Faculty Publications
Department of Computer Information Technology

2014

Using Internet Artifacts to Profile a Child Pornography Suspect

Kathryn C. Seigfried-Spellar
Purdue University, kspellar@purdue.edu

Marcus K. Rogers
Purdue University, rogersmk@purdue.edu

USING INTERNET ARTIFACTS TO PROFILE A CHILD PORNOGRAPHY SUSPECT

Marcus K. Rogers, PhD
Department of Computer Information & Technology
Purdue University
401 N. Grant Street
West Lafayette, IN 47907
Office: 765-496-1072
Fax: 765-496-1212
Email: cyberforeniscs@mac.com

Kathryn C. Seigfried-Spellar, PhD
Department of Criminal Justice
The University of Alabama
410 Farrah Hall
Tuscaloosa, AL 35487
Office: 205-348-5489
Fax: 205-348-7178
Email: kseigspell@as.ua.edu

*Equal contribution; authorship listed alphabetically

ABSTRACT

Digital evidence plays a crucial role in child pornography investigations. However, in the following case study, the authors argue that the behavioral analysis or "profiling" of digital evidence can also play a vital role in child pornography investigations. The following case study assessed the Internet Browsing History (Internet Explorer Bookmarks, Mozilla Bookmarks, and Mozilla History) from a suspected child pornography user's computer. The suspect in this case claimed to be conducting an ad hoc law enforcement investigation. After the URLs were classified (Neutral; Adult Porn; Child Porn; Adult Dating sites; Pictures from Social Networking Profiles; Chat Sessions; Bestiality; Data Cleaning; Gay Porn), the Internet history files were statistically analyzed to determine prevalence and trends in Internet browsing. First, a frequency analysis was used to determine a baseline of online behavior. Results showed 54% ($n = 3205$) of the URLs were classified as "neutral" and 38.8% ($n = 2265$) of the URLs were classified as a porn website. Only 10.8% of the URLs were classified as child pornography websites. However when the IE history file was analyzed by visit, or "hit," count, the Pictures/Profiles (31.5%) category had the highest visit count followed by Neutral (19.3%), Gay Porn (17%), and Child Porn (16.6%). When comparing the frequency of URLs to the Hit Count for each pornography type, it was noted that the accused was accessing gay porn, child porn, chat rooms, and picture profiles (i.e., from Facebook) more often than adult porn and neutral websites. The authors concluded that the suspect in this case was in fact a child pornography user and not an ad hoc investigator, and the findings from the behavioral analysis were admitted as evidence in the sentencing hearing for this case. The authors believe this case study illustrates the ability to conduct a behavioral analysis of digital evidence. More work is required to further validate the behavioral analysis process described, but the ability to infer the predilection for being a consumer of child pornography based on Internet artifacts may prove to be a powerful tool for investigators.

Keywords: Internet child pornography, digital forensics, computer crime investigation, Internet artifacts, profiling, behavioral analysis

1. INTRODUCTION

There is currently no accurate way to determine the number of individuals who are using child pornography (Wortley & Smallbone, 2012). According to the FBI, the United States has seen a 2500% increase in the last ten years in the number of child pornography arrests (2012). In addition, the United Kingdom's Internet Watch Foundation's Hotline (IWF, 2011) reported 12,966 webpages contained child sex abuse images, and 49% of those websites were hosted in North America. As of August 2009, the CyberTipline of the United States' National Center for Missing and Exploited Children (NCMEC) reported receiving over 85,000 tips related to child pornography in 2008 for a total of 625,271 child pornography tips since its establishment in March 1998 (Wolak, Finkelhor, & Mitchell, 2009). Finally, when comparing the National Juvenile Online Victimization (N-JOV) study in 2000 to 2006, the number of offenders arrested solely for child pornography possession or distribution more than doubled from 935 to 2,417 arrests, respectively (Wolak et al., 2009).

Individuals who engage in child pornography do so at varying degrees, with some engaging in more offenses than others. In the United States, an individual may be charged with possession, distribution, or production of child pornography (United States Sentencing Commission [USSC], 2012; Wortley & Smallbone, 2012). Production refers to the creation of sexualized images of children, which includes images created from offenders recording their direct sexual abuse of children (i.e., hands-on contact offender) or through the creation of virtual child pornography (i.e., computer-generated images of child sex abuse). Distribution or trafficking is the dissemination of child sex abuse images, often through peer-to-peer networks or email, and is referred to as "receipt, transportation, and distribution" (R/T/D; USSC, 2012). Lastly, an individual may be charged with possession of child pornography for downloading images from the Internet; however, "possession" may also occur even if the individual did not actively download the image (e.g., individual viewed an image which was cached by the web browser; USSC, 2012).

According to the Federal Child Pornography Offenses report (USSC, 2012), the number of child pornography cases has steadily increased for all child pornography related offenses, with the largest increase seen for possession and distribution (R/T/D). For example, the number of child pornography offenders sentenced to possession and/or "R/T/D" increased from 90 in 1994 to 1649 in 2011 (USSC, 2012). There is no doubt that technological advances, such as the Internet, as well as increased awareness and dedication of resources for targeting child pornography offenders have contributed to its significant growth (USSC, 2012). However, growth of this crime is only expected to increase as the current 39% of the world's population with Internet access continues to grow as well (Internet World Stats, 2014). This growth will only add importance to understanding "why" child pornography users engage in different types of child pornography behaviors.

As heightened efforts by law enforcement continue to increase, Wolak, Finkelhor, and Mitchell (2011) believe a better understanding of the offender population is needed in order to differentiate between those offenders who only engage in child pornography verses those who are also hands-on contact offenders. Relatively new research suggests there are differentiating characteristics between contact and non-contact offenders. McCarthy (2010) compared two groups of child pornography offenders; 51 were contact offenders and 56 were non-contact offenders. Results indicated a significant difference in how the two groups used Internet child pornography; contact offenders were significantly more likely to masturbate to Internet child pornography and download the images onto another external device (other than a computer hard drive; McCarthy, 2010). In addition, the child pornography users who were involved in a higher number of child pornography behaviors (exchanging, paying for images, concealing and organizing collection) were more likely to be in the contact offender group (McCarthy, 2010). Finally, McCarthy (2010) suggested the ratio of adult pornography to child pornography was significantly different between groups in that the contact offenders were more likely to possess a higher ratio of child to adult pornographic images compared to the non-contact group.

Overall, individuals who engage in child pornography do so at varying degrees, with some offenders engaging in more offenses than others. Child pornography offenses may be categorized as production, distribution, or possession, and individuals may be involved in some or all of these offenses (Wortley & Smallbone, 2012). The overabundance of child pornography cases surpasses law enforcement's ability to effectively investigate cases (Eke, Seto, & Williams, 2011). If a suspect is involved in some or all of these child pornography offenses, then law enforcement must be able to determine which crime(s) have been committed. In other words, is the suspect a closet child pornography collector (i.e., possession only) or a hands-on contact offender (i.e., possession and producer)? Therefore, the problem for law enforcement is determining which offenders, who are initially suspected of child pornography possession or distribution charges, may also be hands-on contact offenders.

However, research suggests there are significant differences between contact and non-contact child pornography offenders. The one thing these different child pornography offenses have in common is the use of technology – specifically the Internet and digital devices. Technology may assist child pornography users in the possession, distribution, and production of Internet child pornography, but these same technologies are capable of providing incriminating computer forensic evidence (Rogers & Seigfried-Spellar, 2011). It is these differences that the current study seeks to identify using the actual computer forensic evidence collected from contact and non-contact child pornography cases. By behaviorally analyzing the computer forensic evidence of suspected offenders, law enforcement may be able to better prioritize between crimes by quickly identifying which offenders are more likely to be contact versus non-contact offenders (Rogers & Seigfried-Spellar, 2009; Rogers & Seigfried-Spellar, 2012).

The following case study illustrates the ability to conduct a behavioral analysis based on Internet artifacts of a suspected child pornography user to determine whether the individual is likely to also be a hands-on contact offender. The authors assessed a suspect's Internet Browsing History (specifically Internet Explorer Bookmarks, Mozilla Bookmarks, and Mozilla History) to identity any trends in

pornography use. Finally, the authors discuss the feasibility in conducting a behavioral analysis of Internet artifacts (URLs) to differentiate between Internet child pornography users and child sex offenders.

2. CASE STUDY

The authors were asked by Law Enforcement to examine Internet Artifacts belonging to a computer seized from a suspect who was arrested and indicted for the possession of child pornography. The accused was a former deputy sheriff who claimed he came across the pictures while conducting his own examination of sites that hosted potential child pornography. To back up this claim, the accused indicated he had submitted two police reports to his department and five reports to the National Center for Missing and Exploited Children (NCMEC). These reports were time and date stamped and provided to the authors. The authors were asked to examine the Internet artifacts on the suspect's computer and determine if the evidence indicated behavior that was consistent with someone merely carrying out an investigation or not.

2.1 Tools

The Internet history files were analyzed using TimeFlow Analytical Timeline. TimeFlow is a data analysis tool, which allows researchers to assess trends in data over a period of time (Cohen, 2010). Specifically, events may be analyzed by day, month, or year. In this case study, the events analyzed were URLs visited by the suspect, so TimeFlow allowed the authors to determine any behavioral trends in pornography use by calendar month/year. All data was analyzed using IBM's Statistical Package for the Social Sciences (SPSS).

2.2 Design & Procedure

2.2.1 Phase 1
The first phase consisted of positively identifying artifacts that belonged to the user profile of the accused. The investigators determined that this user profile and account was not shared with any other persons. The Internet artifacts were filtered to remove any entry that was not linked to the accused's user-id. After the filtering process, the Internet Explorer History file contained the most entries, and this file was used as the primary basis

for the analysis and conclusions. The other Internet artifacts (listed above) were examined and analyzed as supplemental data in order to confirm or refute findings drawn from the Internet Explorer History (IE History File).

2.2.2 Phase 2

The IE History file was converted to a comma separate values (CSV) format to facilitate the analysis and examination. Once converted, the IE History file was sorted by the Uniform Resource Locator (URL) name in order to facilitate proper classification. The file contained 5841 entries or events that were used for data analysis. Each entry was classified by both authors based on the URL visited or activity logged. The classifications were then compared and a consensus was reached concerning the appropriate categorization, or else the URL was flagged as unknown. After an initial examination, it was determined that the entries (data) could be classified using a system made up of 9 categories: 0 = Neutral; 1 = Adult Porn; 2 = Child Porn; 3 = Adult Dating; 4 = Pictures/Profiles; 5 = Chat Sessions; 6 = Bestiality, 7 = Data Cleaning; and 8 = Gay Porn (see Table 1).

If the URL name was not recognized as belonging to any of the categories listed from 0-8, it was assigned as "neutral" (0). Given the nature of the analysis, it was deemed appropriate to err on the side of inflating the false negatives (e.g., true child porn or adult porn URLs being classified as neutral). When the URL name was not recognizable and/or no consensus could be reached on the appropriate category, and the nature of site could not be confirmed by any information in the entry (e.g., name of file downloaded or viewed), this entry was flagged as unknown. After classifying the known URLs, any unknown URLs were sent to the Indiana State Police Department's Internet Crimes against Children taskforce who verified whether the URL should be classified as Child Porn or some other category.

2.2.3 Phase 3

The IE History File was additionally sorted by visit count. The visit count field is a rough estimate of the number of times a particular URL was visited. IE, however, does not update this count consistently, and therefore, this number is only used as an estimate.

2.2.4 Phase 4

Phase focused on mapping the category of sites visited (URLs) on a timeline in order to determine if any patterns were present. For this process, the authors used the last-visited meta-data as the time stamp of the URL entry (need a reference here to justify this date).

2.2.5 Phase 5

The content of the seven reports that the accused submitted were studied, and the indicated URLs in the report, along with the dates recorded, were compared to the IE History file entries and the derived timeline.

Table 1 Classification System for Internet Browsing History

Coding	Category	Content
0	Neutral	Neutral sites or system activities not fitting into any of the other categories.
1	Adult Porn	Pornography/stories/other content related to adults or depicting adults.
2	Child Porn	Pornography/stories/other content related to or depicting children under the age of 18 years.
3	Adult Dating	Adult dating sites
4	Pictures/Profiles	Social network profiles of individuals and/or pictures listed on the social network site.
5	Chat Sessions	Sessions related to online chat behaviors.
6	Bestiality	Pornography/stores/other content related to or depicting sexual acts with animals.
7	Data Cleaning	Related to software/tools that could be used to obfuscate data or evidence.
8	Gay/Gay Porn	Pictures/pornography/stories/other content related to "gay, lesbian, or transgender".

3. CONCLUSION OF BEHAVIORAL ANALYSIS

After the URLs were classified, the Internet history files were statistically analyzed to determine prevalence and trends in Internet browsing. First, a frequency analysis was used to determine a baseline of online behavior. As shown in Table 2, 54% (n = 3205) of the URLs were classified as "neutral" and 38.8% (n = 2265) of the URLs were classified as a porn website (see Figure 1). When only considering the frequency of URLs, there were more adult pornography URLs (17.5%) compared to child pornography (10.8%), gay pornography (10.5%), and bestiality (.2%).

Table 2 Frequency of Classification Categories for Internet Browsing History

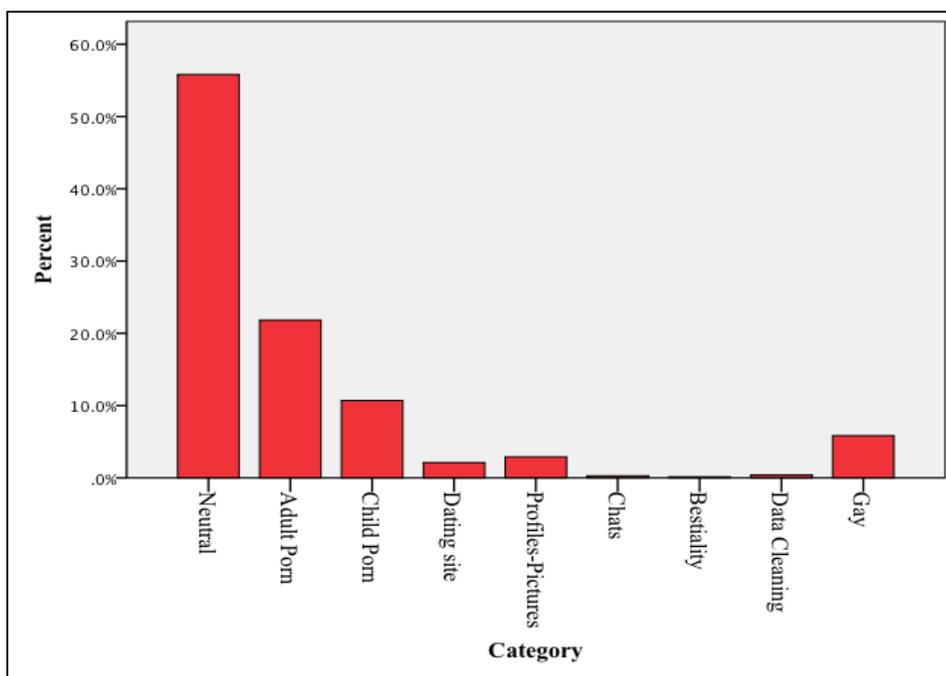| Category	Frequency	Percent
Neutral	3205	54.9
Adult Porn	1021	17.5
Child Porn	628	10.8
Gay Porn	616	10.5
Profiles/Pictures	196	3.4
Adult Dating	124	2.1
Data Cleaning	26	0.4
Chat Sessions	16	0.3
Bestiality	9	0.2
Total	**5841**	**100**

Figure 1 Percentage of Classification Categories for Internet Browsing History

Next, the Internet history files were analyzed using TimeFlow analysis tool. As shown in Figure 2, TimeFlow displays "hot spots" for Internet browser activity based on URL category type. For example, child porn is represented by the neon green "hot spot." Lastly, the IE history file was analyzed by visit, or "hit," count.

Figure 2 TimeFlow Analysis for URL Category by Calendar Month

As shown in Figure 3, the Pictures/Profiles (31.5%) category had the highest visit count followed by Neutral (19.3%), Gay Porn (17%), and Child Porn (16.6%). When comparing the Frequency Graph (Figure 1) to the Hit Count Graph (Figure 3), it was noted that the accused was accessing gay porn, child porn, chat rooms, and picture profiles (i.e., from Facebook) more often than adult porn and neutral websites.

The behavioral patterns obtained from the analysis of the IE History file were consistent with someone that was personally interested in the content of the sites visited, as opposed to fitting the pattern expected from a police investigation, whether formal or not. Based on the frequency analysis and the type of the sites visited, it was concluded that the suspect had preference for same-sex pornography and adolescent male child pornography. The vast majority of the same-sex pornography sites (Gay Porn) contained references to teen boys. This preference was consistent with the classification of a sexual deviance with online paraphilia centered on adolescent males[1]. In addition, the percentage of websites visited that were classified as Child Porn (10.8%), Gay Porn (10.5%) and Picture/Profile (3.4%) provided support that this behavior was preferential.

---

[1] It should be noted that this is not intended to be a clinical diagnosis. This categorization is for investigative purposes.
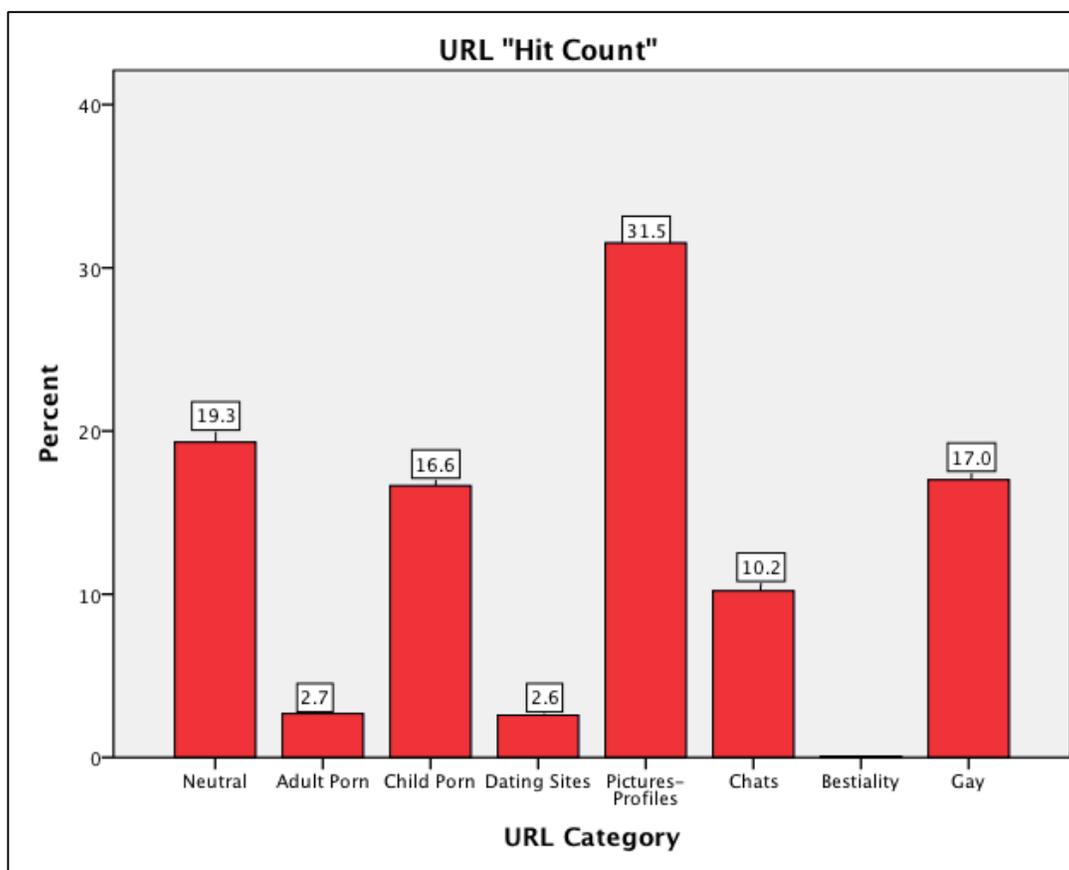
Figure 3 URL Visit or Hit Count by URL Classification Category

Furthermore, the time analysis indicated that the majority of the visited Child Porn sites occurred in 2005-2008, with early spring (March-April) and summer (July-August) accounting for the highest number. If the motivation for this behavior were investigative, then one would expect to see reports being filed at the end of these viewing cycles. However, no reports were submitted during these periods. Furthermore, the fact that the suspect was also visiting adult porn and bestiality sites fits the pattern of a consumer of child pornography, since previous research indicates consumers of child pornography engage in a similar pattern of non-deviant and deviant pornography use, specifically viewing Adult Porn, Bestiality, and Child Porn (see Seigfried-Spellar, 2013; Seigfried-Spellar & Rogers, 2011; Seigfried-Spellar & Rogers, 2013). In addition, the percentage of websites visited for Picture/Profile and Chat Rooms suggest the suspect was moving from fantasy-driven (online cybersex only) to contact-driven (intentions to meet offline) behavior (Briggs, Simon, & Simonsen, 2011).

The findings from the behavioral analysis were admitted as evidence in the sentencing hearing for this case. The federal prosecutor's office successfully argued that the findings painted a much different picture of the suspect and his activities than was proposed by the defense, who argued that the suspect/defendant had been conducting an ad hoc law enforcement investigation. The analysis clearly indicated the behavior was consistent with someone personally interested in sexual pictures of adolescent males. The judge in this case ruled that the defendant had falsely denied conduct (sexual interest in adolescent boys) that was relevant to the sentencing guideline calculation (U.S.S.G. § 3E1.1).

More work is required to further validate the behavioral analysis process described, but the ability to infer the predilection for being a consumer of child pornography based on Internet artifacts may prove to be a powerful tool for investigators.

REFERENCES

1. 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*(1), 72-91.

2. Cohen, S. (2010). *TimeFlow Analytical Timeline*. (Alpha ed.) Retrieved from https://github.com/

3. Cooper, A. (1998). Sexuality and the Internet: Surfing into the new millennium. *CyberPsychology & Behavior*, *1*(2), 187-193.

4. 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.

5. Frei, A., Erenay, N., Dittmann, V., & Graf, M. (2005). Paedophilia on the Internet - A study of 33 convicted offenders in the canton of Lucerne. *Swiss Medical Review, 135*, 488-494.

6. International Telecommunication Union. (2011). *The world in 2011: Facts and figures*. Retrieved from www.itu.int/ict

7. Internet Watch Foundation. (2012). *IWF Operational Trends 2012*. Retrieved from The United Kingdom's Internet Watch Foundation's Website www.iwf.org.uk

8. Klain, E.J., Davies, H.J., & Hicks, M.A. (2001). *Child Pornography: The Criminal-Justice-System Response*. National Center for Missing and Exploited Children. Retrieved from ncmec.org

9. Krone, T. (2005, April). Does thinking make it so? Defining online child pornography possession offenses. *Trends & Issues in Crime and Criminal Justice*, *299*, 1-6.

10. McCarthy, J. A. (2010). Internet sexual activity: A comparison between contact and non-contact child pornography offenders. *Journal of Sexual Aggression, 16*(2), 181-195.

11. O'leary R., & D'Ovidio R. (2007). Online sexual exploitation of children. *The International Association of Computer Investigative Specialists*. Retrieved from www.nga.org

12. Quayle, E., & Taylor, M. (2003). Model of problematic Internet use in people with a sexual interest in children. *CyberPsychology & Behavior, 6*, 93-106.

13. Rogers, M., & Seigfried-Spellar, K. (2009). The future of digital forensics: Merging behavioral science and digital evidence. Position paper presented at the Indo-US Conference and Workshop on Cybersecurity, Cybercrime, and Cyberforensics, Kochi, India, August 2009.

14. Rogers, M.K., & Seigfried-Spellar, K.C. (2011). *Internet child pornography: Legal issues and investigative tactics*. In T Holt (Ed.), Crime On-Line: Correlates, Causes, and Context. Durham, NC: Carolina Academic Press.

15. Rogers, M.K., & Seigfried-Spellar, K.C. (2012). Applied predictive behavioral modeling: The role of behavioral sciences in digital forensics. Presentation at the American Academy of Forensic Sciences 64th Annual Scientific Meeting, Atlanta, GA, February 2012.

16. Seigfried-Spellar, K.C. (2013). Replicating the Seigfried-Spellar and Rogers (2011) Study on deviant pornography use by age of onset and sex. Presentation at the American Academy of Forensic Sciences 65th Annual Scientific Meeting, Washington, D.C, February 2013.

17. Seigfried-Spellar, K., & Rogers, M. (2011). Exploring the progression of nondeviant and deviant pornography use by age of onset and sex. Paper presented at the American Academy of Forensic Sciences 63rd Annual Scientific Meeting, Chicago, IL, February 2011.

18. Seigfried-Spellar, K., & Rogers, M. (2013). Does deviant pornography use follow a Guttman-like progression? *Computers in Human Behavior*, *29*, 1997-2003.

19. United States Department of Justice. (2010). National strategy for child exploitation prevention and interdiction. Retrieved from www.justice.gov/psc/docs/natstrategyreport.pdf

20. United States Sentencing Commission. (2012). Federal child pornography offenses. Retrieved from http://www.ussc.gov

21. Wolak, J., Finkelhor, D., & Mitchell, K.J. (2009). Law enforcement responses to online child sexual exploitation crimes: The national juvenile victimization study. Retrieved from University of New Hampshire Crimes Against Children Research Center at www.unh.edu/ccrc

22. 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*(1), 22-42.

23. Wortley, R., & Smallbone, S. (2012). Child pornography on the Internet. *Problem-Specific Guides Series: Problem-Oriented Guides for Police, 41*, 1-48.