of Justice Statistics.

Olson, L., Daggs, J., Ellevold, B., & Rogers, T. (2007). Entrapping the Innocent: Toward a Theory of Child Sexual Predators' Luring Communication. *Communication Theory* (10503293), 17(3), 231-251.

- Quinsey, V. L., Harris, G. T., Rice, M. E., & Cormier, C. A. (2006). *Violent offenders: Appraising and managing risk (2<sup>nd</sup> ed.)*. Washington, DC: American Psychological Association.
- Saldana, J. (2009). *The coding manual for qualitative researchers*. Los Angeles, CA: Sage.
- Sandelowski, M., & Barroso, J. (2003). Writing the proposal for a qualitative research methodology project. *Qualitative Health Research, 13*(6), 781-820.
- Schwartz, B., Zaitsev, P., & Tkachenko, V. (2012). *High Performance MySQL: Optimization, Backups, and Replication, 3rd Ed.* Sebastopol, CA: O'Reilly & Associates Inc, pp. 133-136.
- Seto M. (2014). Chapter 4: Internet-Facilitated Sexual Offending. In *USDOJ Sex Offender Management Assessment and Planning Initiative*. Washington, DC: U.S. Department of Justice (USDOJ).
- Seto, M. C. (2005). Is more better? Combining actuarial risk scales to predict recidivism among adult sex offenders. *Psychological Assessment, 17*, 156-167.
- Seto, M. C. (2008). *Pedophilia and sexual offending against children: Theory, assessment, and intervention*. Washington, DC: American Psychological Association.
- Seto, M. C. (2010). Child pornography use and Internet solicitation in the diagnosis of pedophilia [letter to the editor]. *Archives of Sexual Behavior, 39*(3), 591-593. doi:10.1007/s10508010-9603-6
- Seto, M. C., Cantor, J. M., & Blanchard, R. (2006). Child pornography offenses are a valid diagnostic indicator of pedophilia. *Journal of Abnormal Psychology, 115*, 610-615.
- Seto, M. C., & Eke, A. W. (2005). The future offending of child pornography offenders. *Sexual Abuse: A Journal of Research and Treatment, 17*, 201-210.

- Seto, M. C., & Eke, A. W. (2015). Predicting recidivism among adult male child pornography offender: Development of the Child Pornography Offender Risk Tool (child pornographyORT). *Law and Human Behavior, 39*, 416-429.  
<http://dx.doi.org/10.1037/lhb0000128>
- Seto, M. C., & Hanson, R. K. (2011). Introduction to Special Issue on Internet-facilitated sexual offending. *Sexual Abuse: A Journal of Research and Treatment, 23*, 3-6.  
Doi:10.1177/1079063211399295
- Seto, M. C., Hanson, R. K., & Babchishin, K. M. (2011). Contact sexual offending by men arrested for child pornography offenses. *Sexual Abuse: A Journal of Research and Treatment, 23*, 124-145. Doi:10.1177/1079063210369013
- Seto, M. C., & Lalumière, M. L. (2001). A brief screening scale to identify pedophilic interests among child molesters. *Sexual Abuse: A Journal of Research and Treatment, 13*, 15-25.
- Seto, M. C., Reeves, L., & Jung, S. (2010). Motives for child pornography offending: The explanations given by the offenders. *Journal of Sexual Aggression, 16*, 169-180. doi: 10.1080/13552600903572396
- Seto, M. C., Wood, J. M., Babchishin, K. M., & Flynn, S. (2012). Online solicitation offenders are different from child pornography offenders and lower risk contact sexual offenders. *Law and Human Behavior, 36*, 320-330. <http://dx.doi.org/10.1037/h0093925>
- United States Department of Justice (2010). The national strategy for child exploitation prevention and interdiction: A report to Congress. Washington, DC: USDOJ.
- Wolak, J., Finkelhor, D., & Mitchell, K. J. (2011). Child pornography possessors: Trends in offender and case characteristics. *Sexual Abuse: A Journal of Research and Treatment, 23*, 22-42
- Yin R. (1989). Case study research design and methods. Newbury Park, CA: Sage Publications.

















## Non-Family Male Households in census blocks in which offenders reside

<b>Offense Type</b>	<b>Average Percentage of Non-Family Households</b>	<b>Average Percentage of Non-Family Male households</b>	<b>Average Percentage of Males Living Alone households</b>	<b>Average Percentage of Males Note Living Alone households</b>
Pornography	31.9%	49.9%	36.7%	13.3%
Chat	32.7%	48.8%	36.9%	11.9%
Both	30.4%	49.8%	37.3%	12.5%
Missing	33.4%	51.1%	38.2%	12.9%

## Non-Family Female Households in census blocks in which offenders reside

<b>Offense Type</b>	<b>Average Percentage of Non-Family Households</b>	<b>Average Percentage of Non-Family Female Households</b>	<b>Average Percentage of Females Living Alone households</b>	<b>Average Percentage of Females Note Living Alone households</b>
Pornography	31.9%	50.1%	40.8%	9.3%
Chat	32.7%	51.2%	42.9%	8.3%
Both	30.4%	50.2%	40.6%	9.7%
Missing	33.4%	48.9%	41.8%	7.1%