Identifying Grooming of Children for Sexual Abuse: Gender Effects and Increased False Positives from Internet Information

Natalie Bennett*, William O'Donohue

University of Nevada, Reno, USA

ABSTRACT

Some child sexual abusers engage in grooming prior to committing abuse. These tactics are methods that perpetrators use to gain access to their future victims and prepare them to be more compliant with the abuse. At times, grooming can be difficult to distinguish from typical adult-child interactions, and definitions of grooming have varied. The current study examined the extent to which participants could be trained to more accurately differentiate grooming behavior based upon participation in three types of training. One-hundred forty-seven undergraduate psychology students were invited to complete an online study. Participants were presented with hypothetical scenarios describing potential grooming behaviors and were asked to make judgments about the appropriateness of each behavior and whether or not it was indicative of grooming. Then participants were randomized to one of three training groups: 1) an experimental presentation which trained participants to focus on the behavior's function and context; 2) a treatment-as-usual (TAU) presentation designed to mimic what a person would find out about grooming on the internet; and 3) a no treatment control. Afterward, participants were again asked to respond to hypothetical scenarios. The experimental and control groups performed equally well at post-test; however, the TAU group performed significantly worse (*p*= .003) suggesting a possible iatrogenic effect of information from the internet. In addition, there was a gender effect in which females outperformed males when labeling grooming behaviors (*p*= .027). There appears to be a need for more training about grooming behaviors and how to detect these.
Keywords: grooming, child sexual abuse, judgment, online training.

Novelty and Significance

What is already known about the topic?

• Perpetrators of child sexual abuse often "groom" their victims to make them more compliant with abuse.
• Grooming behaviors can look similar to normal behaviors seen between adults and children, such as giving gifts, tickling, attention, etc.
• Third-party observers generally have a difficult time differentiating grooming from non-grooming behaviors.

What this paper adds?

• The current study examined how well individuals can be trained to distinguish grooming behaviors, based upon the type of information participants were provided.
• Individuals who were provided information pulled from various websites about grooming performed worse than other groups when judging behaviors, indicating the potential dangers of Internet information regarding grooming.

Grooming in child sexual abuse has been conceptualized as a preparatory stage in which the perpetrator attempts to establish a pattern of secrecy and complicity with the child (Wolf & Pruitt, 2019). It has been more formally defined as antecedent inappropriate behavior on the part of an offender that functions to increase the likelihood of being able to abuse the child (Bennett & O'Donohue, 2014). In this grooming stage the abuser

_____

* Correspondence: Natalie Bennett, Department of Psychology M/S 298, University of Nevada, Reno, 89557; Email: nmbennett06@gmail.com

uses various manipulations that functions to gain increased access to the child, seclude the child, sexualize the relationship, and introduce secrets into the relationship. It is important to note that grooming does not occur in all cases of child sexual abuse, but can be an early indication that an offender is intending to abuse a child (Bennett & O'Donohue, 2014).

Definitions of grooming are highly variable and have evolved over the past few decades (Burgess & Hartman, 2018). Some authors view grooming as containing discrete stages (e.g., Winters & Jeglic, 2017; Colton, Roberts, & Vanstone, 2012; O'Connell, 2003; van Dam, 2001; Brackenridge, 2001) that generally progress from targeting certain potential victims and earning their trust, to isolation, to desensitization to physical touch, ultimately resulting in sexual abuse. For a more complete and critical review of grooming definitions, please refer to Bennett and O'Donohue (2014).

Much of the early research on grooming focused on identifying the topographies of such behavior e.g., Conte, Wolf, & Smith, 1989; Elliott, Browne, & Kilcoyne, 1995; Budin & Johnson, 1989, while more recent research has found that approximately half of child sexual offenders utilize some form of grooming strategy prior to abusing their victims (Winters & Jeglic, 2017). The topographies of such grooming behaviors can range from gift-giving and providing extra attention, to sexually desensitizing the child by talking about sex or offering to bathe or dress the child. Although positive attention is a commonly used grooming tactic, some offenders choose to provide the child with increased negative attention such as name-calling (Wolf & Pruitt, 2019), especially if the prospective victim is male (Spiegel, 2003). The intended function of this diverse behavior may be to motivate the child to want to please the adult. Thus, various topographies of grooming behavior serve at least one common function on the offender's part -to increase the likelihood that the child can be abused.

Because many grooming behaviors tend to appear similar to behaviors seen in typical adult-child relationships (e.g., gift-giving), or tend to occur in secrecy (e.g., sexual desensitization), often the offender can successfully groom a child and subsequently engage in child sexual abuse without other adults noticing or intervening. However, with proper knowledge about grooming a third-party observer could be able to accurately judge whether or not a behavior is indicative of grooming and thus potentially prevent future acts of abuse. As such, one could argue that training individuals -including nonprofessionals- to effectively detect grooming behaviors could potentially present an intervention point where one is able to intervene to prevent the abuse or stop it if it has already started.

Asking an individual to make a judgment about another's behavior is often complex, especially when that behavior may be intentionally misrepresented or disguised. Importantly, some grooming behaviors are inappropriate in and of themselves (e.g., a neighbor teaching about the birds and the bees or gifting underwear) whereas others require much more context in order to determine appropriateness (e.g., asking for hugs). Thus, it is also possible that there may not be sufficient information provided for a respondent to make an accurate judgment. Notably, a behavior can be inappropriate (e.g., providing drugs to a minor) but it may not be indicative of grooming.

In a similar vein, Winters and Jeglic (2016) examined how well undergraduate students were able to judge whether a potential perpetrator in a vignette was a child molester, in the context of hindsight bias. Hindsight bias is a cognitive tendency in which individuals perceive an event that has already occurred as having been more predictable than it actually was before the event actually took place. Results indicated that

participants who were provided outcome information about the perpetrator overestimated the likelihood that they would have correctly predicted their offender status, thus supporting the hypothesis that they demonstrated hindsight bias. Furthermore, in their sample of approximately 400 undergraduates, Winters and Jeglic (2017) found that participants were overall "unable to recognize sexual grooming behaviors" in any of the proposed stages of grooming that they examined, and the authors ultimately concluded that this task of identifying such behavior prospectively remains a difficult if not impossible one. Indeed, previous authors (e.g., Ost, 2004; Gillespie, 2002) have commented that it is unlikely that individuals such as police officers or community members are able to accurately detect grooming behaviors from non-grooming behaviors. If true, this obviously leaves children more vulnerable to abuse, than anyone would like.

Thus, relatively little is yet known about how third-party observers view and categorize potential grooming behaviors, other than that these tasks tend to be difficult. To the authors' knowledge, no scientifically evaluated training program currently exists to teach individuals to correctly detect grooming from non-grooming behaviors. As such, the current study sought to examine whether the performance of individuals tasked with judging hypothetical scenarios could be improved through training. Utilizing the definition of grooming proposed by Bennett and O'Donohue (2014), the goal of such training focused not only on discussion of grooming behaviors, but also on distinguishing the function of a behavior from its topography.

Similar to previous research (e.g., Winters & Jeglic, 2016, 2017) we utilized a sample of undergraduate psychology students. Such a sample was chosen for two reasons: 1) out of convenience, and 2) because undergraduates, especially undergraduates enrolled in psychology courses, may be similar to individuals who would be confronted with the task of identifying grooming behavior in personal or professional contexts (e.g., as daycare providers, teachers, or even in child protective service roles).

Method

Participants

Participants were undergraduate psychology students at a university in the Mountain West, recruited through class announcements, webpage postings, and posted flyers. All research activities were approved by the university's Institutional Review Board. Participants received one hour of credit for participation in the study, which can be applied as extra credit in their psychology courses. A total of 162 participants began the study; 15 dropped out prior to completion of any items on the pre-test. Therefore, data from a total sample size of 147 were analyzed.

Materials

Demographics Questionnaire. Participants were asked to provide information about their age, gender, and race/ethnicity.

Hypothetical Grooming Scenarios. A total of forty scenarios (twenty of which were previously utilized by Removed for blind review, unpublished) were presented to participants in the current study. The scenarios were developed in accordance with the authors' definition of grooming. Thus, the definition of grooming involved two necessary conditions: 1) to be categorized as grooming the behaviors of the adult in the scenario had to be judged by the participants as inappropriate (e.g., an adult

giving a gift is not inappropriate but giving a gift of thong underwear is), and 2) also the behaviors had to be judged as having the function to prepare the child to be compliant with sexual abuse. The scenarios were originally developed by the first author and research assistants and revised by the second author; the scenarios were adjusted until 100% agreement between authors was reached regarding whether the scenario described grooming. Behaviors classified as inappropriate were judged through discussion between authors as to whether or not an average person would agree that the behavior was inappropriate (in any way) for an adult to engage in with a child. Two example scenarios are:

1. A thirty-two year old man is on a tropical vacation with his six-year old daughter. She forgot her swimsuit back home so he takes her to the store and buys her a bikini that she picks out. (Correct answer: appropriate behavior; not grooming).
2. A young girl is celebrating her thirteenth birthday with her parents. When her mother leaves the room, her stepfather tells her that he has gotten her a special birthday present. When she opens it, she discovers that he has bought her several bras. (Correct answer: inappropriate behavior, grooming).

For each scenario, participants were asked to indicate 1) whether the behavior in question was "appropriate" or "inappropriate", and 2) if the behavior in question was "grooming", "not grooming", or "unclear."

Online presentations. Participants in the experimental and TAU groups were asked to watch training videos approximately fifteen minutes in length. Both trainings were presented to participants via a YouTube link. The videos featured both visual and auditory presentation of the material, narrated by the first author and two research assistants.
- *Experimental presentation*. The content for the experimental training video was drawn from the PowerPoint presentation utilized by Bennett & O'Donohue (2015). The training included teaching about distinguishing between function and topography when considering a behavior, as well as information about how to use context clues to help determine the function of a behavior. The modules included in the video, in order, were: (1) Introduction to Child Sexual Abuse and Grooming; (2) Basic Facts about Behavior; (3) Determining the Function of Behavior(4) Common Grooming Techniques.
- *Treatment-as-usual presentation* (TAU). The presentation for the TAU group was covered basic facts about child sexual abuse and a basic description of grooming behaviors. The TAU video was designed to represent the type of information that an average individual would find if they searched the internet or consulted books about child sexual abuse or grooming behavior. The TAU presentation did not include information about examining behavior (e.g., function vs. topography).The modules included in the video, in order, were: (1) Introduction to Child Sexual Abuse; (2) Consequences of Child Sexual Abuse; (3) Definitions of Grooming: Examples included, "Grooming is a patterned behavior designed to increase opportunities for sexual assault, minimize victim resistance or withdrawal, and reduce disclosure or belief" (taken from http://www.kbsolutions.com/Grooming.pdf, July 2018); "Befriending and establishing an emotional connection with a child, and sometimes the family, to lower the child's inhibitions with the object of sexual abuse" (taken from https://en.wikipedia.org/wiki/Child_grooming, July 2018); "A tactic of overcoming the survivor's defenses by slowly desensitizing his or her natural reaction to abusive behaviors" (taken from https://www.abuseandrelationships.org/Content/Behaviors/grooming.html, July 2018); (4) Common Grooming Techniques.

Presentation feedback form. Participants were also asked to provide feedback about their respective training. Consistent with other studies of video-based lectures (e.g., Giannakos, Chorianopoulos, & Chrisochoides, 2015), participants were asked to rate the ease of use and utility of information in the presentation on a five-point Likert scale. Additionally, participants were asked (on a 5-point Likert scale) how closely they paid attention to the training video. Finally, participants were asked to provide open-ended responses regarding what they found the most enjoyable or interesting

about the presentations, what suggestions they have for improving the trainings, and any other general feedback about the presentations.

Procedure

After consenting to participate, participants were asked to provide demographic information. Then they were asked to provide an estimate (in minutes) of the amount of time they had spent learning about grooming behaviors prior to the current study. Next, they completed the pre-test questionnaire. Participants were then randomized to one of three conditions: 1) the experimental group, 2) the TAU group, and 3) the control group. Participants were directed to watch their respective online training (or no training at all in the control group). All participants were then asked to complete the post-test assessment and provide presentation feedback. They were thanked for their participation in the study and provided the investigator's contact information.

Data Analysis

All data analyses were completed in SPSS Version 24. Participants' responses to the scenarios were coded as either consistent or inconsistent with the classifications made by the authors (i.e., correct or incorrect). Responses were divided into two total scores; correct identifications on "appropriateness" items and correct identifications on "grooming" items. Participants' correct ratings of appropriate, inappropriate, grooming, and not grooming behaviors were summed and divided by the total number of each respective item in order to calculate an accuracy percentage for each participant on each type of item: 1) Percentage of correct identifications for items keyed as appropriate (percent correct "appropriate"); 2) Percentage of correct identifications for items keyed as inappropriate (percent correct "inappropriate"); 3) Percentage of correct identifications for items keyed as grooming (percent correct "grooming"); and 4) Percentage of correct identifications for items keyed as not grooming (percent correct "not grooming"). These variables were created in order to determine if participants' judgments were more or less accurate depending on the item's content.

RESULTS

Participants ranged in age from 18 to 31 years. Mean age was 19.99 years (*SD*= 2.17). Additional demographic data are displayed in Table 1. Overall, 60.5% of participants indicated that they had no previous history of training about grooming. The mean for the remaining 52 participants was 100.35 minutes of time spent learning about grooming, with a median of 60 minutes (*SD*= 202.77).

Table 1. Demographic Information of Participants.

		Frequency (%)
Gender	Male	30 (20%)
	Female	116 (79%)
Ethnicity	White / Caucasian	82 (56%)
	Latino / Hispanic	25 (17%)
	Black / African-American	6 (4%)
	Native American / American Indian	2 (1%)
	Asian / Pacific Islander	13 (9%)
	Multi-ethnic / Multi-racial	18 (12%)

There were no differences between groups on demographic variables at pre-test, except for gender. The experimental group had a higher proportion of males (N= 16) than the TAU (N= 6) or control (N= 8) groups. Therefore, analyses exploring gender differences only examined performance at pre-test. There was no difference between groups on the time participants had spent learning about grooming prior to the study.

The amount of time participants indicated they spent learning about grooming prior to the training was not correlated with performance on the appropriateness questions on the pre-test, $r(145)$= .01, p= .950, or appropriateness post-test, $r(145)$= .09, p= .312. It was also not correlated to performance on the grooming pre-test, $r(145)$= -.05, p= .559, or the grooming post-test, $r(145)$= .00, p= .998. Therefore, this variable was not utilized as a covariate in subsequent analyses.

Participants' self-report of how well they paid attention to the training video was not correlated with performance on the appropriateness post-test, $r(97)$= -.02, p= .831. However, it correlated with performance on the grooming post-test, with closer attention being related to better scores, $r(97)$= .24, p= .020. This correlation was then examined by group. It was significant in the experimental group, $r(45)$= .33, p= .023, but not in the TAU group, $r(46)$= .18, p= .222, indicating an interaction between the effect of attention on performance and group.

Participants' performance as measured in accuracy percentages were each significantly correlated with each other. Specifically, pre-test performance on "grooming" items was significantly correlated with pre-test performance on "not grooming items", $r(145)$= .22, p= .007; post-test performance on grooming items, $r(145)$= .61, p <.001; and post-test performance on "not grooming" items $r(145)$= .26, p= .001. Pre-test performance on "not grooming" items was also correlated with post-test performance on "grooming" items, $r(145)$= .29, p <.001; and with post-test performance on "not grooming" items, $r(145)$= .65, p <.001. Finally, post-test performance on "grooming" items was also correlated with post-test performance on "not grooming" items, $r(145)$= .25, p= .002.

Participants in the experimental and TAU groups were asked to rate the ease of using the online training, the usefulness of the training, and how closely they paid attention to the training. No significant differences were noted between groups, indicating that the TAU was at least in some ways an acceptable comparison group for the experiment. Results are presented in Table 2.

An independent samples t-test was run to determine if participants in the experimental condition rated the ease of using of the online format differently than participants in the TAU condition. Results showed no significant difference between groups, $t(98)$= .50, p= .616.

An independent samples t-test was run to determine if participants in the experimental condition rated the usefulness of the training differently than participants in the TAU condition. Results showed no significant difference between groups, $t(98)$= -1.77, p= .080.

Table 2. Mean Scores on Presentation Feedback Items.

	Experimental		TAU	
	M	SD	*M*	SD
Ease of Use	4.80	0.539	4.74	0.565
Usefulness	4.20	0.841	4.48	0.707
Attention Paid	3.79	0.806	3.80	0.857

Note: Scores are out of five; higher scores indicate higher endorsement of item.

An independent samples *I*-test was run to determine if participants in the experimental condition reported paying higher or lower amounts of attention to the training than participants in the TAU condition. Results showed no significant difference between groups, $t(98)= -.08$, $p= .940$).

The mean total scores on the appropriateness and grooming assessments, based on group, are presented in Table 3.

To rule out pre-existing differences between groups on test performance, first a one-way ANOVA was utilized. Results showed that there were no differences between groups on total scores on the appropriateness questions, $F(2,144)= .18$, $p= .98$. Additionally, there were no differences between groups on total scores for the grooming questions, $F(2,144)= .85$, $p= .43$. A one-way ANOVA was then run to determine if there was a difference between groups on total appropriateness scores at post-test. Results showed no significant difference, $F(2,144)= .71$, $p= .49$.

Table 3. Mean number of Correct Answers and Total Accuracy Percentages on Pre- and Post-Assessments by Group.

	Experiment (*n*= 49)		TAU (*n*= 50)		Control (*n*= 48)		Total (*N*= 147)	
	M (%)	*SD*	*M* (%)	*SD*	*M* (%)	*SD*	*M* (%)	*SD*
Pre-test Appropriateness	17.80 (89)	1.84	17.72 (89)	2.20	17.75 (89)	1.96	17.76 (89)	2.00
Post-test Appropriateness	18.00 (90)	1.94	17.20 (86)	2.76	17.98 (89)	1.83	17.72 (89)	2.24
Pre-test Grooming Total	13.76 (69)	3.75	13.08 (65)	3.81	14.00 (68)	3.12	13.61 (68)	3.58
Post-test Grooming Total	14.84 (74)	3.37	13.13 (66)	4.12	13.77 (69)	3.71	13.92 (70)	3.79

Because of the interaction between attention and group on post-test performance on the grooming items, a one-way ANCOVA to control for attention effects could not be conducted. Rather, an independent-samples Kruskal-Wallis test was run to determine if a statistically significant difference existed between groups on post-test performance on grooming items. Results showed that there was a significant effect of training type on post-test grooming identification performance ($p= .013$). Post-hoc Mann-Whitney U tests without a Bonferroni correction were used to compare means between the specific groups. Results indicated a significant difference between experimental and TAU groups on grooming questions ($p= .003$) but not on appropriateness questions ($p= .219$). Results showed no difference between experimental and control groups on post-test appropriateness questions ($p= .449$) or grooming questions ($p= .105$). Results also showed no difference between TAU and control groups on post-test appropriateness questions ($p= .779$) or grooming questions ($p= .353$).

Accuracy percentage scores are reported in Table 4.

Participants' ratings of their attention to the training were not correlated with performance on the four post-test accuracy percentage variables so there was no need to control for this in the current analyses. A one-way ANOVA was conducted to determine if a statistically significant difference existed between groups on post-test performance on accuracy percentage of specific items. Results showed that training video type (i.e., experimental, TAU, or control) had a significant effect on the percentage correct of identifying "appropriate" items on the post-test, $F(2, 144)= 6.40$, $p= .002$. Post hoc Tukey tests revealed that the TAU group differed significantly from both the experimental group, $p= .013$, and the control group, $p= .004$. The TAU group on average obtained lower accuracy scores when rating "appropriate" behavior on the post-test. The experimental and control groups did not differ from each other on this measure.

Table 4. *Mean Accuracy Percentage for Each Item Type, by Group*

		M	SD
	Experimental	82.99	14.00
Pre-Test "Appropriate"	TAU	82.00	14.87
	Control	82.64	16.67
	Experimental	93.88	10.23
Pre-Test "Inappropriate"	TAU	94.00	12.40
	Control	93.75	10.63
	Experimental	68.78	24.72
Pre-Test "Grooming"	TAU	65.60	29.15
	Control	68.54	25.09
	Experimental	68.78	23.33
Pre-Test "Not Grooming"	TAU	62.60	25.38
	Control	68.54	21.14
	Experimental	89.24	13.84
Post-Test "Appropriate"	TAU	79.45	21.17
	Control	90.72	14.94
	Experimental	90.93	11.49
Post-Test "Inappropriate"	TAU	94.00	13.13
	Control	88.89	13.36
	Experimental	76.87	20.01
Post-Test "Grooming"	TAU	72.44	31.45
	Control	64.58	28.77
	Experimental	78.16	22.56
Post-Test "Not Grooming"	TAU	60.40	28.74
	Control	78.96	18.84

ANOVA results also indicated that the effect of training type was significant on the accuracy percentage for "not grooming" items on the post-test, $F(2, 144)= 7.72$, $p <.001$. Post hoc Tukey tests showed that the TAU group differed significantly from the experimental group, $p= .003$, and the control group, $p= .002$. Again, the experimental and control groups did not differ from one another on this measure.

Further analysis of the TAU group's performance on the "non-grooming" items examined the types of errors that participants made. Participants were more likely to incorrectly choose "Unclear" as their response rather than incorrectly choosing "Grooming." Out of a total of 407 incorrect responses, participants chose "Unclear" 67% of the time.

Gender was found to be significantly correlated with pre-test performance on rating the appropriateness of a behavior, $r(145)= .21$, $p= .011$. When looking at performance on specific item type, gender was found to be significantly correlated with pre-test performance on identifying inappropriate behavior, $r(145)= 0.33$, $p <.001$, and grooming behavior, $r(145)= .17$, $p= .044$. Mean scores for these variables are reported in Table 5.

Due to the large difference in sample size and relatively small sample of males ($N= 30$), and due to the data violating normality assumptions, analyses were conducted using Mann-U Whitney tests. Results showed a significant difference between males and females on pre-test measures on total correct appropriateness questions ($p= .031$),

Table 5. Comparison in Performance between Males and Females on Pre-Test Items.

	Male		Female	
	M	SD	M	SD
Appropriateness Total Score	17.07	2.50	18.02	1.57
% Correct "Inappropriate"	9.67	1.63	10.55	0.84
% Correct "Grooming"	5.97	2.58	7.03	2.54

percent correct "inappropriate" items (*p*= .001), and percent correct "grooming" items (*p*= .027). Females significantly outperformed males on each of the three measures.

Further analysis of the male group's performance on the grooming items examined the types of errors that male participants made. Male participants were slightly more likely to incorrectly choose "Unclear" as their response rather than incorrectly choosing "Not Grooming." Out of a total of 121 incorrect responses, participants chose "Unclear" 56% of the time.

Discussion

The results provide interesting insights into the complexity of individuals' judgments about potential grooming behaviors. Overall participant accuracy at judging grooming behaviors at post-test fell between 60 and 70 percent, indicating that participants even after training are incorrectly categorizing behavior nearly one-third of the time. Seen another way, this performance is only 10 to 20 percent above chance levels. Additionally, it is important to note that approximately two-thirds of participants indicated that they had never learned about grooming prior to their involvement in the current study. There appears to be a gap in education of lay people (or in this case, undergraduate psychology students) about grooming. Finally, other discovered factors (e.g., the gender effect of better performance of females) indicated that judgment of an already complicated construct is even more complex than previously thought as gender seems to play an important role.

While the main goal of the current research (i.e., to significantly increase participants' ability to accurately categorize behaviors that may indicate grooming) was not entirely achieved, this project contributed valuable information about the baseline ability level of individuals to make judgments about grooming, as well as how this ability may be impacted when individuals utilize the internet to research this topic. Participants tended to perform better when determining whether a behavior was appropriate; they had an accuracy rate of 89% on appropriateness items compared to 68% on grooming items. These results imply that individuals are much more able to accurately label a behavior as appropriate or inappropriate, but are less able to accurately categorize between grooming and not grooming. Thus, the difficulty for individuals seems to be in judging the function of the behavior -even when it is judged as inappropriate.

The primary hypothesis was that participants who received the experimental training would outperform the TAU and the control group participants at post-test, meaning that they would be more accurate judges of grooming behavior. The results showed that the participants who received the experimental training outperformed the group who received the TAU training on judging whether a behavior was grooming or not, but displayed no significant differences in accuracy rates when compared to the control group. Thus, this suggests that gaining information from the internet can lead to decreases in the ability to correctly identify grooming behavior as compared to doing nothing.

Further analysis revealed that at post-assessment, the TAU group was more likely to 1) judge an appropriate behavior as being inappropriate, and 2) judge a "not grooming" behavior as "unclear" (and less often, incorrectly naming "not grooming" behaviors as "grooming"). These findings imply that when participants are provided information that an individual might find online about child sexual abuse and grooming, they are less able to accurately identify appropriate behavior, and they tend to view appropriate behavior as more ambiguous (i.e., they are more suspicious of appropriate behavior). Thus, overall this leads to an increased rate of false positives. It is possible that this

decline in performance is due to the argument put forth by Bennett and O'Donohue (2014); that is, that the current definitions of grooming behavior are varied, and the conceptual fuzziness of the construct causes confusion for individuals tasked with judging the behavior. Taken into a larger context, this finding implies that individuals exposed to a representation of the material they find on the internet about grooming, without training about understanding behavior in context, may be more likely to become overly suspicious and thus potentially make false accusations of inappropriate behavior. Importantly, individuals who are curious about grooming and search the internet to learn more about it may end up more confused and worse at identifying such behavior than had they not looked at all.

There were no differences when comparing the participants' percentage of correct responses when identifying "grooming" vs. "not grooming" behaviors in the pre-test nor in the post-test. Participants were just as likely to make errors when judging both "grooming" and "not grooming" behaviors, implying no real difference between the potential rates of false negatives and false positives. Instead, participants' accuracy percentages on each type of item (i.e., when assessing grooming and not grooming behaviors in pre- and post-assessments), were highly correlated with one another. This finding is more indicative that participants each possess an overall ability to judge behavior, regardless of whether it is grooming or not grooming behavior; rather than one type of behavior (i.e., grooming or not grooming) being easier to judge than the other.

Participants' report of the attention they paid to the training video correlated with post-test performance, but interestingly only in the experimental group. Attention to online training videos is likely always a concern of this form of dissemination. This problem with participant attention may be reduced by altering the training so that participants must click through each slide, rather than having a video that plays automatically.

The finding that males tended to perform worse, on average, than females at identifying the appropriateness of a behavior and whether it was grooming was important. More specifically, males were more likely to characterize inappropriate behaviors as appropriate, and actual grooming behaviors as unclear. One potential hypothesis for why females may have outperformed males is that females on average tend to spend more time in roles involving childcare (Dush, Yavorsky, & Schoppe-Sullivan, 2018) and therefore may have a better developed repertoire for detecting inappropriate behaviors. Relevant to our sample, it is possible that female undergraduates may be overrepresented in majors such as early childhood education and thus gain correct information in this way. Also females typically have the role of guarding and evaluating sexual approach behavior and thus may be better at detecting ulterior motivations. Additionally, males are much more likely to be perpetrators of CSA than females (Peter, 2009) and an argument could be made that therefore males may be more unwilling or unable to view such behavior as inappropriate.

These findings, although requiring further replication and study, may indicate a larger implication for identification of grooming behavior in the courtroom and in natural settings. For example, a male juror may be less likely to judge an actual grooming behavior as inappropriate or as indicative of grooming and therefore could lead to false negatives (i.e., letting potential perpetrators "get away" with such behavior). Additionally, a male may be less likely to perceive grooming behavior occurring in real time as inappropriate and thus may not act in a way to protect the child.

While the experimental training in its current form does not appear to solve the problem of teaching individuals to more accurately label grooming behavior, it

represents an important first step toward that goal. Furthermore, the participants provided helpful feedback regarding the trainings and scenarios. Participants generally rated the online format of the training videos as easy to use and as useful. The format of using an online presentation was chosen particularly because it could be low-cost and easily disseminated should it be found effective. With regard to the experimental training, participants reported that some of the most useful parts of the video were instruction about using context clues, information about function vs. topography, and specific examples of grooming behaviors. Taken together, this feedback overall provides support for what was originally thought to be important to help individuals make accurate judgments about deviant behavior, i.e., getting the participant to focus on the function and context of the behavior rather than just the topography itself. Many participants noted that the presentation of information about particular types of grooming behaviors (i.e., topographies) was still a highly useful portion of the training as well. Ultimately, the training should continue to include presentation of statistics about commonly used grooming topographies to assist participants in knowing the base rates of discriminative stimuli present in the environment.

With regard to areas for improvement, several participants indicated that the training could be revised to be more engaging, possibly by using more images and different types of video clips. Overall, the training may be enhanced by rerecording some of the audio, adding images and videos, and potentially making the training more interactive to be more engaging to the listener. Future iterations of the video may also want to include presentation of material including both male and female narrators as to investigate any potential bias in the learning experience. Furthermore, a more advanced online training format could be used to provide individualized feedback to participants about their performance on items in order to assist their learning. Specifically, the training could also be tailored by taking into account the individual's baseline judgment error tendencies, e.g., if they are more likely to score false alarms, then the training can focus on training them to reduce bias toward choosing grooming when the behavior is not indicative of it.

This point about feedback is consistent with the behavioral paradigm of multiple exemplar training, which has been shown to be highly effective discrimination training between two stimuli (e.g., Marzullo-Kerth, Reeve, Reeve, & Townsend, 2011), which in this case could help participants differentiate between grooming and not grooming. Stimuli belonging to especially complicated or abstract categories (e.g., grooming) may require a more elaborately designed training which includes exemplars that vary along a grade with respect to their multiple dimensions; this is referred to as concept learning (Holth, 2017). With regard to grooming, this type of training would include presentation of multiple examples of grooming along with specific descriptions of their contextual properties. Participants would be required to make judgments about overall behaviors as well as each individual contextual property; they would also receive immediate feedback from the training program to.

It will also be important for future researchers to consider cultural context both in the actual judgment of grooming behaviors within a specific cultural group as well as the assessment of such judgment. It is possible that many of the grooming behaviors discussed (especially grooming the child's environment) could potentially go under-noticed in cultural groups that are more collectivist and dependent on each other in the care of their children. Research indicates that lower rates of CSA are reported among collectivist cultures, including Asian cultures; it is hypothesized that this finding may

be due to underreporting rather than fewer incidents of abuse occurring (Stoltenborgh, Van Ijzendoorn, Euser, & Bakermans-Kranenburg, 2011). Furthermore, arguments could be made that ideas regarding what is considered acceptable touching varies by cultural group (e.g., a parent kissing their older child on the lips). It would be interesting to examine if certain types of grooming tactics are more or less common among certain cultural groups as well as whether or not such individuals have a more difficult time distinguishing deviant behavior.

Importantly, further development of the scenarios is necessary in order to determine a competency cutoff for the instrument. It has been suggested that proper development and standardization of an assessment tool should be achieved prior to determining a cutoff score to indicate competency on knowledge tests used for training or certification (Biddle, 1993). It would also be interesting for future research to examine the effect of excluding the option of "Unclear" on the grooming assessment. A forced choice between "grooming" and "not grooming" would impact participants' accuracy rates and should be considered. Signal detection analyses could be performed on data collected using this two-choice paradigm which would provide rich information about participants' abilities to discriminate between grooming and non-grooming behaviors.

One of the limitations in this study is the need for further development of the hypothetical grooming scenarios for an assessment of individuals' judgment of grooming. The scenarios used in the study were arguably face valid. However, continued efforts into designing a demonstrably valid assessment of grooming would be useful for efforts regarding the effectiveness of training methods.

An additional research question exists in that it is possible that judgment of grooming behaviors could be improved if participants were allowed to discuss decisions with others rather than having to rely on individual judgment. It could be argued that individuals are more able to come to an accurate determination when they are able to bounce their opinions off one another. This type of paradigm is also more naturalistic as in real life, clinicians can consult with one another. Taken as a whole, these findings show that further research as well as education about grooming are necessary in order to enhance individuals' abilities to distinguish such behavior from appropriate adult-child behavior.

References

Bennett NM & O'Donohue WT (2014). The construct of grooming in child sexual abuse: Conceptual and measurement issues. *Journal of Child Sexual Abuse, 3*, 957-76. Doi: 10.1080/10538712.2014.960632

Bennett NM & O'Donohue WT (2015). *Developing a training for the correct identification of sexual grooming behaviors*. University of Nevada, Reno. Unpublished manuscript.

Biddle RE (1993). How to set cutoff scores for knowledge tests used in promotion, training, certification, and licensing. *Public Personnel Management, 22*, 63-79. Doi: 10.1177/009102609302200105

Brackenridge CH (2001). *Spoilsports: Understanding and preventing sexual exploitation in sport*. London: Routledge.

Budin IE & Johnson CF (1989). Sexual abuse prevention: Offenders' attitudes about their efficacy. *Child Abuse & Neglect, 13*, 77-87. Doi: 0.1016/0145-2134(89)90031-8

Burgess AW & Hartman CR (2018). On the origin of grooming. *Journal of Interpersonal Violence, 33*, 17-23. Doi: 10.1177/0886260517742048

Colton M, Roberts S, & Vanstone M (2012). Learning lessons from men who have sexually abused children. *The Howard Journal of Criminal Justice, 51*, 79-93. Doi: 10.1111/j.1468-2311.2011.00682.x

Conte JR, Wolf S, & Smith T (1989). What sexual offenders tell us about prevention strategies. *Child Abuse and Neglect, 13*, 293-301. Doi: 10.1016/0145-2134(89)90016-1

Dush CMK, Yavorsky JE, & Schoppe-Sullivan SJ (2018). What are men doing while women perform extra unpaid labor? Leisure and specialization at the transitions to parenthood. *Sex Roles*, *78*, 715-730. Doi: 10.1007/s11199-017-0841-0

Gillespie A (2002). Child protection on the internet: Challenges for criminal law. *Child and Family Law Quarterly, 14*, 411-425.

Elliott M, Browne K, & Kilcoyne J (1995). Child sexual abuse prevention: What offenders tell us. *Child Abuse and Neglect, 19,* 579-594. Doi: 10.1016/0145-2134(95)00017-3

Marzullo-Kerth D, Reeve SA, Reeve KF, & Townsend DB (2011). Using multiple-exemplar training to teach a generalized repertoire of sharing to children with autism. Journal of Applied Behavior Analysis, 44(2), 279-294. Doi: 10.1901/jaba.2011.44-279

O'Connell R (2003). A typology of child cybersexploitation and online grooming practices. *Cyberspace Research Unit, University of Central Lancashire*. Available at: http://www.jisc.ac.uk/uploaded_documents/lis_PaperJPrice.pdf

Ost S (2004). Getting to grips with sexual grooming? The new offence under the Sexual Offences Act 2003. *Journal of Social Welfare and Family Law, 26*, 147-159. Doi: 10.1080/014180304200023088

Peter T (2009). Exploring taboos: Comparing male-and female-perpetrated child sexual abuse. *Journal of Interpersonal Violence*, *24*, 1111-1128. Doi: 10.1177/0886260508322194

Salter A (1995). *Transforming trauma: A guide to understanding and treating adult survivors of child sexual abuse*. Newbury Park, CA: SAGE.

Spiegel J (2003). *Sexual abuse of males: The SAM model of theory and practice*. New York: Brunner-Routledge.

Van Dam C (2001). *Identifying child molesters: Preventing child sexual abuse by recognizing the patterns of the offenders*. Binghamton, NY: Haworth Maltreatment and Trauma Press/The Haworth Press, Inc.

Winters GM & Jeglic EL (2016). I knew it all along: The sexual grooming behaviors of child molesters and the hindsight bias. *Journal of Child Sexual Abuse, 25*, 20-36. Doi:10.1080/10538712.2015.1108945

Winters GM & Jeglic EL (2017). Stages of sexual grooming: Recognizing potentially predatory behaviors of child molesters. *Deviant Behavior, 38*, 724-733. Doi:10.1080/01639625.2016.1197656

Wolf MR & Pruitt DK (2019) Grooming hurts too: The effects of types of perpetrator grooming on trauma symptoms in adult survivors of child sexual abuse. *Journal of Child Sexual Abuse, 28*, 345-359, Doi: 10.1080/10538712.2019.1579292